LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

LONDON SCHOOL OF HYGIENE AND TROPICAL MEDICINE

MSc PROJECT REPORT
2017-2018

The time and quantity patterns of carbohydrate intake in the UK adults – the National Diet and Nutrition Survey 2008/09-15/16 programme

---

Supervisor:
Professor Luigi PALLA

**Submitted in part fulfillment of the requirements
for the degree of MSc in Medical Statistics**

Candidate number: 110765
Page count: 42 from Introduction to Conclusions

September 2018

Declaration of Authorship

I, Chaochen WANG, declare that this thesis titled, "The time and quantity patterns of carbohydrate intake in the UK adults – the National Diet and Nutrition Survey 2008/09-15/16 programme" and the work presented in it are my own. I confirm that:

- This work was done wholly while in candidature for a MSc degree in Medical Statistics at London School of Hygiene and Tropical Medicine.

- No part of this thesis has previously been submitted for a degree or any other qualification at London School of Hygiene and Tropical Medicine or any other institution.

- Where I have consulted the published work of others, this is always clearly attributed.

- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

- I have acknowledged all main sources of help.

- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date: 2018-09-05

Acknowledgements

"All models are wrong, but some are useful."

George E. P. Box

Abstract

The National Diet and Nutrition Survey (NDNS) database of detailed four-day food diaries was used in exploration of both temporal and quantity eating patterns of carbohydrates in the general adults living in the UK. 6155 participants aged 19 or older were kept in the analysis.

Multilevel latent class analysis (MLCA) models were applied in finding the latent classes of carbohydrate consumption both at the observation day level, and individual level.

Contents

List of Figures

List of Tables

List of Abbreviations

AIC	**A**kaike **I**information **C**riterion
A1c	Haemoglobin **A1c**: Glycated hemoglobin
BMI	**B**ody **M**ass Index
BIC	**B**ayesian **I**information **C**riterion
CI	**C**onfidence **I**nterval
DM	**D**iabetes **M**ellitus
FSA	**F**ood **S**tandards **A**gency
HDL	**H**igh **D**ensity **L**ipoprotein cholesterol
LDL	**L**ow **D**ensity **L**ipoprotein cholesterol
LCA	**L**atent **C**lass **A**nalysis
LCGA	**L**atent **C**lass **G**rowth **A**nalysis
LTA	**L**atent **T**ransition **A**nalysis
MAFF	**M**inistry of **A**griculture, **F**isheries and **F**ood
MLCA	**M**ultilevel **L**atent **C**lass **A**nalysis
ML	**M**aximum **L**ikelihood
NDNS	the **N**ational **D**iet and **N**utrition **S**urvey
OR	**O**dds **R**atio
PHE	**P**ublic **H**ealth **E**ngland
PSUs	**P**rimary **S**ampling **U**nits
RMLCA	**R**epeated **Measures LCA**
TC	**T**otal **C**holesterol
TG	**T**ri**g**lycerides
WC	**W**aist **C**ircumference

Chapter 1

Introduction

Background

The widely accepted standard these days seems to be that we eat three times a day. However, whether this is really an ideal temporal eating pattern for everyone has never been answered with scientific evidence. More importantly, how many temporal patterns of eating are there in the population, proportions of people who actually manage/fail to follow this doctrine, are people consistently following one specific temporal eating pattern or do they switch, have not been studied and described thoroughly either.

Although nutritional studies have extensively examined the influence of the quantity and quality of dietary and nutrients intake and their alteration on morbidity and mortality, investigations on temporal eating patterns and their effects are still scarce. The importance of the circadian rhythm in regulating physiological responses has been recognised for long, while the impact of which on nutrition and metabolism is still largely unknown (Johnston, 2014; Garaulet and Gómez-Abellán, 2014; Asher and Sassone-Corsi, 2015). Some recent evidence have found that meal timing is associated with a wide variety of health outcomes. Skipping breakfast is associated with higher risk of type 2 diabetes (Uemura et al., 2015). Shift workers have a higher risk of developing metabolic syndrome (De Bacquer et al., 2009) and type 2 diabetes (Pan et al., 2011). Evening intake of energy is positively associated with overweight/obesity (Almoosawi et al., 2016).

More recently, discernible temporal eating patterns that differed by sociodemographic and eating profiles were revealed by latent class analysis using nutrition survey data (Leech et al., 2017; Mansukhani and Palla, 2018). Based on total energy consumption, the presence of 3 groups of eaters: grazers, early eaters, and late eaters were identified. So far, the temporal eating patterns were only based on averaging the total energy intake calculated from dietary recall, and therefore could

not capture the day-to-day variation in temporal eating patterns. Thus, the question of how much variability of a person follows one or several specific temporal eating patterns in his/her everyday life remains unanswered. Many factors, such as day of the week or season, or culture may contribute to daily variation in dietary intake, however, most of the variation in an individual's diet may seems to be without an obvious pattern. Intakes of macro-nutrients (carbohydrate, fat, and protein), due to the reason of their large contribution to the total energy intake, may have somewhat moderate degrees of day-to-day variation (Willett, 2012). Novel analytic methods that can account for this within person day-to-day variation is needed.

In the current report, we focused on temporal eating patterns for carbohydrate consumption in the UK adults. Eating more carbohydrate in the morning has been found to be negatively associated with metabolic syndrome (Almoosawi et al., 2013). On the other hand, high total consumption of carbohydrate has been linked with higher risk of type 2 diabetes (Alhazmi et al., 2012). Whether the amount or the timing (or both) of carbohydrate consumption during the day actually matters, would be the question concerned by the public.

The National Diet and Nutrition Survey (NDNS)

The National Diet and Nutrition Survey (NDNS) programme (NatCen Social Research, 2018) was initially established in 1992 and started off as a joint initiative between the Ministry of Agriculture, Fisheries and Food (MAFF) and the Department of Health. In 2008, a new continuous cross-sectional survey was started, the NDNS Rolling Programme (RP). The NDNS RP is funded by Public Health England (PHE), an executive agency of the Department of Health, and the UK Food Standards Agency (FSA). The survey covers a representative sample of around 1000 people per year. Fieldwork began in 2008 and is now beginning its 11th year. NDNS provides essential evidence on the diet and nutrition of the UK population to enable PHE to identify and address nutritional issues in the population and monitor progress towards public health nutrition objectives.

The NDNS RP has now completed and analysed its eighth year. The sample was randomly drawn from a list of all the addresses, clustered into postcode sectors from across the UK. Overall, for years 1-8 combined, a sample of 39,300 addresses was selected from 799 (year 1-4), 323 (year 5-6), and 316 (year 7-8) postcode sectors. At each address, one household was selected at random (in cases where there were two or more households). For each household, either an adult and a child, or a child only, was selected to participate.

These individuals were asked to keep a four-day diary on their food and drink consumption on consecutive days. An interview and a nurse visit were also conducted to collect information regarding height and weight, smoking and drinking habits, physical activity, blood pressure, prescribed medicines, dietary supplements, fasting blood sample, and 24-hour urine sample.

Aims and objectives

Our goal is to explore and make use of the NDNS RP (2008/09-15/16) database to describe and identify the potential relationship between the timing of eating within the day and specific nutrient–carbohydrate intake. We aimed at finding time patterns of consumption of carbohydrate and defining latent groups in the UK adults. Subsequently, an additional potential aim, is to investigate the association between carbohydrate eating patterns with hypertension and obesity.

Chapter 2

Methods

Dietary diary collected in the NDNS RP

Participants were asked to keep a record of everything eaten or drunk over four consecutive days. Interviewers undertook three visits with each participant. At the first visit, the interviewer explained the method followed a protocol, taking participants through the sections in the diary including how to describe details of food and drink and portion size and an example day. The second was a brief visit to check for compliance, answer questions or deal with problems and review the diary to identify and edit possible omissions and missing detail. The third visit was to collect the diary and again review and edit possible omissions.

In the diary, participants were asked to record portion sizes in household measures (e.g. one tablespoon of beans, one Kit Kat finger-size), or for packaged foods to note the weight indicated on the packet. For homemade dishes, participants were asked to record on a separate page in the diary the individual ingredients and quantities for the whole dish along with a brief description of the cooking method and how much of dish they had consumed. In each eating occasion, in addition to the details of what and how much was eaten, participants were also asked to record: when was it, where they were, and who they were eating with. An example, used as guidance for participants, of a food diary for one day is shown in **Appendix E**.

Definition of carbohydrate intake

Detailed dairy checking was performed to code and convert the food consumption into energy and nutrients intake. Intakes of nutrients were calculated from the food consumption records using a specially adapted Nutrient Databank (Smithers, 1993), which was originally developed by the Ministry of Agriculture, Fisheries and Food (MAFF) for the Dietary and Nutritional Survey of British Adults. Further details of data coding and editing are outlined in Appendix A of the NDNS official reports

(Department of Health, 2018). Specifically, the main variables that we adopted in the current analysis were defined as:

- Total energy intake = (protein(gram) $\times$ 17) + (fat(gram) $\times$ 37) + (carbohydrate(gram) $\times$ 16) + (alcohol(gram) $\times$ 29) kJ;

- Carbohydrate intake = total sugars (gram) + starch (gram);

Time across a typical survey day was divided into 7 time slots in the dietary diary of NDNS RP: 6 am to 9 am, 9 am to 12 noon, 12 noon to 2 pm, 2 pm to 5 pm, 5 pm to 8 pm, 8 pm to 10 pm, and 10 pm to 6 am next morning. To produce a sequence of discrete responses regarding the carbohydrate intake we are interested, the energy consumption within each time slot over the four days of survey for each participant were calculated. The percentages of energy that contributed by carbohydrate within each time slot were then estimated. Since we planed to apply latent class analysis (LCA) in the current study, in which the observed indicators for latent classes must be categorical, we then dichotomised the responses according to the carbohydrate contribution to the energy intake at cut-off value of 50%, i.e. if within a time slot where there is any energy intake occured, carbohydrate consumption was categorised into whether it's energy contribution was lower or higher/equal to 50% of total energy intake within that time slot. Consequently, for each day of the recording, there were 7 data points generated by the diary, each data point included one of the following responses:

- Not eating any food (energy intake = 0 kJ);

- Eating, and carbohydrate contributed less than 50% of the total energy intake;

- Eating, and carbohydrate contributed higher or equal to 50% of the total energy intake.

Survey Data

Survey selection method

The NDNS RP participants were drawn from the UK Postcode Address File, a list of all the addresses in the UK. The addresses were clustered into Primary Sampling Units (PSUs), small geographical areas, based on postcode sectors, randomly selected from across the UK. A list of 27 or 28 addresses was then randomly selected from each PSU.

Overall, for years 1 to 8 combined, a sample of 39,300 addresses was selected from 1,438 PSUs. The sampling selection process was:

- Randomly select PSUs from the Postcode Address File;

- Randomly select 27 or 28 addresses in that postcode area;

- Randomly select one household at that address;

- Selected addresses were randomly allocated to one of two groups to determine whether an adult (aged 19 years or older) and a child (aged 1.5 to 18 years), or a child only, were selected for interview.

Response rates

The response rates for completion of the dietary diary (three or four days) were 56%, 53%, 53%, for years 1 to 4, 5 to 6, and 7 to 8, respectively. A total of 6,155 adults aged 19 years and older were kept in the current study.

Strata and weightings

It is necessary to apply weighting factors to the data collected in the NDNS RP for two reasons: to remove any bias in the observed results which may be due to differences in the probability of households and individuals being selected to take part; and to attempt to reduce differential non-response bias by age, sex and geographical region.

The strata that used to calibrate proportions in the sample include: age-group (1.5-3, 4-6, 7-10, 11-15, 16-18, 19-24, 25-29, 30-39, 40-49, 50-59, 60-64, 65-69, and over 70 years); sex (men or women); and regions (Northern Ireland, Scotland, Wales, and the nine regions of England).

Two steps of weighting system are designed in the NDNS RP to assure that the combined sample will be representative of the UK population:

1. An overall selection weight, which is the product of the address, dwelling unit, catering (household) unit, and individual selection weights, was generated to correct for the unequal selection probabilities. These weights are the inverse of the selection probabilities at each level of the random sampling process and they can be used to compensate for differences in the chance of selection of an individual.

2. An iterative procedure was used to adjust the selection weights until the distribution of the weighted sample matched that of the population for age-group, sex and geographical region. Population distributions were taken from the mid-year population estimates (Office for National Statistics, 2018).

Another two sets of weights were generated to correct for differential non-response (either due to refusal or inability) to 1) nurse visit, and 2) giving a blood sample. Response rates to the nurse visit among those completed a dietary diary was approximately 75%, to blood sample in adults were 51%, 57%, and 50% for years 1 to 4, 5 to 6, and 7 to 8, respectively. In creating the nurse/blood sample weight, a logistic regression model was used by the NDNS RP study team to model the relationship between response to nurse visit/giving blood sample and a set of predictor variables (socio-demographic, participant and catering/household unit characteristics). The model generated a predicted probability for each participant, which is the probability would agree to a nurse visit/provide blood sample, given the characteristics of the individual and the household unit. Participants with characteristics associated with non-response were under-represented in the sample and therefore receive a low predicted probability. The inverse of these predicted probabilities were used as a set of non-response weights so that participants with a low predicted probability got a larger weight, increasing their representation in the sample. Then the nurse/blood sample weights were re-scaled so that the sum of the weights equalled the number of participants who had a nurse visit/who provided a blood sample. The final nurse/blood weights should therefore make the sample participants representative of all eligible persons in the population.

Further details of the weighting system developed by the NDNS RP are described in the Appendix B of the reports published by Public Health England (PHE) (Bates et al., 2014; Roberts et al., 2018; Department of Health, 2018).

Socio-demographic status, lifestyle, physical activity, anthropometric measurements and biochemical analyses

Computer assisted personal interviews were conducted for the selected individuals by trained interviewers to collect background information on smoking habits (current, ex-smokers, and never), ethnicity (white, non-white), education level (lower than degree/degree or above level), living with a partner or not and other socio-demographic related variables. Participants also had their height, weight, blood pressure, waist circumferences (WC) measured by the nurses.

Specifically, blood pressure was measured in a sitting position using an automated, validated machine, the Omron HEM907, after a five minute rest. The means of second and third readings, taken at one minute intervals, were used in the current report. Hypertension was defined as with systolic blood pressure of 140 mmHg or above, and/or diastolic blood pressure of 90 mmHg or above, and/or taking any medication specifically to reduce blood pressure.

A self-completion questionnaire - the Recent Physical Activity Questionnaire (Besson et al., 2009) (RPAQ, developed by the MRC Epidemiology Unit Cambridge) was used to estimate physical activity from year 2 of the survey. The RPAQ was designed to assess usual physical activity in the last month in four domains: home, work, commuting to work, and leisure activities. Detailed descriptions of the assessment of adult physical activity in the NDNS RP and the processing of data are available in Appendices G and V of the published reports (Bates et al., 2014; Roberts et al., 2018; Department of Health, 2018).

Blood samples were stored at 4 °C, and sent directly by post to the Department of Haematology and Department of Clinical Biochemistry and Immunology, Addenbrooke's Hospital, Cambridge within two hours of their collection. Serum samples were obtained by centrifugation of the coagulated blood sample. Serum total, High Density Lipoprotein (HDL) and Low Density Lipoprotein (LDL) cholesterol, triglycerides (TG), fasting blood glucose, haemoglobin A1C were measured. A1C value of 6.5% was used as the cut off point for diagnosing diabetes.

Body mass index (BMI) was calculated as weight in kilograms divided by height in square meters. BMI was then categorised into less than 25 kg/m^2 (normal weight), 25 to 30 kg/m^2 (overweight), and higher or equal to 30 kg/m^2 (obese).

Ethical approval

Ethical approval for the survey was obtained form the Oxfordshire A Research Ethics Committee. The letters of approval for the original submission and subsequent substantial amendments, together with approved documents, were sent to all Local Research Ethics Committees covering areas where fieldwork was being conducted. Research governance approval was sought for all participating NHS laboratories and obtained where required by the Research and Development Committee for each laboratory. Ethical approval for the current project was obtained from the MSc Research Ethics Committee of London School of Hygiene & Tropical Medicine (LSHTM MSc Ethics Ref: 15624).

Statistical methods

Latent Class Analysis (LCA)

Latent class analysis is a statistical technique that identifies categorical latent (unobserved) class variables on the basis of observed categorical variables (Collins and

Lanza, 2010). It belongs to the family of latent variable models, and is directly analogous to the factor analysis model. The major difference is that the latent variable in LCA is categorical, not continuous as in factor analysis. The basic assumptions in LCA are independent observations, and local independence, as shown in the fundamental expression of a typical LCA model:

$$P(U_{i1} = s_1, U_{i2} = s_2, \cdots, U_{ik} = s_K) = \sum_{t=t}^{T} P(C_i = t) \prod_{k=1}^{K} P(U_{ik} = s_k | C_i = t) \quad (2.1)$$

Where,

- $P(U_{i1} = s_1, U_{i2} = s_2, \cdots, U_{ik} = s_k)$ is the probability of observing a particular vector of responses;

- $P(C_i = t)$ is the probability that a randomly selected ith observation will be in class t;

- $P(U_{ik} = s_k | C_i = t)$ is the probability of a particular observed response pattern $U_{ik} = s_k$ conditional on membership in latent class t.

Equation 2.1 indicates that responses for an observation to the measuring variables are independent of one another given its membership in latent class t. However, in the NDNS RP data set, the assumption of independent observations is violated. Each individual completed their dietary diary for four consecutive days, their diary recordings were later converted into four sequences of categorical responses reflecting the type of carbohydrate consumption at each time slot of the day. The four observed sequences (observed days) are nested in the participants and therefore are not independent. This nested data structure requires multilevel techniques.

Multilevel Latent Class Analysis (MLCA)

Multilevel latent class analysis accounts for the nested structure of the data by allowing latent class intercepts to vary across level 2 units and thereby examining if and how level 2 units influence the level 1 latent classes. These random intercepts allow the probability of membership in a particular level 1 (observation days) latent class to vary across level 2 units (e.g., here in the current context are the individuals). Essentially this allows the probability that an observation day will belong to a particular day level latent class to vary across individual levels.

Parametric approach

Proposed by Vermunt (Vermunt, 2003; Vermunt, 2008) and Asparouhov and Muthén (Muthén and Asparouhov, 2009), a traditional, parametric approach can

be applied using a logistic regression model. For example, let's assume that there are two types of observation days in the dietary survey–high and low carbohydrate eating days. In an unconditional logistic regression model, the probability of the outcome (i.e. an observed high carbohydrate eating day vs. a low carbohydrate eating day) is constant within individual level, which means for each person throughout his/her survey there is some probability of following a high carbohydrate eating day. A random effect considers the individuals (level 2) to be drawn from the adult population in the UK, and the probability of the outcome (i.e. high carbohydrate eating days) across individuals is considered to be a random variable (Snijders and Bosker, 2011).

Thus, for a binary outcome $C_{ij} = 0, 1$ (low $= 0$ or high $= 1$ carbohydrate eating days), where i denotes the observation days ($i = 1, 2, 3, 4$), and j denotes the individual ($j = 1, 2, \cdots, 6155$). The 2-level random intercept logistic regression model can be expressed as:

$$
\begin{aligned}
\text{logit}[P(C_{ij} = 1)] &= \beta_{0j} + \beta_1 x_{ij} \qquad \textbf{(day level)} \\
\beta_{0j} &= \gamma_0 + \gamma_1 w_j + u_{0j} \textbf{ (individual level)} \\
\Rightarrow P(C_{ij} = 1) &= \frac{\exp\left(\gamma_0 + \beta_1 x_{ij} + \gamma_1 w_j + u_{0j}\right)}{1 + \exp\left(\gamma_0 + \beta_1 x_{ij} + \gamma_1 w_j + u_{0j}\right)}
\end{aligned}
\tag{2.2}
$$

Where we define:

- $\text{logit}(x) = \frac{\log(x)}{\log(1-x)}$;

- $P(C_{ij} = 1)$ as the probability that the randomly selected ith observation day of jth individual is a high carbohydrate eating day;

- β_{0j} as the random intercept, for outcome $C_{ij} = 1$;

- the random deviation of the individuals u_{0j} are assumed be normally distributed (i.e. $u_{0j} \sim N(0, \sigma_{u_0}^2)$), the magnitude of the u_{0j} variance ($\sigma_{u_0}^2$) indicates the influence of the individuals (level 2);

- x_{ij}, w_j is the predictors for day level (weekdays or weekends) and individual level, such as age, and sex.

Same framework can be used to consider random effects in an LCA model, but instead of saying that C_{ij} is either low or high carbohydrate eating days as if we already know, it is now replaced by a latent variable G_{ij} which indicates the typologies of carbohydrate eating patterns. Then we can use the day level data to assess the log-odds of belonging to kth type of carbohydrate eating pattern on a specific day of survey, and we allow the log-odds to vary across individuals.

Therefore, for some persons the log-odds of having a *k*th type of carbohydrate eating pattern during the survey can be high, but for the other persons, the log-odds of following the *k*th type of carbohydrate eating pattern can be low.

If the day level LCA model (carbohydrate eating temporal pattern typologies) is best defined by $T(T \geqslant 2)$ latent classes, then $T - 1$ random intercept will be specified by a two-level multinomial logistic regression model. Similar to the typical LCA models, the latent class variable in a MLCA is defined by multiple observed indicators (here is defined by the responses of eating carbohydrate within each time slots, throughout 4 consecutive days of survey period). Considering the latent class indicators are indicator variables (U_{ijk}), the MLCA model can be written as follows:

$$P(U_{ij1} = s_1, U_{ij2} = s_2, \cdots, U_{ijk} = s_K) = \sum_{t=1}^{T} P(G_{ij} = t) \prod_{k=1}^{K} P(U_{ijk} = s_k | G_{ij} = t)$$

(2.3)

Where,

- U_{ijk} represents the response of eating carbohydrate (one of the following: not eating any food, $< 50\%$ of the energy, or $\geqslant 50\%$ of the energy) on *i*th day of the survey ($i \in (1, 2, 3, 4)$) in *j*th individual at the *k*th time slot of the day ($k \in (1, 2, 3, \cdots, 7)$);

- G_{ij} denotes the latent class membership for *j*th individuals on *i*th day of the survey, the total number of day level latent class is T;

- $P(U_{ijk} = s_k | C_{ij} = t)$ is the probability of a specific response pattern, conditional on membership in latent class *t*.

The $P(G_{ij} = t)$ in equation 2.3 is what we have already defined in equation 2.2:

$$P(G_{ij} = t) = \frac{\exp(\gamma_0 + \beta_1 x_{ij} + \gamma_1 w_j + u_{0j})}{1 + \exp(\gamma_0 + \beta_1 x_{ij} + \gamma_1 w_j + u_{0j})}$$

(2.4)

Non-Parametric approach

Since the parametric approach discussed above can be extremely computationally demanding (Van Horn et al., 2008; Vermunt, 2008), an alternative approach is using a non-parametric MLCA (Davidian et al., 2008). In this approach, separate latent class models are specified for level 1 (observation days) and level 2 (individuals). Similar with the parametric MLCA approach, there are $T - 1$ random intercepts, where T is the number of level 1 latent classes. However, rather than assuming the random intercepts following a normal distribution, the non-parametric MLCA

assumes a multinomial (discrete) distribution of the level 2 latent classes. This approach is less computationally demanding compared with the parametric approach. These level 2 (individual) latent classes reflect differences in the probability of belonging to a specific day level latent class, so that individuals that contain observation days with similar probabilities for the level 1 latent classes will be grouped together. The non-parametric MLCA model can be defined as follows:

$$P(C_{ij} = t | CB_j = m) = \frac{\exp(\gamma_{tm})}{\sum_{r=1}^{T} \exp(\gamma_{tm})} \tag{2.5}$$

Where,

- CB_j is individual level latent class membership for jth individual;

- γ_{tm} is day level and individual level indicators.

According to Finch and French's simulation study (Finch and French, 2014), non-parametric approach generally resulted in more accurate recovery of the underlying latent structure of the data at both levels and provided better latent class model compared with parametric approach. In the current project, we are interested in exploring both meaningful individual (level 2) latent classes and the daily (level 1) carbohydrate consumption classification. Therefore, non-parametric MLCA was employed 1) to identify latent classes of observation days (level 1) based on the subjects' responses to the 4-day food and drink diary and 2) to form distinct latent classes of individuals (level 2) based on the distribution of day-level carbohydrate eating temporal latent classes within individuals.

Strategy of conducting MLCA

To identify the best-fitting model, we used the following sequential modelling strategy (Henry and Muthén, 2010):

- Firstly, we ignored the multilevel structure of the data and estimated a series of traditional LC models to determine the number of classes at the observational day-level;

- Next, a series of MLCA models were fitted to account for the multilevel structrure of the data. In these models, the number of day-level classes was based on the best fitting LCA model from the first step, and the LCA model at the individual level was estimated to identify the number of individual level latent classes;

- Thirdly, when number of individual level latent classes is defined based on the previous stage, day-level classes was modified (one class lower and one

class higher than in the second step) to see the effect of changing level 1 classes and confirm the best fitting model.

The number of classes in either level 1 were determined by 1) the evaluation of model fit indices, including the Bayesian information criterion (BIC), and entropy which is a statistic that summarizes latent class probabilities where values near 1 indicate better latent class separation; 2) the Lo-Mendell-Rubin Likelihood Ratio Test (LMR-LRT) (Lo, Mendell, and Rubin, 2001; Nylund, Asparouhov, and Muthén, 2007) which compares q vs. $q - 1$ classes models, where q is the number of latent classes and most importantly, 3) pattern interpretability. In the steps of performing multilevel LCA, where LMR-LRT is not available, same rules of model fit indices and pattern interpretability were used to determine the optimal combination of latent classes in observation day level and individual level. MLCA models were fitted in Mplus 7.4 (Muthén and Muthén, 2017), the Mplus syntax and outputs are shown in **Appendix B**.

Characteristics of day level latent classes and individual level latent classes

Day level latent classes identified by the first step of MLCA were tabulated according to day of week and also whether the diary was recorded during weekends or not. A contingency table giving the frequency of responses across the 7 time slots of the survey days was produced. Descriptive statistics for the dietary day level recordings according to the latent class memberships were presented. Pearson χ^2 test was used to compare the distribution of categorical variables. One-way Analysis of Variance (ANOVA) was used to compare the means across the multiple groups for continuous variables.

Person level point estimates and 95% confidence intervals (CIs) were determined by applying individual, nurse visiting, and blood sample weights accordingly which account for the probability of participant selection and the clustered survey design. Descriptive statistics for sample characteristics are presented as weighted means (95% CI) or weighted percentages (95%CI). After examining the distribution of the data, the following variables were log-transformed to improve normality: fasting blood glucose, A1C, TC, LDL, HDL, TG, and average physical activity duration per day. Weighted geometric means (95% CI) were used for all log-transformed variables.

To see whether there is any temporal pattern for food intake eating could also be defined at individual level, weighted estimates of nutrients consumption across

the 7 time slots of the day were calculated for each individual level latent class. Contributions (%) of the average energy intake within time slots were evaluated by determining the percentages of energy coming from carbohydrate, fat, protein, and alcohol intake.

For continuous variables, the *F* test was used to determine differences between latent classes with Bonferroni correction to account for multiple testing across > 2 classes when applicable. For categorical variables, differences between latent classes were assessed using the adjusted Pearson χ^2 test for survey data.

Association between individual level latent classes and the prevalence of hypertension, and measurements of obesity

Associations between individual level carbohydrate eating classes and hypertension (yes/no), body mass index (BMI, kg/m^2), and waist circumference (WC, cm) were explored in men and women separately. Point estimates of weighted means and proportions and 95%CI of the characteristics were determined by applying either nurse visiting weights (for outcomes of hypertension, BMI, and WC) or blood sample weights (for diagnosis of DM) accordingly. Similarly, *F* tests (for continuous variables) and adjusted Pearson χ^2 tests (for categorical variables) were used to determine sex-specific differences by hypertension status, and BMI categories.

Survey-designed logistic regression models (for hypertension), and linear regression models (for WC, BMI), were used to test for associations between latent classes of carbohydrate eating patterns and hypertension, BMI, and WC, in the NDNS RP sample, separately. Since diabetic participants might or might not modify their carbohydrate eating habits, we also fitted all the above mentioned regression models restricted to those without diabetes.

For the multiple regression models, model fitting strategies are as follows:

1. The crude association between the carbohydrate eating groups and the outcomes was first examined.

2. Potential confounders of the association between carbohydrate eating groups (exposure) and the outcomes were selected depending on the descriptive statisitical analyses conducted above, i.e. those are associated with both the exposure and the outcome and also not on the causal pathway were selected as potential confounders. Those are strongly related with the outcomes but may not associated with carbohydrate eating groups may reduce the standard errors and so improve the precisions are also considered in the linear regression models.

3. Confounding and/or interaction effect from each of the potential factors were checked one by one. Interaction effect were tested using the adjusted Wald test testing whether the regression coefficients of the interaction terms are simultaneously equal to zero.

4. A preliminary model that includes all of the variables suggested to be confounders in the previous step was established.

5. The remaining variables were added to the preliminary model one by one to see if any of them may be a confounder in condition of the presence of the other covariates.

6. For logistic regression models (hypertension) under the survey data, goodness-of-fit was assessed using the adapted `svylogitgof` command in Stata (Archer and Lemeshow, 2006). Other diagnostics for regular logistic regression models, such as estimating the pseudo-R^2, AIC or BIC, checking the standardized Pearson residuals, or covariate pattern residuals are currently not available for weighted survey data.

7. For linear regression models (WC, BMI), assumption of independent observations is violated as soon as we weighted the sample. General checking such as QQ plots of the residuals (normality), plotting the residuals against fitted values (constant variance) are not available as well. Outliers, leverage, and Cook's distance cannot be check either, however, participants with extreme weightings (if exist) were checked by removing them and refit the models as a sensitivity analysis.

8. Since under survey design data, the sampling-weighted least squares are not maximum likelihood, it would not be possible to compare models using likelihood ratio test. Instead, adjusted Wald tests with $p < 0.05$ were used as criteria for variable inclusion in the final model. Another Stata command `linktest` was also used to decide whether quadratic and cubic terms of continuous variables were necessary in improving the fitting of model (Pregibon, 1980).

Data manipulation and preparation **(Appendix A)** were done in R version 3.5.1 (R core Team, 2018). All statistical analyses, except for MLCA models, were performed with `svyset` command as implemented in Stata software version 15.1 (StataCorp LLC, 2017). The process of model fitting, covariates selection, interaction effect testings for the association between carbohydrate eating patterns and hypertension is shown as an example in **Appendix D**. All p values were two-sided.

Chapter 3

Results

Model selection, and interpretation

A series of traditional LCA of the responses to carbohydrate intake within 7 time slots of day was first examined. These initial analyses ignored the clustering of observation days within participants of the survey. **Table 3.1** shows the latent class solutions for one to five classes (see rows under the Fixed effects model section). The BIC declines with the number of day level classes increases. However, the improvement of BIC dropped to less than 1000 from 3 classes to 4 classes solutions (658.9) and from 4 classes to 5 classses solutions (361.7). Entropy index indicates that the 4 classes model could explain about 51% percent of the data, while p values of Lo-Mendell-Rubun LRT suggest that the more classes we fit, the better model we will have until up to 6 classes ($p = 0.06$ and is not shown in the table). From the parsimony point of view, we extended the model with random effects building on 2 classes, 3 classes and 4 classes solutions.

The results of the random effect included models are presented in **Table 3.1** under the Random effects model section. It is obvious that the BIC improves with the addition of the random effects which account for the nested structure of the data. Entropy indicates that 4 classes in individual level and 2 classes in the day level may be the best solution mathematically. However, after these solutions were checked in more details, the potentially most substantively interpretable model was found to be the 3×3 random effect model, which is the model with 3 latent classes in the day level, and 3 latent classes in the individual level. We must emphasize that different researchers may have made decision slightly different from ours, we provided the descriptions and figures for other solutions in the **Appendix C** for reference.

In the 3×3 random effect model solution we have chosen, there were 39.5%, 20.4%, and 40.1% observations classified into 3 latent groups in the day level. The overall counts and percentages for each responses within every time slot and the distributions of the solution are presented in **Table 3.2**. The trajectories illustrating the

change of the probabilities of each response to carbohydrate eating during the hours of the day are shown separately by three types of days in **Figure 3.1**.

TABLE 3.1: Fit criteria for each model specification.

Model	Number of day level classes				
	1 class	2 classes	3 classes	4 classes	5 classes
Fixed effects model					
No. of free parameters	14	29	44	59	74
Log-likelihood	-173793.306	-172669.771	-172039.204	-171633.941	-171377.292
BIC	347728.092	345632.608	344523.060	343864.121	343502.409
Lo-Mendell-Rubun LRT	–	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Entropy	1	0.310	0.392	0.510	0.481
Random effects model					
2 individual level classes					
No. of free parameters		59	89	119	
Log-likelihood		-169331.132	-168700.96	-168366.193	
BIC		339258.502	338301.338	337934.968	
Entropy		0.581	0.569	0.555	
3 individual level classes					
No. of free parameters		89	134	179	
Log-likelihood		-166936.279	-166348.815	-166062.761	
BIC		334771.968	334051.799	333934.448	
Entropy		0.677	0.630	0.644	
4 individual level classes					
No. of free parameters		119	179		
Log-likelihood		-165441.731	-164845.696		
BIC		332086.045	331500.318		
Entropy		0.729	0.659		

Note:

Abbreviations: No, number; BIC, Bayesian information criterion; Entropy, a pseudo-r-squared index;

Lo-Mendel-Rubin LRT, likelihood ratio test comparing q classes models with $q - 1$ classes models.

Class 1 days **(Figure 3.1-A)** were given the name of "high percentage carbohydrate day" since in these days of survey, the probabilities of carbohydrate contributed higher or equal to 50% of the energy consumed were always higher than that in the other two types of days. Specifically, high percentage carbohydrate days were characterised with probabilities of over 0.6 in time slots between 6 am to 9 am, 9 am to 12 noon, and also 2 pm to 5 pm, during which the time slots may be interpreted as breakfast, morning snack, and afternoon snack time periods for many participants. Moreover, even during late night time period, such as 8 pm to 10 pm, and 10 pm to 6 am time slots, the probabilities of having higher carbohydrate contained food were still as high as 0.412, and 0.246, respectively.

TABLE 3.2: Day level latent class solution for three classes LCA model.
(No individual level model)

Time slots of	Responses to			Class 1 days (39.5%)	Class 2 days (20.4%)	Class 3 days (40.1%)
the day	carbohydrate intake	*n*	(%)	High perc- entage carb	Low perc- entage carb	Regular meals
6 am – 9 am						
	Not eating any food	7655	31.2	0.129	0.450	0.320
	Carbohydrate < 50%*	4500	18.4	0.130	0.267	0.128
	Carbohydrate ⩾ 50%†	12328	50.4	0.741	0.283	0.552
9 am – 12 noon						
	Not eating any food	5447	22.2	0.237	0.079	0.401
	Carbohydrate < 50%	7227	29.5	0.158	0.492	0.173
	Carbohydrate ⩾ 50%	11809	48.2	0.605	0.429	0.426
12 noon – 2 pm						
	Not eating any food	4783	19.5	0.156	0.356	0.019
	Carbohydrate < 50%	11112	45.4	0.405	0.413	0.560
	Carbohydrate ⩾ 50%	8588	35.1	0.439	0.231	0.421
2 pm – 5 pm						
	Not eating any food	6926	28.3	0.130	0.123	0.659
	Carbohydrate < 50%	8277	33.8	0.249	0.602	0.076
	Carbohydrate ⩾ 50%	9280	37.9	0.621	0.276	0.266
5 pm – 8 pm						
	Not eating any food	3043	12.4	0.114	0.199	0.034
	Carbohydrate < 50%	14240	58.2	0.516	0.590	0.639
	Carbohydrate ⩾ 50%	7200	29.4	0.370	0.211	0.328
8 pm – 10 pm						
	Not eating any food	8722	35.6	0.322	0.291	0.480
	Carbohydrate < 50%	8898	36.3	0.266	0.551	0.212
	Carbohydrate ⩾ 50%	6863	28.0	0.412	0.158	0.308
10 pm – 6 am						
	Not eating any food	16295	66.6	0.680	0.590	0.751
	Carbohydrate < 50%	4144	16.9	0.074	0.294	0.101
	Carbohydrate ⩾ 50%	4044	16.5	0.246	0.115	0.148

Note:

Abbreviation: LCA, latent class analysis, carb is short for carbohydrates.

* Carbohydrate < 50% indicates that within the time slot, carbohydrate contributed < 50% total energy intake.

† Carbohydrate ⩾ 50% indicates that within the time slot, carbohydrate contributed ⩾ 50% total energy intake.

Class 2 days **(Figure 3.1-B)** were named as "low percentage carbohydrate day" because first of all, in these days the probability of participants skipping breakfast was 0.45. And after 9 am, within these days, the probability of having low carbohydrate contained food (carbohydrate contributed < 50% of total energy intake), was always higher than having high carbohydrate contained food (carbohydrate contributed ⩾ 50% of total energy intake). In class 2 days, participants also turned to have morning snacks (with only 0.079 possibility of **not** eating any food and

similar probabilities of having either high or low carbohydrate contained food). This phenomenon may also be interpreted as having a long and late breakfast (brunch) in these mornings. The probability of **not** eating any food was the lowest for low carbohydrate days during the midnight time slot (10 pm to 6 am), with probability of 0.590 compared with 0.680 and 0.751 in class 1 and class 3 days, respectively.
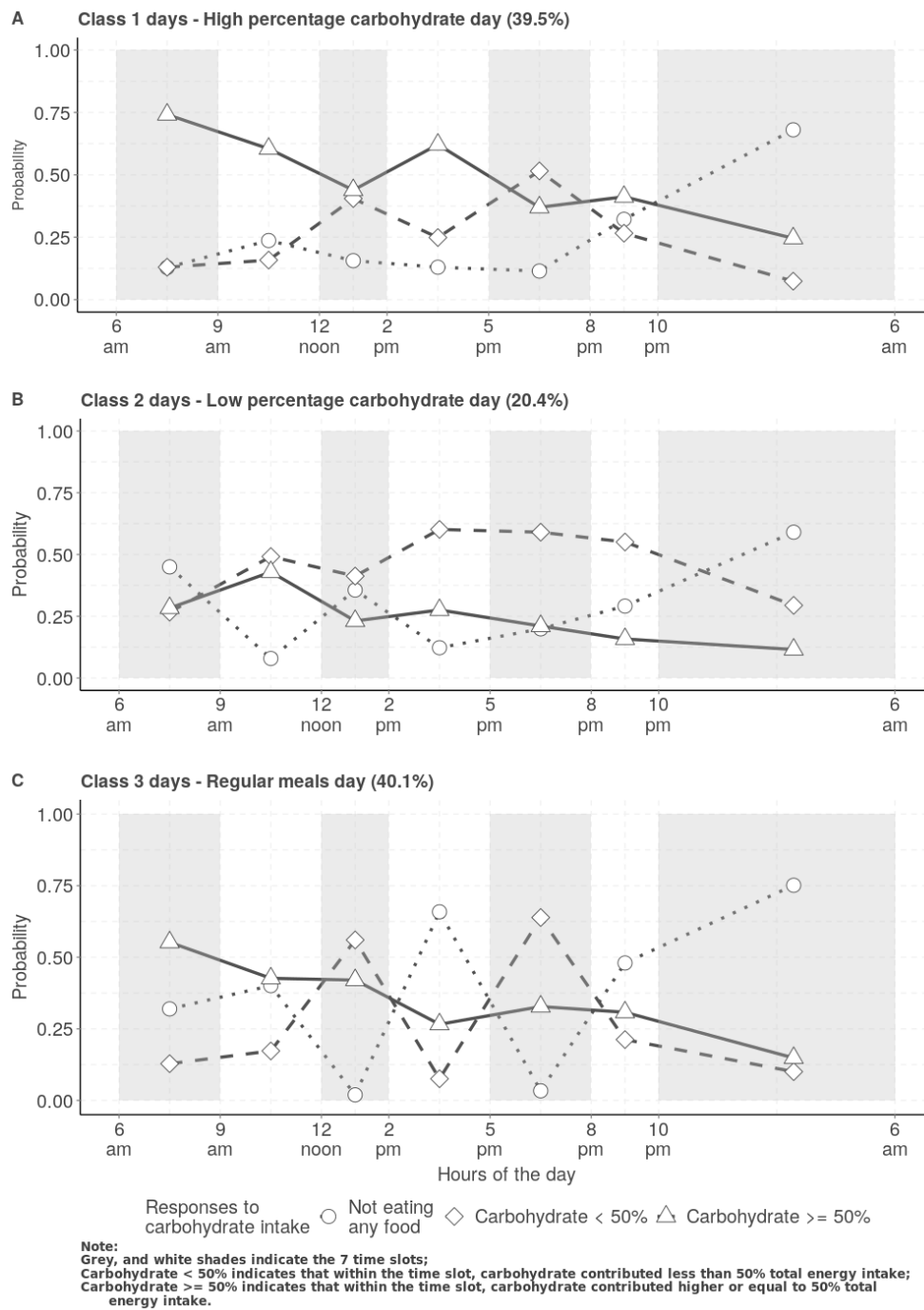
FIGURE 3.1: Day Level Latent Classes Solution.

TABLE 3.3: Means (standard deviations), and counts (%) of the characteristics of different types of carbohydrate eating days.

	High percentage carb	Low percentage carb	Regular meals	*P* value[*]
Counts (%)	9667 (39.5)	5002 (20.4)	9814 (40.1)	
Country (%)				< 0.001
England	5627 (58.2)	2972 (59.4)	5291 (53.9)	
Northern Ireland	1194 (12.4)	527 (10.5)	1400 (14.3)	
Scotland	1527 (15.8)	813 (16.3)	1774 (18.1)	
Wales	1318 (13.6)	690 (13.8)	1349 (13.7)	
Day of Week (%)				< 0.001
Monday	1303 (13.5)	715 (14.3)	1370 (14.0)	
Tuesday	1266 (13.1)	674 (13.5)	1290 (13.1)	
Wednesday	1225 (12.7)	740 (14.8)	1233 (12.6)	
Thursday	1272 (13.2)	752 (15.0)	1425 (14.5)	
Friday	1458 (15.1)	797 (15.9)	1479 (15.1)	
Saturday	1537 (15.9)	703 (14.1)	1495 (15.2)	
Sunday	1605 (16.6)	621 (12.4)	1522 (15.5)	
Weekend, Yes (%)	3142 (32.5)	1324 (26.5)	3017 (30.7)	< 0.001
Total energy (kJ)	7539.98 (2875.87)	7160.22 (2922.15)	7439.68 (2978.91)	< 0.001
Carbohydrate (g)	222.79 (89.84)	209.70 (86.17)	206.59 (84.42)	< 0.001
Protein (g)	71.36 (29.79)	69.55 (30.20)	73.29 (32.94)	< 0.001
Fat (g)	65.44 (33.27)	63.94 (33.76)	67.24 (34.73)	< 0.001
Alcohol (g)	11.76 (27.31)	8.85 (24.25)	13.80 (33.00)	< 0.001
Total sugars (g)	98.63 (56.03)	88.03 (50.50)	86.39 (50.96)	< 0.001
Starch (g)	124.07 (55.84)	121.59 (56.13)	120.11 (54.62)	< 0.001
Non-milk extrinsic sugar[†]	59.45 (49.31)	50.07 (43.41)	50.41 (44.84)	< 0.001
Fruit (g)	107.40 (137.97)	103.15 (129.08)	92.76 (126.02)	< 0.001
Yellow Red Green Vegetables (g)	26.52 (46.44)	26.84 (47.99)	26.16 (45.99)	0.681

Note: carb is short for carbohydrate.

[*] *P* values were obtained from Pearson χ^2 test for categorical variables, and one-way ANOVA comparing the means in multiple groups for continuous variables;

[†] Non-milk extrinsic sugar is defined as: additionally added free sugar, such as table sugar, honey, glucose, fructose and glucose syrups, sugars added to food and sugars in fruit juices.

Class 3 days **(Figure 3.1-C)** were called "regular meals day" due to the following reasons: 1) participants' dietary recordings showed that in these days there was almost 0 possibility of not eating any food at lunch (0.019 between 12 noon and 2 pm) and dinner (0.034 between 5 pm and 8 pm); 2) the probabilities of not eating during morning snack time (9 am to 12 noon) and afternoon snack time (2 pm to 5 pm) were also the highest among the three types of days (0.401 and 0.659). 3) during these days, participants may have some high carbohydrate contained food between 8 pm and 10 pm (probability = 0.308), but the probability of not eating any food during 10 pm to 6 am next morning was 0.751, the highest among the three types of days.

Features of the three types of carbohydrate eating temporal patterns

The details of the characteristics of the three types of carbohydrate eating time pattern were listed in **Table 3.3**. Specifically, regular meals day turned to be recorded

slightly more often in Northern Ireland, and Scotland. In terms of day of week distribution in the three types of days, there is strong evidence ($p < 0.001$) that high carbohydrate days appeared more frequently in weekends (32.5%) compared with low carbohydrate day (26.5%) and regular meals day (30.7%).

As expected, consumption of total energy (7539.98 kJ), total carbohydrate (222.79 g), total sugar (98.63 g), starch (124.07 g), and non-milk extrinsic sugar (59.45 g) were the highest among high percentage carbohydrate days (all $p < 0.001$). On the other hand, the consumption of protein (73.29 g), total fat (67.24 g), and alcohol (13.80 g) were the highest in the regular meals days. Moreover, in high percentage carbohydrate days, participants turned to consume the highest amount of fruit (107.40 g). There was no evidence of any difference for the consumption of yellow, red, or green vegetables across the three types of days ($p = 0.681$).

Individual level LCA solution

In the random effect models we utilized the non-parametric approach, in which we added a level 2 (individual level) latent classes based on the random means from the level 1 (day level) latent class solution. The results of the individual level LCA solution for 2 and 3 classes are presented in **Figure 3.2**, and **3.3**.
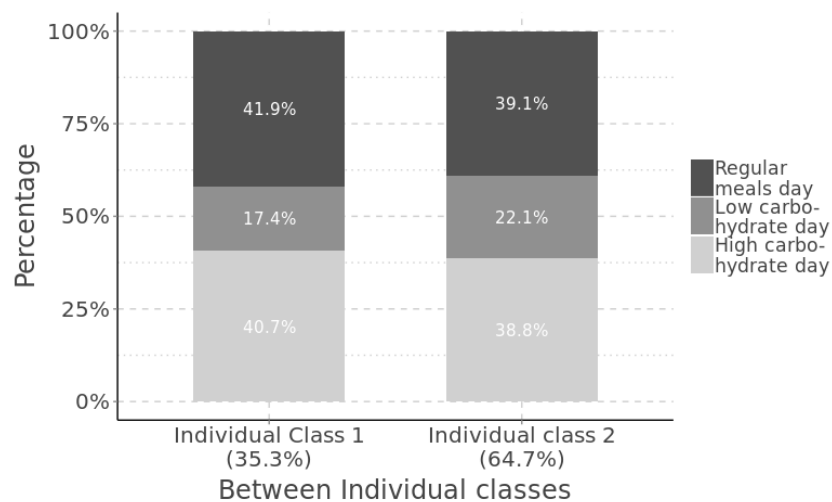
FIGURE 3.2: Multilevel Latent Class Solution, 3 classes in day level, 2 classes in individual level.

With two individual level latent classes **(Figure 3.2)**, one individual class is comprised of individuals with a relatively slightly higher proportion of having "low carbohydrate day" (22.1%) compared to the other (17.4%). This class represents
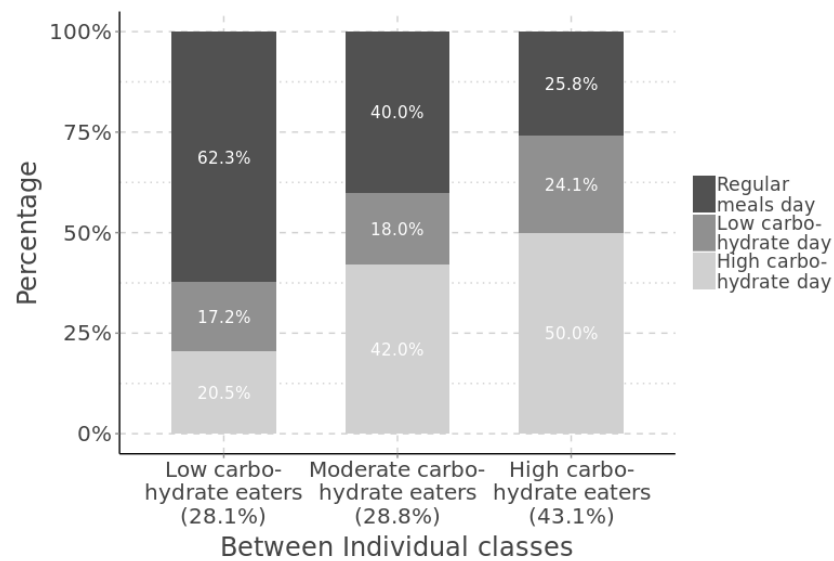
FIGURE 3.3: Multilevel Latent Class Solution, 3 classes in day level, 3 classes in individual level.

nearly 65% of the individuals. However, we believe these individual classes are not very distinguishable to each other.

With three individual level latent classes **(Figure 3.3)**, a low-carbohydrate eaters class, a moderate-carbohydrate eaters class, and a high-carbohydrate eaters class emerges. 43.1% participants were identified as high-carbohydrate eaters, in these individuals, about 50% of the days (2 out of 4 days) of their dietary diary could be classified as having high carbohydrate days. Nearly 1 out of 4 days of their dietary diary were either "regular meals day" or "low carbohydrate day". 28.1% participants fell into the low carbohydrate eaters class in the left hand side of **Figure 3.3**, their recordings of food intake showed that in more than 60% of their days, they were having "regular meals" which was characterised as with highest amount of fat and alcohol consumptions as already described in **Table 3.3**. Moderate carbohydrate eaters have comparable proportions (42.0% vs. 40.0%) of having high carbohydrate days and regular meals day, 18.0% of their dietary diary were found to be low carbohydrate days.

After recognising that there were three potential latent groups of carbohydrate eaters in the UK adults, whose food consumption pattern were also probably switching from one to another during the survey, their average carbohydrate contribution to total energy intake (as well as the subtypes of carbohydrate actually consumed) within the 7 pre-defined time slots of the day were still of interest. Survey-design-weighted mean energy intake within each time slot of the day and

their composition of contribution are illustrated in **Figure 3.4**, weighted mean nutrients intakes are listed in **Table 3.4**.
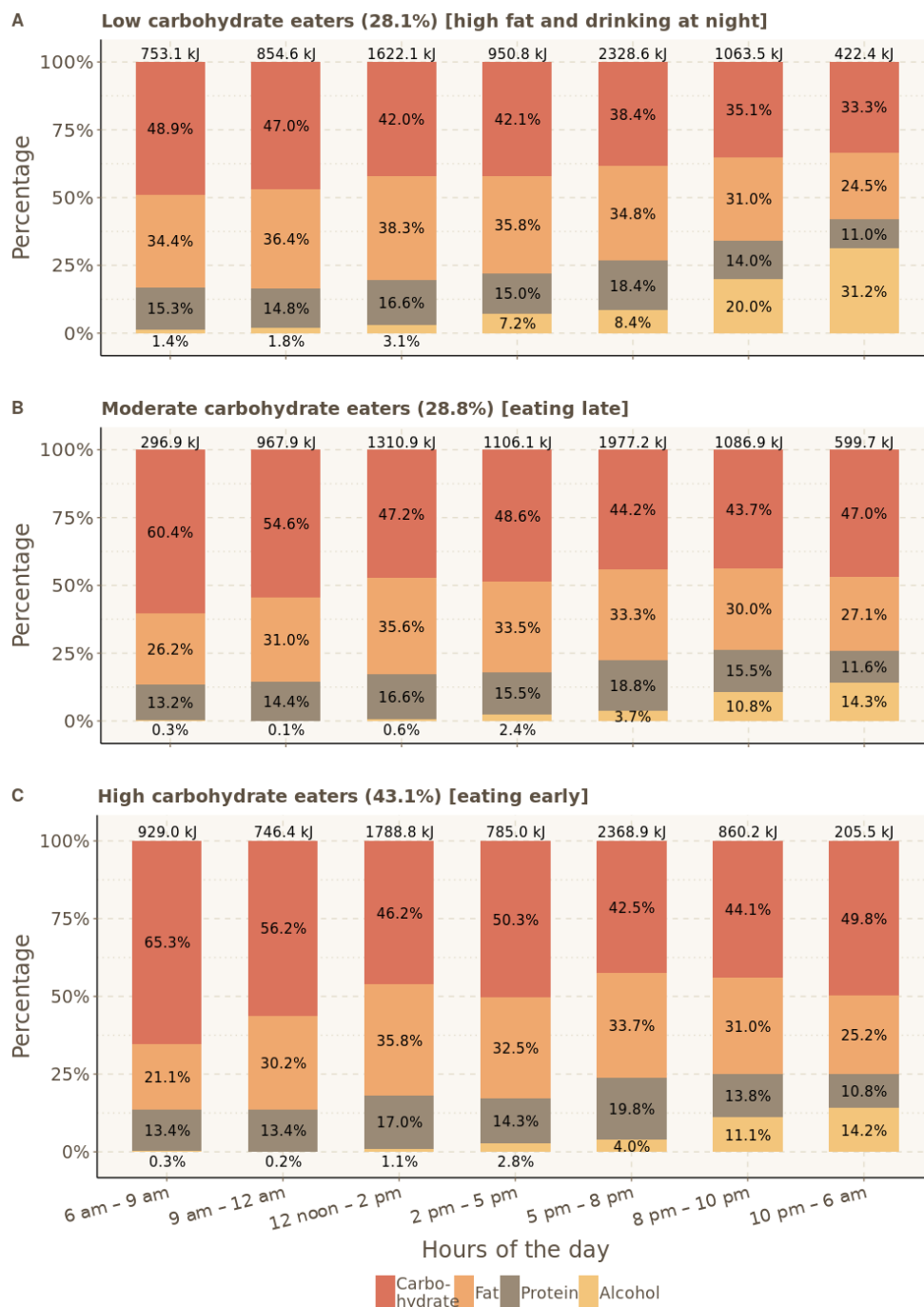
FIGURE 3.4: Sources of Energy Contribution at Each Time Slot by Individual Carbohydrate Eating Groups.

Among the three types of carbohydrate eaters, the mean of total energy intake over the 4 days of dietary survey was the highest (7985.8 kJ, 95%CI: 7283.3, 8146.3) in the low carbohydrate eaters group, and the lowest (7341.8 kJ, 95%CI: 7172.5, 7511.2) in the moderate eaters group **(Table 3.4)**. Sources of energy for each type of

carbohydrate eaters by the 7 time slots were also different. Low carbohydrate eaters **(Figure 3.4-A)** never had carbohydrate contributed more than 50% of their total energy throughout the day. Energy from fat were the highest for low carbohydrate eaters most of the time during the day (except for time between 10 pm to 6 am next morning). Most impressively, energy from alcohol were always the highest in low carbohydrate eaters, percentages for energy from alcohol for the 7 times slots were 1.4% (6-9 am), 1.8% (9-12 noon), 3.1% (12-2 pm), 7.2% (2-5 pm), 8.4% (5-8 pm), 20.0% (8-10 pm), and 31.2% (10 - 6 am), respectively. Contribution from different energy sources are quite similar for moderate and high carbohydrate eaters, but their absolute amount of energy consumption at each time slot were largely different. Moderate carbohydrate eaters **(Figure 3.4-B)** were characterised as consuming the lowest energy (296.9 kJ) before 9 am, but having higher energy consumption (967.9 kJ) between 9 am and 12 noon time compared with low and high carbohydrate eaters. Moderate carbohydrate eaters may turn to have later breakfast, later lunch, and probably later dinner as well. They had the highest total energy consumption (599.7 kJ) at night (10 pm - 6 am) across three types of eaters. High carbohydrate eaters **(Figure 3.4-C)** consumed the highest total energy (929.0 kJ) during 6 am to 9 am in the morning and the lowest total energy between 10 pm to 6 am (205.5 kJ). Specifically, carbohydrate contribution to total energy intake were 65.3% (6-9 am), 56.2% (9-12 noon), 46.2% (12-2 pm), 50.3% (2-5 pm), 42.5% (5-8 pm), 44.1% (8-10 pm), and 49.9% (10-6am). We also noticed that high carbohydrate eaters consumed their energy mainly from three time slots: 6-9 am, 12-2 pm, and 5-8 pm.

As expected, in total, the mean of carbohydrate intake was 203.8 g, 218.3 g, and 233.4 g for low, moderate, and high carbohydrate eaters, respectively **(Table 3.4)**. Energy contribution from carbohydrate was close to 50% in the high carbohydrate eaters, but was only 40.6% in the low carbohydrate eaters. In terms of the subtypes (components) of the carbohydrate consumed at each time slot, high carbohydrate eaters consumed as much as more than 2 times (compared to low carbohydrate eaters) and nearly 4 times (against moderate carbohydrate eaterss) the amount of sugar (37.9g 95% CI: 36.8, 39.2) and non-milk extrinsic sugar (i.e. free sugar) (11.1g 95%CI: 10.7, 11.6) between 6-9 am. Moderate carbohydrate eaters had their carbohydrate intake more spread out. They consumed more sugar and starch during 9-12 noon, 2-5 pm, 8-10 pm, and 10-6 am. Low carbohydrate eaters turned to have similar temporal pattern of consuming carbohydrates but the absolute amount of fibre, sugar, free sugar, and starch were usually lower than that in the high carbohydrate eaters except for time slots of 2-5 pm, and 10-6 am. Strong evidence ($p < 0.001$) suggested that the mean of total fibre consumption for low,

TABLE 3.4: Weighted means and percentages (95%CI) of the nutrients
intake according to individual level carbohydrate eating classes.
(NDNS RP 2008/09-15/16, sample size = 6155)

Variables	Low carbo-hydrate eaters (n = 1730)	Moderate carbo-hydrate eaters (n = 1772)	High carbo-hydrate eaters (n = 2653)	P value [*]
Total energy (kJ)	7985.8 (7823.3, 8146.3)	7341.8 (7172.5, 7511.2)	7677.8 (7555.8, 7799.8)	< 0.001
Carbohydrate (g)	203.8 (199.8, 207.8)	218.3 (212.9, 223.7)	233.4 (229.6, 237.2)	< 0.001
6 am – 9 am	23.0 (21.8, 24.3)	11.2 (10.0, 12.3)	37.9 (36.8, 39.2)	
Fibre (g)	1.4 (1.3, 1.5)	0.6 (0.5, 0.7)	2.0 (1.9, 2.2)	
Sugar (g)	10.2 (9.6, 10.9)	5.3 (4.8, 5.8)	19.7 (19.0, 20.4)	
NMES (g)[†]	4.7 (4.3, 5.1)	3.2 (2.9, 3.6)	11.1 (10.7, 11.6)	
Starch (g)	12.8 (12.0, 13.5)	5.9 (5.1, 6.6)	18.3 (17.6, 19.1)	
9 am – 12 noon	25.1 (23.9, 26.3)	33.0 (31.4, 34.6)	26.2 (25.1, 27.2)	
Fibre (g)	1.5 (1.4, 1.6)	1.6 (1.5, 1.7)	1.3 (1.2, 1.3)	
Sugar (g)	11.6 (10.9, 12.3)	15.7 (14.8, 16.6)	14.2 (13.6, 14.8)	
NMES (g)[†]	5.7 (5.2, 6.2)	9.6 (8.9, 10.2)	8.1 (7.7, 8.5)	
Starch (g)	13.5 (12.8, 14.3)	17.3 (16.4, 18.3)	11.9 (11.3, 12.6)	
12 noon – 2 pm	42.6 (40.9, 44.3)	38.7 (37.0, 40.4)	51.6 (50.2, 52.9)	
Fibre (g)	3.1 (2.9, 3.2)	2.3 (2.2, 2.5)	3.6 (3.5, 3.7)	
Sugar (g)	14.7 (14.0, 15.4)	14.9 (14.0, 15.7)	19.4 (18.7, 20.0)	
NMES (g)[†]	7.3 (6.7, 7.8)	9.1 (8.4, 9.8)	10.3 (9.8, 10.8)	
Starch (g)	27.9 (26.6, 29.1)	23,8 (22.6, 24.9)	32.2 (31.2, 33.1)	
2 pm – 5 pm	25.0 (23.6, 26.4)	33.6 (31.6, 35.6)	24.7 (23.6, 25.7)	
Fibre (g)	1.6 (1.5, 1.7)	1.9 (1.7, 2.0)	1.3 (1.2, 1.4)	
Sugar (g)	11.9 (11.3, 12.7)	14.5 (13.5, 15.5)	13.4 (12.8, 13.9)	
NMES (g)[†]	6.9 (6.4, 7.5)	9.9 (9.0, 8.6)	8.6 (8.2, 9.1)	
Starch (g)	13.1 (12.1, 13.9)	19.1 (17.7, 20.4)	11.3 (10.6, 11.9)	
5 pm – 8 pm	55.9 (54.1, 57.9)	54.6 (52.1, 57.0)	62.9 (61.3, 64.4)	
Fibre (g)	4.4 (4.2, 4.5)	3.7 (3.5, 3.9)	4.9 (4.7, 5.0)	
Sugar (g)	18.7 (17.9, 19.5)	18.6 (17.6, 19.5)	21.8 (20.9, 22.5)	
NMES (g)[†]	10.2 (9.6, 10.8)	11.8 (10.9, 12.6)	12.1 (11.4, 12.7)	
Starch (g)	37.3 (35.8, 38.8)	35.9 (34.1, 37.9)	41.1 (39.9, 42.2)	
8 pm – 10 pm	23.3 (21.9, 24.6)	29.7 (27.6, 31.7)	23.7 (22.5, 24.9)	
Fibre (g)	1.4 (1.3, 1.6)	1.6 (1.5, 1.8)	1.3 (1.5, 1.8)	
Sugar (g)	10.9 (10.3, 11.5)	13.2 (12.2, 14.2)	12.4 (11.8, 13.0)	
NMES (g)[†]	7.3 (6.8, 7.8)	9.4 (8.5, 10.4)	8.3 (7.8, 8.8)	
Starch (g)	12.3 (11.4, 13.3)	16.4 (15.0, 17.8)	11.3 (10.5, 12.1)	
10 pm – 6 am	8.8 (7.7, 9.8)	17.6 (15.2, 19.9)	6.4 (5.8, 7.1)	
Fibre (g)	0.34 (0.29, 0.39)	0.74 (0.63, 0.85)	0.24 (0.21, 0.27)	
Sugar (g)	5.3 (4.6, 6.1)	10.0 (8.6, 11.5)	4.1 (3.7, 4.5)	
NMES (g)[†]	3.9 (3.3, 4.6)	7.7 (6.4, 8.9)	2.9 (2.6, 3.3)	
Starch (g)	3.5 (2.9, 3.9)	7.5 (6.3, 8.8)	2.3 (1.9, 2.7)	
Carbohydrate (%)	40.6 (40.2, 41.0)	47.3 (46.8, 47.8)	48.3 (47.9, 48.6)	< 0.001
Fibre (g)	13.7 (13.4, 14.0)	12.5 (12.1, 12.9)	14.7 (14.4, 14.9)	< 0.001
Protein (g)	79.9 (77.9, 81.8)	69.3 (67.6, 71.0)	73.7 (72.5, 74.8)	< 0.001
Fat (g)	74.7 (73.1, 76.4)	63.8 (62.1, 65.5)	65.7 (64.4, 67.0)	< 0.001
Alcohol (g)	20.8 (18.3, 23.2)	10.7 (9.4, 11.9)	8.9 (8.1, 9.8)	< 0.001

[*] P values were obtained from χ^2 test for categorical variables, and one-way ANOVA comparing the means
in multiple groups for continuous variables;

[†] NMES, Non-milk extrinsic sugar is defined as: additionally added free sugar, such as table sugar, honey,
glucose, fructose and glucose syrups, sugars added to food and sugars in fruit juices.

moderate, and high carbohydrate eaters were different: 13.7g (13.4, 14.0), 12.5g (12.1, 12.9), and 14.7g (14.4, 14.9) with all 95% CI being exclusive to each other. It is also noteworthy that low carbohydrate eaters consumed the highest average amount of protein (79.9 g, 17.2% of total energy), fat (74.7g, 35.4% of total energy), and alcohol (20.8 g, 6.8% of total energy) as we have described for **Figure 3.4**.

The social-demographic characteristics of the UK adults according to their individual level latent class membership are shown in **Table 3.5**. Moderate carbohydrate eaters were relatively younger ($p < 0.001$), and slightly less from England ($p = 0.007$). Gender distribution across the three types of carbohydrate eaters was fairly even ($p = 0.119$). The distribution of the carbohydrate eater types turned out to be changing with the year of survey. Low carbohydrate eaters represented 32.5% of the population in the first year of survey, but later dropped to lower than 30% (lowest in the third year, 22.6%) until the most recent year. Proportion of high carbohydrate eaters increased from 41.2% to the highest (50.6%) in the second year, but then started to decline to 38.4% in the 8th year of survey ($p = 0.015$). There was no evidence of difference in employment status across three types of carbohydrate eaters. However, strong evidence suggested that high carbohydrate eaters had the highest proportion (61.3%) of living with partner ($p < 0.001$); moderate carbohydrate eaters had the lowest average income (27180.8 £/year), highest proportion of non-white population (20.5%), and lower education level (23.3% with degree of higher education) compared with either low or high carbohydrate eaters.

Weighted means, percentages of anthropometric measurements, average of main nutrients intake, as well as biochemical characteristic profiles according to the latent carbohydrate eater groups are given in **Table 3.6**. Low carbohydrate eaters had higher mean BMI (27.8 kg/m^2) and larger mean WC (98.9/89.9 cm in men/women) compared with 27.2, 27.3 kg/m^2, and 95.9/88.7 (men/women), 98.1/87.2 (men/women) cm in moderate and high carbohydrate eaters. Moderate carbohydrate eaters had the highest prevalence of being a current smoker (27.8%), shortest time of daily physical activity (geometric mean: 0.87 hours/day), and the lowest prevalence of hypertension (20.2%).

From the results of blood tests, 6.9% of low carbohydrate eaters were found to be diabetic (diagnosed by A1C > 6.5%), while the percentages of diabetes in the moderate and high carbohydrate eaters were 3.5%, and 4.1% ($p < 0.011$), respectively. Although there was some evidence ($p = 0.027$) that fasting blood glucose level may be slightly higher in non-diabetic low carbohydrate eaters, the geometric mean for

TABLE 3.5: Weighted means, percentages, and 95% CIs of the social-demographic characteristics by carbohydrate eating latent class memberships in the UK adults.
(NDNS RP 2008/09-15/16, sample size = 6155)

Variables	Low carbo-hydrate eaters (n = 1730)	Moderate carbo-hydrate eaters (n = 1772)	High carbo-hydrate eaters (n = 2653)	*P* value [*]
Total (%)	28.4 (26.9, 29.9)	28.7 (27.1, 30.3)	43.0 (41.3, 44.7)	
Country (%)				0.007
England	84.5 (81.7, 86.9)	82.0 (79.3, 84.5)	84.7 (82.3, 86.8)	
Northern Ireland	2.1 (1.6, 2.8)	4.2 (3.2, 5.6)	2.2 (1.7, 3.0)	
Scotland	9.1 (7.0, 11.8)	8.6 (6.7, 11.1)	8.0 (6.3, 10.2)	
Wales	4.3 (3.3, 5.6)	5.1 (4.0, 6.4)	5.1 (4.0, 6.4)	
Age (years)	51.0 (49.9, 52.1)	40.3 (39.1, 41.6)	51.7 (50.7, 52.7)	< 0.001
Sex (%)				0.119
Men	50.0 (46.9, 53.1)	50.2 (47.0, 53.5)	46.6 (44.0, 49.1)	
Women	50.0 (46.9, 53.1)	49.8 (46.5, 53.0)	53.4 (50.9, 56.0)	
Survey years (% in rows)				0.015
1	32.5 (28.4, 36.9)	26.3 (21.9, 31.2)	41.2 (36.6, 46.0)	
2	26.8 (22.6, 31.3)	22.6 (18.6, 27.3)	50.6 (45.8, 55.4)	
3	22.6 (18.8, 26.9)	33.7 (28.6, 39.2)	43.6 (38.7, 48.7)	
4	27.9 (24.1, 32.2)	27.6 (23.8, 31.8)	44.4 (40.2, 48.7)	
5	27.9 (24.2, 32.0)	28.7 (24.4, 33.5)	43.3 (38.2, 48.6)	
6	28.0 (24.0, 32.4)	31.5 (26.9, 36.6)	40.5 (35.8, 45.3)	
7	29.1 (25.2, 33.4)	29.0 (24.5, 34.0)	41.8 (37.1, 46.7)	
8	31.1 (27.3, 35.3)	30.5 (25.9, 35.5)	38.4 (34.1, 42.8)	
Paid employment[†] (%)				0.907
Yes	40.3 (37.0, 43.6)	40.8 (37.1, 44.5)	39.8 (37.1, 42.6)	
No	59.7 (56.4, 63.0)	59.2 (55.5, 62.9)	60.2 (57.4, 62.9)	
Live with partner[‡] (%)				< 0.001
Yes	56.9 (53.6, 60.1)	38.4 (35.2, 41.8)	61.3 (58.7, 63.7)	
No	43.1 (39.9, 46.4)	61.6 (58.2, 64.8))	38.7 (36.3, 41,3)	
Household income, £/year	36558.5 (34800.2, 38316.8)	27180.8 (25597.9, 28763.7)	32171.6 (31024.9, 33318.2)	< 0.001
Ethnicity (%)				
White	94.2 (92.4, 95.6)	79.5 (76.4, 82.3)	91.9 (90.1, 93.4)	< 0.001
Non-White	5.8 (4.4, 7.6)	20.5 (17.7, 23.6)	8.1 (6.6, 9.9)	
Education (%)				
Degree or higher	29.0 (26.1, 32.1)	23.3 (20.5, 26.3)	26.2 (24.1, 28.5)	0.019
Lower than degree	71.0 (67.9, 73.9)	76.7 (73.7, 79.5)	73.8 (71.5, 75.9)	

Note:

Abbreviations: CI, confidence intervals; NDNS RP, national diet and nutrition survey rolling programme.

Variables were weighted by individual weights.

[*] For continuous variables, the *F* test was used to determine differences between latent classes with Bonferroni correction to account for multiple testing across > 2 classes. For categorical variables, differences between latent classes were assessed using the adjusted Pearson χ^2 test for survey data.

[†] Paid employment was defined as being in paid employment during the last 4 weeks prior to the survey.

[‡] Live with partner was defined as either living with a married husband/wife or a legally recognised civil partnership.

A1C was probably lower in moderate carbohydrate eaters (4.72, 95%CI: 5.39, 5.47). Total cholesterol, HDL, and LDL were all lower in the moderate carbohydrate eaters, while no evidence of any difference of TG was found across three types of carbohydrate eaters.

TABLE 3.6: Weighted means, percentages, and 95% CIs of the anthropometric measurements, average main nutrients intake and biochemical characteristics by carbohydrate eating latent class memberships in the UK adults. (NDNS RP 2008/09-15/16, sample size = 6155)

Variables	Low carbohydrate eaters (n = 1730)	Moderate carbohydrate eaters (n = 1772)	High carbohydrate eaters (n = 2653)	*P* value [*]
BMI (kg/m^2)	27.8 (27.4, 28.2)	27.2 (26.7, 27.7)	27.3 (26.9, 27.6)	0.006
WC (cm)				
Men	98.9 (97.4, 100.5)	95.9 (94.1, 97.8)	98.1 (96.9, 99.2)	0.056
Women	89.9 (88.7, 91.3)	88.7 (87.1, 90.3)	87.2 (86.1, 88.2)	0.005
Smoking status (%)				
Current	20.4 (18.0, 23.0)	27.8 (25.0, 30.9)	17.1 (15.4, 19.0)	< 0.001
Ex-smoker	29.3 (26.5, 32.2)	16.8 (14.6, 19.2)	26.1 (24.9, 28.3)	
Never	50.3 (47.2, 32.2)	55.4 (52.2, 58.6)	56.8 (54.3, 59.3)	
Physical activity (hours/day) [¶]	1.08 (0.97, 1.19)	0.87 (0.77, 0.97)	1.07 (0.98, 1.16)	0.005
Hypertension[†], Yes (%)	33.8 (30.2, 37.5)	20.2 (17.0, 24.0)	30.9 (26.9, 31.0)	< 0.001
Total energy intake (kJ)	7985.8 (7823.3, 8146.3)	7341.8 (7825.3, 8146.3)	7677.0 (7555.8, 7799.8)	< 0.001
Glucose (mmol/l)	5.17 (5.12, 5.23)	5.05 (4.99, 5.13)	5.10 (5.05, 5.15)	0.027
A1C (%)	5.47 (5.44, 5.51)	5.43 (5.39, 5.47)	5.50 (5.48, 5.53)	0.010
DM [§]	6.9 (5.0, 9.3)	3.5 (2.3, 5.3)	4.1 (2.9, 5.6)	0.011
TC (mmol/l)	4.95 (4.84, 5.05)	4.72 (4.62, 4.83)	4.95 (4.87, 5.03)	0.001
HDL (mmol/l)	1.39 (1.35, 1.43)	1.32 (1.28, 1.35)	1.39 (1.36, 1.42)	0.003
LDL (mmol/l)	2.88 (2.79, 2.97)	2.77 (2.68, 2.86)	2.93 (2.86, 3.00)	0.024
TG (mmol/l)	1.14 (1.08, 1.19)	1.11 (1.05, 1.17)	1.10 (1.06, 1.15)	0.629

Abbreviations: CI, confidence intervals; NDNS RP, national diet and nutrition survey rolling programme; BMI body mass index; WC, waist circumference; A1C, haemoglobin A1c; DM, diabetes mellitus; TC, total cholesterol, HDL, high density lipoproteins; LDL, low density lipoproteins; TG, triglycerides.

Glucose, A1C, TC, HDL, LDL, TG, and physical activity were expressed in geometric means (95% CI) because the data were positively skewed.

Variables from the blood tests (glucose and A1C) were weighted by blood sample weights, the other variables were weighted by nurse visiting weights. Glucose and A1C levels are estimated in subgroups of people without diabetes.

[*] For continuous variables, the *F* test was used to determine differences between latent classes with Bonferroni correction to account for multiple testing across > 2 classes. For categorical variables, differences between latent classes were assessed using the adjusted Pearson χ^2 test for survey data.

[¶] Physical activity was calculated as mean time spent at moderate or vigorous physical activity including both work-related and recreational activities during the most recent month before the survey.

[†] Hypertension was defined as either systolic blood pressure $\geqslant$ 140 mmHg or diastolic blood pressure $\geqslant$ 90 mmHg, or under treatment for hypertension.

[§] DM was defined by A1C > 6.5%.

Association between individual level latent classes and hypertension, and obesity.

Hypertension

Table 3.7 presents the characteristics of men and women participants in the NDNS RP 2008/09-15/16 by hypertension status. The weighted prevalences of hypertension were 30.4% in men and 27.5% in women. Among both sexes, there were strong evidence of differences by hypertension status for age, education level, living with a partner or not, smoking status, BMI, abdominal obesity (WC), and prevalence of diabetes ($p < 0.01$). No difference was found among either men or women for ethnicity. Strong evidence of difference was suggested in women for average household income (32741.5 £/year in non-hypertensive vs. 27862.0 £/year in hypertensive, $p < 0.001$), and physical activity level (geometric mean: 0.81 hours/day in non-hypertensive compared with 0.53 hours/day in hypertensive, $p < 0.001$) but not in men. Interestingly, in both sexes, hypertensive participants had higher proportion of being classified as low or high carbohydrate eaters; total energy and carbohydrate intake was higher in people with hypertension ($p < 0.001$).

The sex-specific associations of carbohydrate eating patterns with hypertension (both in total and in participants without diabetes) are shown in **Table 3.8**. In the crude models, moderate carbohydrate eaters had statistically significant lower odds of having hypertension than low carbohydrate eaters in both men and women irrespective to diabetes status. Among men, after adjustment for selected confounders, which includes: age, live with partner or not, education level, BMI, smoking status, and total energy intake, the odds ratio (OR) comparing moderate with low carbohydrate eaters was 0.68 (95% CI: 0.43, 1.07) and remained borderline significant ($p = 0.093$). 95% CI of the adjusted OR became narrower (OR: 0.64, 95% CI: 0.41, 1.01, $p = 0.054$) when BMI was replaced with WC in model 2. When diabetic men were excluded in the models, the ORs (95%CI) for moderate and high carbohydrate eaters compared with low carbohydrate eaters were 0.65 (0.41, 1.03) and 0.73 (0.51, 1.06), respectively. The negative associations between moderate carbohydrate eating pattern and hypertension were also observed in women, however, without any statistically significant evidence in the fully adjusted models. High carbohydrate eaters also had lower adjusted odds compared with low carbohydrate eaters, while the 95% CIs for the adjusted ORs were all wide and included the null value suggesting no evidence of any association in either men or women for high carbohydrate eating pattern and hypertension.

TABLE 3.7: Weighted means, percentages, and 95 % CIs of the characteristics by hypertension status in the UK adults. (NDNS RP 2008/09-15/16, sample size = 6155)

	Men (n = 2537)			Women (n = 3618)		
	Non-hypertensive	Hypertensive	P value*	Non-hypertensive	Hypertensive	P value*
Weighted prevalence (%)	69.6 (66.6, 72.5)	30.4 (27.5, 33.4)		72.5 (69.8, 75.0)	27.5 (25.0, 30.2)	
Age (years)	43.2 (41.7, 44.7)	59.9 (58.0, 61.7)	<0.001	43.9 (42.7, 45.1)	64.9 (63.4, 66.5)	<0.001
Ethnicity (%)			0.534			0.126
White	89.6 (86.5, 92.0)	91.1 (86.2, 94.4)		85.7 (82.7, 88.3)	90.2 (85.0, 93.7)	
Non-white	10.4 (8.0, 13.5)	8.9 (5.6, 13.8)		14.3 (11.7, 17.3)	9.8 (6.3, 15.0)	
Education (%)			0.006			<0.001
Degree or higher	30.3 (26.6, 34.2)	21.5 (17.3, 26.5)		33.0 (29.9, 36.3)	19.7 (15.8, 24.3)	
Lower than Degree	69.7 (65.8, 73.4)	78.5 (73.5, 82.7)		67.0 (63.7, 70.1)	80.3 (75.7, 84.2)	
Household income, £/year	34006.5 (31972.9, 36040.1)	32280.5 (29875.6, 34685.4)	0.284	32741.5 (31009.9, 34473.1)	27862.0 (25557.0, 30167.0)	<0.001
Live with partner‡, Yes, (%)	56.1 (51.8, 61.4)	66.6 (61.3, 71.5)	0.002	48.7 (45.1, 52.3)	58.9 (53.6, 63.9)	0.002
Smoking status			<0.001			<0.001
Current	19.7 (16.6, 23.1)	12.9 (9.5, 17.2)		15.2 (13.1, 17.6)	8.5 (6.2, 11.6)	
Ex-smoker	24.2 (21.1, 27.6)	38.8 (33.4, 44.5)		21.6 (19.1, 24.4)	32.2 (27.3, 37.4)	
Never	56.2 (52.1, 60.1)	48.3 (42.7, 54.0)		63.2 (60.1, 66.2)	59.3 (54.0, 64.4)	
Physical activity (hours/day) †	1.52 (1.33, 1.72)	1.29 (1.08, 1.53)	0.134	0.81 (0.73, 0.89)	0.53 (0.42, 0.64)	<0.001
BMI (kg/m²)	26.8 (26.4, 27.2)	29.5 (28.9, 29.9)	<0.001	26.4 (26.1, 26.8)	29.8 (29.2, 30.5)	<0.001
WC (cm)	95.0 (93.9, 96.2)	104.6 (103.2, 106.1)	<0.001	85.7 (84.8, 86.6)	95.7 (94.2, 97.2)	<0.001
DM§ (%)	3.7 (2.4, 5.7)	12.6 (8.9, 17.5)	<0.001	1.8 (1.0, 3.3)	7.9 (5.1, 11.9)	<0.001
Carbohydrate eating patterns (%)			<0.001			<0.001
Low	28.3 (24.8, 32.2)	37.1 (32.0, 42.5)		26.9 (24.1, 29.9)	32.0 (27.2, 37.2)	
Moderate	30.8 (26.9, 35.0)	19.3 (15.3, 24.1)		29.6 (26.4, 33.0)	18.4 (14.5, 22.9)	
High	40.8 (36.9, 44.9)	43.6 (38.2, 49.2)		43.5 (40.3, 46.8)	49.7 (44.1, 55.2)	
Total energy intake (kJ)	9021.4 (8791.9, 9251.0)	8366.2 (8094.9, 8637.4)	<0.001	6802.6 (6681.1, 6924.0)	6396.7 (6217.1, 6576.2)	<0.001
Carbohydrate intake (g)	259.2 (252.9, 265.3)	235.3 (227.8, 242.8)	<0.001	198.0 (194.2, 201.8)	184.5 (178.8, 190.1)	<0.001

Note:

Abbreviations: CI, confidence intervals; NDNS RP, national diet and nutrition survey rolling programme; BMI body mass index; WC, waist circumference.

Variables are weighted by nurse visiting weights.

* Significant sex-specific differences by hypertension status assessed using an *F* test for continuous variables or design-adjusted Pearson χ^2 test.

‡ Live with partner was defined as either living with a married husband/wife or a legally recognised civil partnership.

† Physical activity was calculated as mean time spent at moderate or vigorous physical activity including both work-related and recreational activities.

§ DM was defined by AIC > 6.5%.

TABLE 3.8: ORs (95%CI) of carbohydrate eating patterns with hypertension in the UK adults, with or without diabetes .
(NDNS RP 2008/09-15/16, sample size = 6155)

	Carbohydrate eating patterns				
Model	**Low**	**Moderate**	*P* value[*]	**High**	*P* value[*]
Men (n = 2537)					
Hypertension					
Crude model	1	0.48 (0.33, 0.70)	< 0.001	0.82 (0.59, 1.13)	0.217
Model 1[†]	1	0.68 (0.43, 1.07)	0.093	0.80 (0.56, 1.15)	0.227
Model 2	1	0.64 (0.41, 1.01)	0.054	0.75 (0.53, 1.08)	0.124
Hypertension in non-diabetics					
Crude model	1	0.49 (0.33, 0.73)	< 0.001	0.82 (0.59, 1.14)	0.241
Model 1[†]	1	0.69 (0.43, 1.09)	0.110	0.78 (0.54, 1.14)	0.197
Model 2	1	0.65 (0.41, 1.03)	0.066	0.73 (0.51, 1.06)	0.096
Women (n = 3618)					
Hypertension					
Crude model	1	0.52 (0.36, 0.75)	< 0.001	0.96 (0.72, 1.28)	0.773
Model 1[‡]	1	0.79 (0.45, 1.39)	0.415	0.89 (0.61, 1.30)	0.552
Model 2	1	0.78 (0.45, 1.36)	0.384	0.88 (0.62, 1.26)	0.483
Hypertension in non-diabetics					
Crude model	1	0.51 (0.35, 0.74)	< 0.001	0.98 (0.73, 1.31)	0.875
Model 1[‡]	1	0.79 (0.44, 1.42)	0.435	0.89 (0.61, 1.29)	0.534
Model 2	1	0.79 (0.45, 1.39)	0.415	0.87 (0.61, 1.25)	0.452

Note:

Abbreviations: OR, odds ratio; CI, confidence interval; BMI, body mass index; WC, waist circumference; NDNS RP, national diet and nutrition survey rolling programme.

Diabetes was defined by A1C > 6.5%. BMI was replaced with WC in Model 2s, other covariates remained the same with the corresponding Model 1s.

[*] *P* values were obtained from wald tests from logistic regression models.

[†] Adjusted for age (continuous), live with partner or not (binary), education level (higher or equal to degree level or not), BMI, smoking status (current, ex-smoker, never) , total energy intake (kJ);

[‡] Adjusted for age, live with partner or not, average household income (continuous), education level, BMI, smoking status, total energy intake (kJ), alcohol consumption (g/day).

Obesity (BMI and WC)

Table 3.9 shows the characteristics for participants according to their obesity status stratified by sex. The survey design-weighted prevalence for being overweight and obese in the UK adults were estimated to be 43.4% and 25.7% in men, and 30.9% and 27.4% in women. Obviously, abdominal obesity (WC) increased significantly with the elevated BMI level in both men and women. Overweight or obese participants were older, having lower total energy intake and lower carbohydrate intake compared with normal weight men and women ($p < 0.001$). Moreover, edu-

TABLE 3.9: Weighted means, percentages, and 95 % CIs of the characteristics by BMI categories in the UK adults. (NDNS RP 2008/09-15/16, sample size = 6155)

	Men (n = 2537)				Women (n = 3618)			
	Normal weight	Overweight	Obese	P value[*]	Normal weight	Overweight	Obese	P value[*]
Weighted prevalence (%)	30.9 (28.0, 33.9)	43.4 (40.4, 46.4)	25.7 (23.2, 28.4)	—	41.7 (39.0, 44.4)	30.9 (28.4, 33.5)	27.4 (25.1, 29.9)	—
BMI (kg/m²)	22.6 (22.3, 22.8)	27.3 (27.2, 27.5)	33.7 (33.3, 34.2)	< 0.001	22.2 (22.0, 22.4)	27.3 (27.2, 27.5)	35.0 (34.6, 35.4)	< 0.001
WC (cm)	84.5 (83.6, 85.4)	97.1 (96.4, 97.8)	112.7 (111.6, 113.9)	< 0.001	76.9 (76.2, 77.5)	89.0 (88.3, 89.7)	103.7 (102.6, 104.7)	< 0.001
Age (years)	40.3 (38.2, 42.4)	49.6 (47.9, 51.2)	45.0 (43.4, 46.7)	< 0.001	45.0 (43.4, 46.7)	50.4 (48.6, 52.3)	50.9 (49.1, 52.7)	< 0.001
Ethnicity (%)				0.466				0.879
White	88.7 (83.9, 92.2)	89.1 (85.6, 91.9)	91.9 (87.3, 94.9)		88.4 (84.9, 91.19)	88.6 (84.5, 91.7)	87.3 (82.5, 90.9)	
Non-white	11.3 (7.8, 16.1)	10.9 (8.1, 14.4)	8.1 (5.1, 12.7)		11.6 (8.9, 15.1)	11.4 (8.3, 15.5)	12.7 (9.1, 17.5)	
Education (%)				0.022				< 0.001
Degree or higher	29.5 (24.5, 35.0)	28.3 (24.3, 32.7)	20.1 (16.0, 25.0)		35.7 (31.8, 39.8)	24.2 (20.4, 28.4)	19.4 (16.1, 23.2)	
Lower than Degree	70.5 (65.0, 75.5)	71.7 (67.3, 75.7)	79.9 (75.0, 84.0)		64.3 (60.2, 68.2)	75.8 (71.6, 79.6)	80.6 (76.8, 83.9)	
Household income, £/year	33695.9 (30462.3, 36929.7)	35059.6 (32949.7, 37169.5)	30295.5 (27948.5, 32642.6)	0.011	34594.1 (32326.8, 36861.4)	29777.7 (27808.3, 31747.2)	27230.6 (25259.1, 29202.0)	< 0.001
Live with partner‡, Yes, (%)	40.3 (34.8, 46.1)	65.3 (60.8, 69.6)	65.6 (60.1, 70.8)	< 0.001	47.6 (43.2, 52.1)	52.2 (47.5, 57.0)	51.7 (46.7, 56.6)	< 0.001
Smoking status				< 0.001				0.042
Current	32.0 (26.8, 37.7)	18.7 (15.5, 22.4)	19.2 (15.0, 24.3)		19.5 (16.4, 22.9)	17.8 (14.8, 21.4)	16.4 (13.1, 20.3)	
Ex-smoker	17.3 (13.5, 22.0)	28.6 (24.8, 32.7)	32.9 (27.9, 38.4)		19.0 (15.9, 22.5)	24.4 (20.8, 28.3)	26.9 (22.8, 31.6)	
Never	50.6 (44.8, 56.4)	52.7 (48.2, 57.1)	47.9 (42.1, 53.7)		61.6 (57.4, 65.5)	57.8 (53.3, 62.2)	56.7 (51.8, 61.4)	
Physical activity† (hours/day)	1.58 (1.33, 1.85)	1.42 (1.24, 1.62)	1.41 (1.15, 1.70)	0.547	0.84 (0.74, 0.94)	0.71 (0.62, 0.79)	0.65 (0.53, 0.78)	0.038
Carbohydrate eating patterns (%)				0.072				0.253
Low	25.9 (21.0, 31.5)	30.6 (26.6, 35.0)	31.4 (26.6, 36.6)		24.8 (21.5, 28.5)	26.8 (22.8, 31.2)	29.5 (25.3, 34.1)	
Moderate	34.2 (28.6, 40.4)	25.5 (21.9, 29.6)	25.5 (20.6, 31.0)		27.6 (23.8, 31.8)	26.3 (22.3, 30.8)	29.8 (25.4, 34.6)	
High	39.9 (34.2, 45.8)	43.8 (39.6, 48.2)	43.1 (37.7, 48.7)		47.6 (43.3, 51.9)	46.9 (42.4, 51.4)	40.7 (36.0, 45.6)	
Total energy intake (kJ)	9351.2 (8961.7, 9740.7)	8786.9 (8595.1, 8978.7)	8465.3 (8196.4, 8734.1)	0.001	7048.9 (6894.4, 7203.4)	6570.1 (6406.2, 6734.0)	6566.4 (6360.7, 6772.1)	< 0.001
Carbohydrate intake (g)	268.7 (258.3, 279.2)	250.1 (244.1, 256.1)	239.1 (231.3, 246.8)	< 0.001	205.8 (200.2, 211.3)	190.1 (185.3, 194.9)	189.8 (183.7, 195.9)	< 0.001

Note:

Abbreviations: CI, confidence intervals; NDNS RP, national diet and nutrition survey rolling programme; BMI body mass index; WC, waist circumference.

Variables are weighted by nurse visiting weights.

[*] Significant sex-specific differences by BMI categories assessed using an F test (with Bonferroni correction to account for multiple testing across > 2 groups) for continuous variables or design-adjusted Pearson χ^2 test for categorical variables

‡ Live with partner was defined as either living with a married husband/wife or a legally recognised civil partnership.

† Physical activity was calculated as mean time spent at moderate or vigorous physical activity including both work-related and recreational activities.

cation level (p = 0.022 for men, < 0.001 for women), average household income (p = 0.011 for men, < 0.001 for women) were decreasing with increasing BMI. Living with partner or not was strongly positively associated with obesity in men but not in women. Men with obesity were also found to have the lowest proportion of never being a smoker (47.9 %), and the highest proportion of being ex-smoker (32.9%). Association between smoking status and obesity in women was only with very weak evidence (p = 0.042) but similar pattern as in men was also observed (higher proportion of ex- and current smokers in overweight or obese women). No difference was found for length of physical activity across obesity levels in men, while in women, somewhat weak inverse association (p=0.038) was confirmed. Interestingly, predefined carbohydrate eating patterns were seemingly not associated with BMI in men (p = 0.072) or in women (p = 0.253).

TABLE 3.10: Associations of carbohydrate eating patterns with BMI in the UK adults, with or without diabetes. (NDNS RP 2008/09-15/16, sample size = 6155)

| Model | Carbohydrate eating patterns | | | | |
	Low	Moderate	P value[*]	High	P value[*]
Men (n = 2537)					
BMI					
Crude model	–	-0.78 (-1.62, 0.06)	0.068	-0.28 (-0.96, 0.41)	0.426
Model 1[†]	–	-0.20 (-1.06, 0.66)	0.654	-0.43 (-1.13, 0.26)	0.220
BMI in non-diabetics					
Crude model	–	-0.65 (-1.49, 0.19)	0.127	-0.21 (-0.89, 0.48)	0.557
Model 1[†]	–	-0.10 (-0.97, 0.77)	0.820	-0.39 (-1.10, 0.31)	0.269
Women (n = 3618)					
BMI					
Crude model	–	-0.30 (-1.18, 0.57)	0.496	-0.76 (-1.44, -0.82)	0.028
Live with partner[‡]	–	-0.93 (-2.33, 0.46)	0.188	-1.76 (-2.78, -0.73)	0.001
Live alone[‡]	–	1.17 (-0.35, 2.70)	0.132	0.57 (-0.58, 1.719	0.332
BMI in non-diabetics					
Crude model	–	-0.24 (-1.12, 0.65)	0.601	-0.71 (-1.39, -0.03)	0.040
Live with partner[‡]	–	-0.86 (-2.28, 0.55)	0.232	-1.62 (-2.65, -0.58)	0.002
Live alone[‡]	–	1.22 (-0.34, 2.78)	0.124	0.43 (-0.71, 1.56)	0.462

Note:

Abbreviations: BMI body mass index; NDNS RP, national diet and nutrition survey rolling programme. Diabetes was defined by A1C > 6.5%.

[*] P values were obtained from wald tests from linear regression models.

[†] Adjusted for age, live with partner or not, education level, hypertension (yes or no), smoking status, total energy intake, alcohol consumption;

[‡] Adjusted for age, average household income, education level, hypertension, smoking status, total energy intake, alcohol consumption.

Results of the multivariable linear regression analyses showed inverse associations between latent classes of carbohydrate eating patterns and BMI among men **(Table 3.10)**. However, the 95%CI of the regression coefficients were all wide and included the null value 0, indicating no statistically supported evidence for the inverse association. However, evidence of interaction effect was found in whether live with a partner or not on the association between carbohydrate eating patterns and BMI (p for interaction = 0.014 and 0.036 for women in total and without diabetes). For women who were living with her partner, latent classes of carbohydrate eating patterns were negatively associated with BMI. Compared with women eating low carbohydrate food pattern, women having a high carbohydrate eating pattern were averagely associated with 1.76 kg/m^2 lower BMI after adjustment of age, average household income, education level, smoking status, total energy intake and alcohol consumption. 95%CI for the adjusted BMI difference was 0.73 to 2.78 kg/m^2, p = 0.001. After excluding diabetic women, BMI was still 1.62 kg/m^2 (95%CI: 0.58, 2.65, p = 0.002) smaller in high carbohydrate eaters versus low carbohydrate eaters on average. On the contrary, latent classes of carbohydrate eating patterns were positively associated with BMI in women who were living by themselves, although the regression coefficients were not statistically different from 0.

Similarly, when looking at the association between carbohydrate eating pattern and abdominal obesity (WC), men who were classified as moderate carbohydrate eaters were found to had about 3 cm (95%CI: 0.52, 5.49 cm, p = 0.018) smaller WC compared with low carbohydrate eaters in the crude model **(Table 3.11)**. However, after adjustment of age, living with partner or not, average household income, education level, hypertension, smoking status, total energy intake, and alcohol consumption, the association attenuated to no difference. The interaction effect of whether living with partner or not on the association between carbohydrate eating patterns and WC was again found in women in total (p for interaction = 0.009) and without diabetes (p for interaction = 0.012). Among women who were living with their partners, high carbohydrate eaters had 4.71 cm (95%CI: 2.43, 7.00, p < 0.001) smaller WC on average compared with low carbohydrate eaters. The association remained when restricted the sample to non-diabetic women (-3.74 cm, 95% CI: -5.97, -1.51, p = 0.001). However, for women who were living alone, moderate carbohydrate eaters had 3.17 cm (95%CI: 0.05, 6.30, p = 0.047) larger WC on average compared with low carbohydrate eaters. The evidence for the positive association between moderate carbohydrate eaters and WC in women became weaker but with borderline significant when excluding diabetic women (3.08 cm, 95%CI: -0.09, 6.25, p = 0.057).

TABLE 3.11: Associations of carbohydrate eating patterns with WC in the UK adults , with or without diabetes. (NDNS RP 2008/09-15/16, sample size = 6155)

	Carbohydrate eating patterns				
Model	**Low**	**Moderate**	*P* **value**[*]	**High**	*P* **value**[*]
Men (n = 2537)					
WC					
Crude model	–	-3.00 (-5.49, -0.52)	0.018	-0.90 (-2.84, 1.04)	0.364
Model 1[†]	–	1.06 (-1.50, 3.64)	0.415	-1.55 (-3.42, 0.31)	0.103
WC in non-diabetics					
Crude model	–	-2.51 (-5.00, -0.21)	0.048	-0.51 (-2.47, 1.44)	0.606
Model 1[†]	–	1.42 (-1.17, 4.01)	0.283	-1.29 (-3.18, 0.60)	0.181
Women (n = 3618)					
WC					
Crude model	–	-1.28 (-3.26, 0.70)	0.206	-2.81 (-4.50, -1.12)	0.001
Live with partner[‡]	–	0.28 (-2.85, 3.41)	0.861	-4.71 (-7.00, -2.43)	< 0.001
Live alone[‡]	–	3.17 (0.05, 6.30)	0.047	0.73 (-1.84, 3.30)	0.577
WC in non-diabetics					
Crude model	–	-0.91 (-2.88, 1.07)	0.368	-2.41 (-4.06, -0.76)	0.004
Live with partner[‡]	–	1.11 (-2.02, 4.23)	0.487	-3.74 (-5.97, -1.51)	0.001
Live alone[‡]	–	3.08 (-0.09, 6.25)	0.057	0.16 (-2.36, 2.69)	0.899

Note:

Abbreviations: WC, waist circumference; NDNS RP, national diet and nutrition survey rolling programme.

Diabetes was defined by A1C > 6.5%.

[*] *P* values were obtained from wald tests from linear regression models.

[†] Adjusted for age, live with partner or not, average household income, education level, hypertension (yes or no), smoking status, total energy intake, alcohol consumption;

[‡] Adjusted for age, age^2, education level, hypertension, smoking status, total energy intake, alcohol consumption.

Chapter 4

Discussion and Conclusion

Carbohydrate eating patterns

Using multilevel LCA as a novel technique, and the NDNS RP dietary database, this project examined carbohydrate eating temporal patterns firstly in the day level, based on which, individual level carbohydrate eating patterns were also defined subsequently.

Among the dietary diaries collected, there were three distinct latent classes specifically for carbohydrate intake: 1) high probabilities of having high carbohydrate contained food across the hours of day (high percentage carbohydrate day); 2) low carbohydrate food dominant through out the hours of day (low percentage carbohydrate day); and 3) always having lunch and dinner day (regular meals day). And from these day level classifications and their features, one might anticipate that individuals who managed to follow the regular meals day might be eating a relatively healthier diet because it seemed to be a regular temporal eating habit; at this time point, we also believed that those who followed either high or low carbohydrate percentage days would probably consume higher total energy than those who followed mostly regular meals days.

However, when the MLCA extended the model to individual level, three types of persons were further defined depending on their 4-day-diary: 1) low carbohydrate eaters, who mostly followed "regular meals day"; 2) moderate carbohydrate eaters, who had similar probabilities of following either "regular meals day" or "high percentage carbohydrate day"; 3) high carbohydrate eaters, who followed "high percentage carbohydrate day" for half of their survey. For the first time, as far as we know, the day-to-day food intake pattern variation within individuals was successfully captured by MLCA models. Results from the MLCA models showed

that from the perspective of carbohydrate consumption, people were indeed changing their diet from day to day even within a short term period of survey. The MLCA models allowed the probability of following a certain type of carbohydrate eating day to vary across individuals. This properly accounted for the fact that for some people, their probability of following a type of food eating pattern during the survey could be higher/lower than that in the others. This finding also suggested that assuming a person will always follow a certain type of food intake pattern is not appropriate.

Surprisingly, low carbohydrate eaters whose dietary recordings suggested that they were mostly following a regular temporal meals pattern turned out to be consuming the highest amount of total energy among the three types of carbohydrate eaters. Detailed profiling of energy composition according to the time slots revealed that low carbohydrate eaters actually had higher proportion of energy contributed by both alcohol and fat. High percentage of fat consumption was shown in all 7 time slots, energy coming from alcohol exceeded more than one fifth of the total energy after 8 pm. These findings explained why they were actually consuming the highest energy among the three types of carbohydrate eaters. However, we also found that participants consuming low carbohydrate food had higher prevalence of diabetes, hypertension, and obesity. These health issues might possibly lead them (or advices were given from their physicians) to replace carbohydrates in their diet to other energy sources such as fat, protein, or even alcohol. Therefore, there is a possibility that they chose to follow low carbohydrate diets out of health purpose, but many of them were replacing carbohydrates with higher energy condensed food or even alcohol at night which might indeed be a public health concern.

Next, when looking into the details of the timing and composition of the energy intake among the moderate carbohydrate eaters, we realised that although these individuals did not consume as much alcohol as low carbohydrate eaters did at night, they consumed the highest amount of energy, especially during time period as late as after 10 pm. People fell into moderate carbohydrate eaters group seemed to have the tendency of having their food or meals later than high carbohydrate eaters. They consumed the highest amount of carbohydrates and also total energy among three types of carbohydrate eaters during the following time slots: 9-12 noon, 2-5 pm, 8-10 pm, and 10pm-6am. These individuals in the NDNS RP were younger, largely single, with lower average income, and lower education level. They might possibly correspond to the "late eaters" defined by previous studies (Leech et al., 2017; Mansukhani and Palla, 2018).

Finally, the high carbohydrate eaters identified by our MLCA models had the

highest absolute total amount of carbohydrate intake. Most (nearly 70%) of their total energy intake occurred during 6-9 am, 12-2 pm, and 5-8 pm time slots although their total energy consumption was not the highest (between low and moderate carbohydrate eaters). High carbohydrate eaters were also found to be the ones that consumed the least amount of energy after 8 pm. Therefore, contrary to what was anticipated at the beginning, people who followed high percentage carbohydrate days for most of their time, were actually eating a healthier diet compared with the other two eating patterns.

Our analyses looking for different temporal carbohydrate also highlighted the complexity of eating pattern behaviours in the population and the utility of exploratory, data-driven methods to objectively identify eating patterns that reflect both timing and quantities of food intake, which may have not been detected so far in the literature.

Association between carbohydrate eating patterns and health outcomes

Among men, who were classified as moderate carbohydrate eaters probably had lower odds of having hypertension, after adjustment of age, live with partner or not, educational level, BMI, smoking status and total energy intake. As discussed above, moderate carbohydrate eaters turned to have meals (or energy consumption) later in time compared with high carbohydrate eaters, but low carbohydrate eaters also consumed large amount of energy (from both fat and alcohol) at night. Therefore, considering that moderate carbohydrate eaters were younger than low carbohydrate eaters (although age was adjusted in the full models), there is probably reverse causality exists here (also due to the nature of cross-sectional study). That is, they were potentially both late eaters, however, with increased age (and so as increased health-related problems/concern) some of them modified their habits, such as quit smoking, replace carbohydrate food with other energy sources which lead to the phenomenon of lower odds of hypertension in moderate carbohydrate eaters. Although these hypotheses cannot be determined by the cross-sectional data from NDNS RP, if they were true, the energy sources they used to replace carbohydrate in their diet were apparently not very wisely chosen.

Among women, whether living with a partner became an interaction factor for the associations between carbohydrate eating patterns and BMI and abdominal obesity (WC). Directions of the associations were opposite to each other depending on

whether women were living with a partner or not. This interaction effect was more obvious when looking at abdominal obesity measurement. High carbohydrate eaters had lower BMI and WC in those lived with their partner, while moderate carbohydrate eaters had higher WC in those who lived alone after adjustment of age, education level, smoking status, total energy intake, and alcohol consumption. High carbohydrate eaters who were characterised with high and early in the day energy consumption and low fat and alcohol intake may reflect a healthier diet and lifestyle, but this might be different between women who lived alone and those who lived with their partners. It was often assumed that live alone may associated with lower diversity of food intake, and a higher likelihood of having an unhealthy dietary pattern (Hanna and Collins, 2015). Therefore, there may be differences in the actual contents consumed in high carbohydrates eating women, or there may be other social, psychological or lifestyle related factors related with living alone which were not measured or did not included in the models, so that the inverse association between high carbohydrate eating pattern and BMI or abdominal obesity were only observed among women who lived with partners, further analysis is needed. Whereas the reason why moderate carbohydrate eaters' WC was larger than low carbohydrate eaters only among women who lived alone is unknown, given that the evidence of this association was weak and borderline significant, whether it was just a false positive finding should be explored in other studies.

Limitations and strengths

There are several limitations in the current project that merits consideration.

Firstly, we ignored the order of observation days in the MLCA models. The food consumption diaries were accomplished by participants for at least 3 out of 4 consecutive days. In the multilevel analyses, these 3 or 4 days' observations were treated as if they were exchangeable in the models. Since one's diet might change according to season, day of the week (weekdays or weekends), or sometimes depends on what they had consumed the day before, exchangeability of their diaries is a strong assumption and cannot be overcome by the MLCA models adopted here. Other statistical techniques, which could take the order or the longitudinal nature of the data into account, such as repeated measures latent class analysis (RMLCA), latent transition analysis (LTA) (Collins and Lanza, 2010), or latent class growth analysis (LCGA) (Davidian et al., 2008; Jung and Wickrama, 2008; Andruff et al., 2009) are not applicable for the NDNS RP dataset. For the purpose of maintaining the response rate, flexibility was allowed for the NDNS

RP participants in their choice of which day to begin their diary. But the above mentioned alternative models require that the 3 or 4 repeated measurements of food diaries to be recorded at the same time points longitudinally. Specifically, LCGA models (also called growth mixture modelling) will also need to model the change of the odds of the probability over time as a function (quadratic, or cubic) which is apparently not the objective in the current project.

Secondly, the classification of individuals to latent carbohydrate eating classes were defined by maximum posterior probability assignment rule. This approach assigns individuals to the class for which they have the highest posterior probability of membership (Nagin, 2005). The other approach is multiple pseudo-class draws (Wang, Hendricks Brown, and Bandeen-Roche, 2005), which was proposed in attempt to account for uncertainty in class assignment. However, the maximum probability rule is still believed to be able to minimize the number of incorrect assignments (Goodman, 2007). A Monte Carlo simulation study (Bray, Lanza, and Tan, 2015) also demonstrated that maximum probability assignment is less biased than multiple pseudo-class draws. Moreover, this simulation study also found that an inclusive LCA (i.e. LCA with covariates) would perform better than non-inclusive LCA, and has the potential to reduce bias of class assignment. However, whether this advantage can be extended to MLCA in the current project is unknown, and also due to the reason that the associations between the latent classes and distal outcomes (hypertension, and BMI) in our analyses are mostly still under exploration, therefore, the model for LCA with covariates can never be well-known before our analyses. Thus, as a first step, given the complexity of the NDNS RP dataset itself, we chose to conduct the MLCA modellings without any covariate in either day level or individual level models. Future studies may be advised to consider incorporating other predictors in the MLCA models to see whether the classifications in both levels can be improved or not.

Thirdly, information regarding the detailed occupations of the participants was not available and because of which we could not exclude those who had a shift work involving night work. But considering that completing a 4-consecutive-day food diary would be a large burden for the participants, those had a shift work involving night work might possibly performed relatively lower compliance and could not complete their diary for at least 3 days. But this hypothesis cannot be verified in the NDNS RP dataset. If people working with shift work were included in the NDNS RP sample, the classification in both levels as well as their associations with either hypertension or BMI would be biased. Furthermore, under-reporting due to the burden of writing down everything they eat at each occasion is also a

concern in the food diaries collected. To evaluate the influence of under-reporting in the NDNS RP sample, study team of NDNS RP conducted doubly labelled water (DLW) sub-study within the survey. The details of this sub-study can be accessed from the Appendix X of the Official Reports provided by Public Health England (Bates et al., 2014; Roberts et al., 2018; Department of Health, 2018). Briefly, in healthy adults, if, for a given period of time, energy consumed in food matches total energy expended, they are in energy balance. In NDNS RP sub-study, estimates of energy intake (EI) from the four-day diary were compared with measurements of total energy expenditure (TEE) using the DLW technique in a sub-sample of survey participants. Results of the sub-study showed reasonable agreement between EI and TEE (overall ratio: 0.73). Reasons such as misreporting of actual consumption; under-reported or modified usual intake due to the burden of the survey/DLW sub-study are potential factors may contributed to under-reporting of their EI. We cannot extrapolate the estimated under-reporting to the whole sample since other individuals' diet might be affected by under-reporting differentially.

Lastly, it should be noted that findings from data-driven, exploratory methods may not be generalizable to populations in other countries. Since the carbohydrate eating patterns in both day level and individual level probably only represent the socio-cultural and life-style characteristics of the adult population in the UK. Further researches are warranted to explore and better understand the eating patterns in other populations. Besides, the cross-sectional study design also refrained us from deducing any causal effect between the carbohydrate eating patterns and hypertension or BMI.

Strengths of this study include the large, nationally representative sample of men and women in the UK. We applied a novel, objective approach using MLCA to examine the carbohydrate eating patterns while using standardised criteria to determine the numbers of latent classes. The process of finding the classifications through model-based, data-driven procedure minimises reliance on researchers' preconcieved notions of eating patterns. Eating patterns were determined from 4 consecutive days of well-designed, fully-examined food diaries. MLCA correctly accounted for the multilevel structure of the data which means that the 4-day diet diaries were nested within the participants. Our findings also captured the day-to-day variation of following different carbohydrate eating patterns within individuals. Moreover, the examination of associations between the individual level carbohydrate eating patterns and hypertension, BMI, abdominal obesity were conducted correctly considering the probability of participant selection and the

clustered survey design. Health outcomes such as weight, height, waist circumferences, and blood pressure, diabetes status were measured by trained nurses and blood tests which could also minimise the misclassification and under-reporting of these measurements.

Conclusions

We have successfully defined carbohydrate eating patterns in the general population in the UK adults using the NDNS RP database in both observation day level and individual level. Low carbohydrate eaters turned to have more energy that contributed by both fat and alcohol. Moderate carbohydrate eaters consumed the lowest total energy, while they had the tendency of having meals later in time-of-day. High carbohydrate eaters consumed most of their carbohydrate as well as energy earlier in time-of-day. These dietary patterns specifically for carbohydrate intake were found to be differed by timing, quantity and contributions to energy consumption. Compared with low carbohydrate eaters, men had moderate carbohydrate eating pattern may associated with a lower prevalence of hypertension, women in this latent class who lived alone may associated with a larger waist circumferences. Among women who lived with partners, high carbohydrate eating pattern was associated with both lower BMI and smaller waist circumferences. Longitudinal designed studies are needed in finding whether the eating patterns themselves are changing over time and in investigating how such circadian eating patterns may relate to the change of blood pressure, obesity, and even mortality.

Bibliography

Alhazmi, Amani et al. (2012). "Macronutrient intakes and development of type 2 diabetes: a systematic review and meta-analysis of cohort studies". In: *Journal of the American College of Nutrition* 31.4, pp. 243–258.

Almoosawi, S et al. (2013). "Time-of-day and nutrient composition of eating occasions: prospective association with the metabolic syndrome in the 1946 British birth cohort". In: *International Journal of Obesity* 37.5, p. 725.

Almoosawi, S et al. (2016). "Chrono-nutrition: a review of current evidence from observational studies on global trends in time-of-day of energy intake and its association with obesity". In: *Proceedings of the Nutrition Society* 75.4, pp. 487–500.

Andruff, Heather et al. (2009). "Latent Class Growth Modelling: A Tutorial". In: *Tutorials in Quantitative Methods for Psychology* 5.1, pp. 11–24.

Archer, Kellie J, Stanley Lemeshow, et al. (2006). "Goodness-of-fit test for a logistic regression model fitted using survey sample data". In: *Stata Journal* 6.1, pp. 97–105.

Asher, Gad and Paolo Sassone-Corsi (2015). "Time for food: the intimate interplay between nutrition, metabolism, and the circadian clock". In: *Cell* 161.1, pp. 84–92.

Bates, Beverley et al. (2014). *National Diet and Nutrition Survey: Results from Years 1, 2, 3 and 4 (combined) of the Rolling Programme (2008/2009-2011/2012): A survey carried out on behalf of Public Health England and the Food Standards Agency*. Public Health England.

Besson, Herve et al. (2009). "Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults." In: *The American Journal of Clinical Nutrition* 91.1, pp. 106–114.

Bray, Bethany C, Stephanie T Lanza, and Xianming Tan (2015). "Eliminating bias in classify-analyze approaches for latent class analysis". In: *Structural Equation Modeling* 22.1, pp. 1–11.

Collins, L.M. and S.T. Lanza (2010). *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. Wiley Series in Probability and Statistics. Wiley.

Davidian, Marie et al. (2008). "Growth mixture modeling: Analysis with non-Gaussian random effects". In: *Longitudinal Data Analysis*. Chapman and Hall/CRC, pp. 157–180.

De Bacquer, Dirk et al. (2009). "Rotating shift work and the metabolic syndrome: a prospective study". In: *International Journal of Epidemiology* 38.3, pp. 848–854.

Department of Health (2018). *National Diet and Nutrition Survey Rolling Programme*. https://www.gov.uk/government/collections/national-diet-and-nutrition-survey.

Finch, W Holmes and Brian F French (2014). "Multilevel latent class analysis: Parametric and nonparametric models". In: *The Journal of Experimental Education* 82.3, pp. 307–333.

Garaulet, Marta and Purificación Gómez-Abellán (2014). "Timing of food intake and obesity: a novel association". In: *Physiology & Behavior* 134, pp. 44–50.

Goodman, Leo A (2007). "On the Assignment of Individuals to Latent Classes". In: *Sociological Methodology* 37.1, pp. 1–22.

Hanna, Katherine L and Peter F Collins (2015). "Relationship between living alone and food and nutrient intake". In: *Nutrition Reviews* 73.9, pp. 594–611.

Henry, Kimberly L and Bengt Muthén (2010). "Multilevel latent class analysis: An application of adolescent smoking typologies with individual and contextual predictors". In: *Structural Equation Modeling* 17.2, pp. 193–215.

Johnston, Jonathan D (2014). "Physiological responses to food intake throughout the day". In: *Nutrition Research Reviews* 27.1, pp. 107–118.

Jung, Tony and KAS Wickrama (2008). "An introduction to latent class growth analysis and growth mixture modeling". In: *Social and personality psychology compass* 2.1, pp. 302–317.

Leech, Rebecca M et al. (2017). "Temporal eating patterns: a latent class analysis approach". In: *International Journal of Behavioral Nutrition and Physical Activity* 14.1, p. 3.

Lo, Yungtai, Nancy R Mendell, and Donald B Rubin (2001). "Testing the number of components in a normal mixture". In: *Biometrika* 88.3, pp. 767–778.

Mansukhani, R and L Palla (2018). "Investigating eating time patterns in UK adults from The 2008–2012 National Diet and Nutrition Survey". In: *Proceedings of the Nutrition Society* 77.OCE1.

Muthén, Bengt and Tihomir Asparouhov (2009). "Multilevel regression mixture analysis". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 172.3, pp. 639–657.

Muthén, Linda K and Bengt O Muthén (2017). *Mplus: Statistical analysis with latent variables: User's guide*. Muthén & Muthén Los Angeles.

Nagin, Daniel S (2005). *Group-based modeling of development*. Harvard University Press.

NatCen Social Research (2018). *National Diet and Nutrition Survey Years 1-8, 2008/09-2015/16*. http://doi.org/10.5255/UKDA-SN-6533-8.

Nylund, Karen L, Tihomir Asparouhov, and Bengt O Muthén (2007). "Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study". In: *Structural Equation Modeling* 14.4, pp. 535–569.

Office for National Statistics (2018). *Mid 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, and 2016 Population Estimates*. https : / / www . ons . gov . uk / peoplepopulationandcommunity / populationandmigration / populationestimates.

Pan, An et al. (2011). "Rotating night shift work and risk of type 2 diabetes: two prospective cohort studies in women". In: *PLoS Medicine* 8.12, e1001141.

Pregibon, Daryl (1980). "Goodness of link tests for generalized linear models". In: *Applied Statistics*, pp. 15–14.

R core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Roberts, Caireen et al. (2018). "National Diet and Nutrition Survey: results from years 7 and 8 (combined) of the Rolling Programme (2014/2015–2015/2016)". In:

Smithers, Gillian (1993). "MAFF's nutrient databank". In: *Nutrition & Food Science* 93.2, pp. 16–19.

Snijders, T.A.B. and R.J. Bosker (2011). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. SAGE Publications.

StataCorp LLC (2017). *Stata Statistical Software: Release 15*. Version 15.1. URL: https://www.stata.com/.

Uemura, Mayu et al. (2015). "Breakfast skipping is positively associated with incidence of type 2 diabetes mellitus: evidence from the Aichi Workers' Cohort Study". In: *Journal of Epidemiology* 25.5, pp. 351–358.

Van Horn, M Lee et al. (2008). "Using multilevel mixtures to evaluate intervention effects in group randomized trials". In: *Multivariate Behavioral Research* 43.2, pp. 289–326.

Vermunt, Jeroen K. (2003). *Multilevel Latent Class Models*. Vol. 33. 1, pp. 213–239. DOI: 10.1111/j.0081-1750.2003.t01-1-00131.x.

Vermunt, Jeroen K (2008). "Latent class and finite mixture models for multilevel data sets". In: *Statistical Methods in Medical Research* 17.1, pp. 33–51.

Wang, Chen-Pin, C Hendricks Brown, and Karen Bandeen-Roche (2005). "Residual Diagnostics for Growth Mixture Models: Examining the Impact of a Preventive Intervention on Multiple Trajectories of Aggressive Behavior". In: *Journal of the American Statistical Association* 100, pp. 1054–1076.

Willett, Walter (2012). "Nature of variation in diet". In: *Nutritional Epidemiology*. Oxford University Press. Chap. 3, pp. 34–48.

Appendix A

R code for importing and manipulating the data

```r
# NDNS analysis, data management -------------------------------------------
# Change the data path accordingly ------------------------------------------
setwd(".../UKDA-6533-stata11_se/stata11_se/")
library(epiDisplay)
library(plyr)
library(tidyverse)
# Read the data into memory -------------------------------------------------
library(haven)
data <- read_dta("ndns_rp_yr1-4a_foodleveldietarydata_uk.dta")
data56 <- read_dta("ndns_rp_yr5-6a_foodleveldietarydata.dta")
data78 <- read_dta("ndns_rp_yr7-8a_foodleveldietarydata.dta")
names(data)
names(data56)
names(data78)
names(data)[names(data)=="seriali"] <- "id"
names(data56)[names(data56)=="seriali"] <- "id"
names(data78)[names(data78)=="seriali"] <- "id"


# Extract the variables needed ----------------------------------------------
df14d <- data[,c(113,1,2,3,5,6,7,8,9,21, 22, 23, 24, 53, 55,
57,58,59,60,61,62,63,64,65)]
var <- names(df14d)
df56d <- data56 %>%
  select(var)
df78d <- data78 %>%
  select(var)
dfs1 <- rbind(df14d, df56d, df78d)

dfs2 <- dfs1[dfs1$Age>=19,] # keep participants who aged 19 or older
rm(data, data56, data78) # remove the unneeded big dataset

dfs2

# Calculate the time (minute and hour) when they eat --------------------

dfs2$MealTime_chr <- as.character(dfs2$MealTime)
dfs2$MealTime_hm <- unlist(strsplit(dfs2$MealTime_chr," "))[c(FALSE,
          TRUE)]
```

```r
dfs2$MealHourN <- as.numeric(unlist(strsplit(dfs2$MealTime_hm,
                      ":"))[c(TRUE, FALSE, FALSE)])
dfs2$MealMinN <- as.numeric(unlist(strsplit(dfs2$MealTime_hm,
                       ":"))[c(FALSE, TRUE, FALSE)])

dfs2$MealMinN0 <- (60*dfs2$MealHourN)+dfs2$MealMinN

dfs3 <- dfs2[order(dfs2$id,dfs2$DayNo,dfs2$MealMinN0),]

length(unique(dfs3$id)) ## number of participants = 6155


# Create a subset data with only the first observation of each participant --------

NDNS <- dfs3[!duplicated(dfs3$id), ]
with(NDNS, tab1(SurveyYear, graph = FALSE, decimal = 2))

# #SurveyYear :
#             Frequency Percent Cum. percent
# NDNS Year 1      801   13.01        13.01
# NDNS Year 2      812   13.19        26.21
# NDNS Year 3      782   12.71        38.91
# NDNS Year 4     1055   17.14        56.05
# NDNS Year 5      625   10.15        66.21
# NDNS Year 6      663   10.77        76.98
# NDNS Year 7      703   11.42        88.40
# NDNS Year 8      714   11.60       100.00
#    Total        6155  100.00       100.00

# how many men and women -----------------------------------------------

with(NDNS, tab1(Sex, graph = FALSE, decimal = 2))

# Sex :
#           Frequency Percent Cum. percent
# 1              2537   41.22        41.22 Men
# 2              3618   58.78       100.00 Women
#    Total       6155  100.00       100.00

# create a variable combine id and day No --------------------------------


dfs3 <- dfs3 %>%
  mutate(id_dy = paste(id, DayNo, sep = "D"))

# For each subject, the total energy/carbohydrate intake for each
# eating time can be calculated --------

old<-Sys.time()
Energy <- ddply(dfs3, .(id_dy, id, SurveyYear, DayNo, Age, Sex,
                    DiaryDaysCompleted, MealHourN, DayofWeek),
              summarise, Tot_Energ =sum(EnergykJ),
                  Tot_Carb = sum(Carbohydrateg),
                  Tot_Sugar = sum(Totalsugarsg),
                  Tot_Starch = sum(Starchg),
                  Tot_Fibre = sum(Englystfibreg),
                  Tot_Fat   = sum(Fatg),
                  Tot_Prot  = sum(Proteing),
```

```
                        Tot_Alc   = sum(Alcoholg),
                        Tot_NMES  = sum(Nonmilkextrinsicsugarsg))
new<-Sys.time()-old
print(new)
# Time difference of 6.429822 min

# reset the time intervals into time slots -------------------------------

### Breakfast:        6am to 9am
### morning snack:    9am to 12noon
### lunch:            12noon to 2pm
### afternoon snack: 2pm to 5pm
### dinner:           5pm to 8pm
### night snack:      8pm to 10pm
### midnight:         10pm to 6am
Energy <- Energy %>%
  mutate(TimeSlot = cut(MealHourN, breaks = c(6, 9, 12, 14, 17, 20, 22),
                                   right = FALSE))
levels(Energy$TimeSlot) <- c(levels(Energy$TimeSlot), "[22, 6)")
Energy$TimeSlot[is.na(Energy$TimeSlot)] <- "[22, 6)"
tab1(Energy$TimeSlot)
# For each subject, the total energy/carbohydrate intake for each
# time slot can be calculated --------
old<-Sys.time()
Energy <- ddply(Energy, .(id_dy, id, SurveyYear, DayNo, Age, Sex,
                          DiaryDaysCompleted, TimeSlot, DayofWeek),
                summarise,
              Tot_Energ = sum(Tot_Energ),
              Tot_Carb = sum(Tot_Carb),
              Tot_Sugar = sum(Tot_Sugar),
              Tot_Starch = sum(Tot_Starch),
              Tot_Fibre = sum(Tot_Fibre),
              Tot_Fat   = sum(Tot_Fat),
              Tot_Prot  = sum(Tot_Prot),
              Tot_Alc   = sum(Tot_Alc),
              Tot_NMES  = sum(Tot_NMES))
new<-Sys.time()-old
print(new)
# Time difference of 3.74195 mins

# Calculate the energy from total carbohydrates --------------------------
Energy <- Energy %>%
  mutate(KJcarbo = Tot_Carb*16) %>%
  mutate(CarKJpercentage = KJcarbo/Tot_Energ) %>%
  mutate(Carbo = cut(CarKJpercentage, breaks = c(0, 0.50, 2), right = FALSE))

Energy0 <- Energy[!(Energy$Tot_Energ == 0), ] # discard those eating occassion
# with 0 energy intake
write.csv(Energy0, file = "Energy_slots.csv") # for later analysis
Energy0$Carbo <- factor(Energy0$Carbo, labels = c("< 50%", ">= 50%"))

vecid<-unique(Energy0$id)

# Filter the data by observation day--------------------------------------------

dta_day1 <- Energy0 %>%
  filter(DayNo == 1) %>%
  select(c("id", "id_dy", "Age", "Sex",
```

```r
"DayofWeek",   "TimeSlot", "Carbo")) %>%
   mutate(DayofWeek = factor(DayofWeek,
levels = c("Monday", "Tuesday",
"Wednesday", "Thursday",
"Friday", "Saturday", "Sunday")))

dta_day2 <- Energy0 %>%
   filter(DayNo == 2) %>%
   select(c("id", "id_dy", "Age", "Sex",
"DayofWeek",   "TimeSlot", "Carbo")) %>%
   mutate(DayofWeek = factor(DayofWeek,
levels = c("Monday", "Tuesday",
"Wednesday", "Thursday",
"Friday", "Saturday", "Sunday")))

dta_day3 <- Energy0 %>%
   filter(DayNo == 3) %>%
   select(c("id", "id_dy", "Age", "Sex",
"DayofWeek",   "TimeSlot", "Carbo")) %>%
   mutate(DayofWeek = factor(DayofWeek,
levels = c("Monday", "Tuesday",
"Wednesday", "Thursday",
"Friday", "Saturday", "Sunday")))

dta_day4 <- Energy0 %>%
   filter(DayNo == 4) %>%
   select(c("id", "id_dy", "Age", "Sex",
"DayofWeek",   "TimeSlot", "Carbo")) %>%
StringTok  mutate(DayofWeek = factor(DayofWeek,
levels = c("Monday", "Tuesday",
"Wednesday", "Thursday",
"Friday", "Saturday", "Sunday")))

vecid1<-unique(dta_day1$id) # n = 6153
vecid2<-unique(dta_day2$id) # n = 6153
vecid3<-unique(dta_day3$id) # n = 6151
vecid4<-unique(dta_day4$id) # n = 6026
Noday1 <- setdiff(vecid, vecid1) # two subjects did not have day 1 data
Noday2 <- setdiff(vecid, vecid2) # two subjects did not have day 2 data
Noday3 <- setdiff(vecid, vecid3) # four subjects did not have day 3 data
Noday4 <- setdiff(vecid, vecid4) # 129 subjects did not have day 4 data

# Long to wide data -------------------------------------------------------
dta_d1_wide <- dta_day1 %>%
   spread(key = TimeSlot,
value = Carbo)
head(dta_d1_wide)
names(dta_d1_wide)[6:12] <- c("H6_9", "H9_12", "H12_14", "H14_17",
"H17_20", "H20_22", "H22_6")
names(dta_d1_wide)
dta_d2_wide <- dta_day2 %>%
   spread(key = TimeSlot,
value = Carbo)
head(dta_d2_wide)
names(dta_d2_wide)[6:12] <- c("H6_9", "H9_12", "H12_14", "H14_17",
"H17_20", "H20_22", "H22_6")
names(dta_d2_wide)
dta_d3_wide <- dta_day3 %>%
```

```r
    spread(key = TimeSlot,
value = Carbo)
head(dta_d3_wide)
names(dta_d3_wide)[6:12] <- c("H6_9", "H9_12", "H12_14", "H14_17",
"H17_20", "H20_22", "H22_6")
names(dta_d3_wide)
dta_d4_wide <- dta_day4 %>%
    spread(key = TimeSlot,
value = Carbo)
head(dta_d4_wide)
names(dta_d4_wide)[6:12] <- c("H6_9", "H9_12", "H12_14", "H14_17",
"H17_20", "H20_22", "H22_6")
names(dta_d4_wide)
# recode NA to not eating ---------------------------------------------
for (i in 6:ncol(dta_d1_wide))
if(is.factor(dta_d1_wide[,i]))
levels(dta_d1_wide[,i]) <- c("1", "2", "0")
dta_d1_wide[is.na(dta_d1_wide)] <- "0"
for (i in 6:ncol(dta_d2_wide))
if(is.factor(dta_d2_wide[,i]))
levels(dta_d2_wide[,i]) <- c("1", "2",  "0")
dta_d2_wide[is.na(dta_d2_wide)] <- "0"
for (i in 6:ncol(dta_d3_wide))
if(is.factor(dta_d3_wide[,i]))
levels(dta_d3_wide[,i]) <-  c("1", "2", "0")
dta_d3_wide[is.na(dta_d3_wide)] <- "0"
for (i in 6:ncol(dta_d4_wide))
if(is.factor(dta_d4_wide[,i]))
levels(dta_d4_wide[,i]) <-  c("1", "2", "0")
dta_d4_wide[is.na(dta_d4_wide)] <- "0"
dta_all <- rbind(dta_d1_wide, dta_d2_wide, dta_d3_wide, dta_d4_wide)
dta_all <- dta_all[order(dta_all$id,dta_all$id_dy),]
# Export the data for Mplus -------------------------------------------
write_csv(dta_all, path = "NDNS_Tslots.csv")
write_delim(dta_all, "NDNS_Tslots.dat", na = ".", delim = " ")
```

Appendix B

Mplus code and output for Multilevel LCA models

```
Mplus VERSION 7.4
MUTHEN & MUTHEN
07/28/2018   9:55 AM

INPUT INSTRUCTIONS

TITLE:    3-class at level 1 (CW), 3-classes at level 2 (CB) random effects model
ordered polytomous variables for carb intake at each time slot over four
days of NDNS survey 2008/09 - 2015/16
variable 0 = not eating
1 = eating & carb provided < 50% calorie
2 = eating & carb provided >= 50% calorie

DATA:     File is H:\summer_project\Mplus\TimeSlots\NDNS_Tslots.dat;


VARIABLE: NAMES = id id_dy Age Sex H6_9 H9_12 H12_14 H14_17 H17_20
H20_22 H22_6;

USEVAR = H6_9 H9_12 H12_14 H14_17 H17_20
H20_22 H22_6;

auxiliary = Age Sex;

CATEGORICAL =  H6_9 H9_12 H12_14 H14_17 H17_20
H20_22 H22_6;

CLUSTER = id;

IDVARIABLE = id_dy;

BETWEEN = CB;

WITHIN = H6_9 H9_12 H12_14 H14_17 H17_20
H20_22 H22_6;

CLASSES = CB(3) CW(3);

MISSING are .;


ANALYSIS:
```

```
type = mixture twolevel;
starts = 50 25;
process = 8(starts);



MODEL:
%within%
%overall%
%between%
%overall%
CW ON CB;




Savedata:
file is H:\summer_project\Mplus\TimeSlots\Multilevel\NDNSslot_CW3CB3.txt;
save is cprob;
format is free;
```

3-class at level 1 (CW), 3-classes at level 2 (CB) random effects model
ordered polytomous variables for carb intake at each time slot over four
days of NDNS survey 2008/09 - 2015/16
variable 0 = not eating
1 = eating & carb provided < 50% calorie
2 = eating & carb provided >= 50% calorie

SUMMARY OF ANALYSIS

Number of groups   1
Number of observations   24483

Number of dependent variables   7
Number of independent variables   0
Number of continuous latent variables   0
Number of categorical latent variables   2

Observed dependent variables

Binary and ordered categorical (ordinal)
H6_9   H9_12   H12_14   H14_17   H17_20   H20_22
H22_6

Observed auxiliary variables
AGE   SEX

Categorical latent variables
CB   CW

Variables with special functions

Cluster variable   ID
ID variable   ID_DY

Within variables
H6_9   H9_12   H12_14   H14_17   H17_20   H20_22
H22_6

```
Estimator                                                     MLR
Information matrix                                        OBSERVED
Optimization Specifications for the Quasi-Newton Algorithm for
Continuous Outcomes
Maximum number of iterations                                   100
Convergence criterion                                   0.100D-05
Optimization Specifications for the EM Algorithm
Maximum number of iterations                                   500
Convergence criteria
Loglikelihood change                                    0.100D-02
Relative loglikelihood change                           0.100D-05
Derivative                                              0.100D-02
Optimization Specifications for the M step of the EM Algorithm for
Categorical Latent variables
Number of M step iterations                                     1
M step convergence criterion                            0.100D-02
Basis for M step termination                            ITERATION
Optimization Specifications for the M step of the EM Algorithm for
Censored, Binary or Ordered Categorical (Ordinal), Unordered
Categorical (Nominal) and Count Outcomes
Number of M step iterations                                     1
M step convergence criterion                            0.100D-02
Basis for M step termination                            ITERATION
Maximum value for logit thresholds                             15
Minimum value for logit thresholds                           -15
Minimum expected cell size for chi-square               0.100D-01
Maximum number of iterations for H1                          2000
Convergence criterion for H1                            0.100D-03
Optimization algorithm                                        EMA
Integration Specifications
Type                                                     STANDARD
Number of integration points                                  15
Dimensions of numerical integration                            0
Adaptive quadrature                                           ON
Random Starts Specifications
Number of initial stage random starts                         50
Number of final stage optimizations                           25
Number of initial stage iterations                            10
Initial stage convergence criterion                    0.100D+01
Random starts scale                                     0.500D+01
Random seed for generating random starts                       0
Parameterization                                           LOGIT
Link                                                       LOGIT
Cholesky                                                      OFF

Input data file(s)
H:\summer_project\Mplus\TimeSlots\NDNS_Tslots.dat
Input data format   FREE


SUMMARY OF DATA

Number of missing data patterns             1
Number of y missing data patterns           0
Number of u missing data patterns           1
Number of clusters                       6155
```

COVARIANCE COVERAGE OF DATA

Minimum covariance coverage value   0.100

UNIVARIATE PROPORTIONS AND COUNTS FOR CATEGORICAL VARIABLES

```
H6_9
Category 1      0.313           7655.000
Category 2      0.184           4500.000
Category 3      0.504          12328.000
H9_12
Category 1      0.222           5447.000
Category 2      0.295           7227.000
Category 3      0.482          11809.000
H12_14
Category 1      0.195           4783.000
Category 2      0.454          11112.000
Category 3      0.351           8588.000
H14_17
Category 1      0.283           6926.000
Category 2      0.338           8277.000
Category 3      0.379           9280.000
H17_20
Category 1      0.124           3043.000
Category 2      0.582          14240.000
Category 3      0.294           7200.000
H20_22
Category 1      0.356           8722.000
Category 2      0.363           8898.000
Category 3      0.280           6863.000
H22_6
Category 1      0.666          16295.000
Category 2      0.169           4144.000
Category 3      0.165           4044.000
```


RANDOM STARTS RESULTS RANKED FROM THE BEST TO THE WORST LOGLIKELIHOOD VALUES

Final stage loglikelihood values at local maxima, seeds, and initial stage
start numbers:

```
-166348.815   153942          31
-166348.815   573096          20
-166348.815   253358          2
-166348.816   318230          46
-166348.816   246261          38
-166348.873   285380          1
-166348.908   903420          5
-166349.394   120506          45
-166349.394   966014          37
-166349.394   207896          25
-166349.395   195873          6
-166349.513   68985           17
-166349.514   366706          29
-166352.737   76974           16
-166357.057   127215          9
-166482.723   533738          11
```

```
-166495.844  645664          39
-166668.918  372176          23
```

THE BEST LOGLIKELIHOOD VALUE HAS BEEN REPLICATED. RERUN WITH AT LEAST TWICE THE
RANDOM STARTS TO CHECK THAT THE BEST LOGLIKELIHOOD IS STILL OBTAINED AND REPLICATED.

THE MODEL ESTIMATION TERMINATED NORMALLY

MODEL FIT INFORMATION

Number of Free Parameters   134

Loglikelihood

HO Value   -166348.815
HO Scaling Correction Factor   1.8182
for MLR

Information Criteria

Akaike (AIC)   332965.630
Bayesian (BIC)   334051.799
Sample-Size Adjusted BIC   333625.950
(n* = (n + 2) / 24)

MODEL RESULTS USE THE LATENT CLASS VARIABLE ORDER

CB   CW

Latent Class Variable Patterns

CB   CW
Class   Class

1   1
1   2
1   3
2   1
2   2
2   3
3   1
3   2
3   3

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASS PATTERNS
BASED ON ESTIMATED POSTERIOR PROBABILITIES

Latent Class
Pattern

1 1   4050.97975   0.16546
1 2   1561.55249   0.06378
```

```
1  3          1286.46696            0.05255
2  1          2746.94031            0.11220
2  2          3011.00217            0.12298
2  3          1341.59686            0.05480
3  1          2748.25320            0.11225
3  2          4770.55950            0.19485
3  3          2965.64876            0.12113
```

FINAL CLASS COUNTS AND PROPORTIONS FOR EACH LATENT CLASS VARIABLE
BASED ON ESTIMATED POSTERIOR PROBABILITIES

Latent Class
Variable   Class

```
CB            1          6898.99902            0.28179
              2          7099.53906            0.28998
              3         10484.46094            0.42823
CW            1          9546.17285            0.38991
              2          9343.11426            0.38162
              3          5593.71240            0.22847
```

FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASS PATTERNS
BASED ON THEIR MOST LIKELY LATENT CLASS PATTERN

Class Counts and Proportions

Latent Class
Pattern

```
1  1              4262              0.17408
1  2              1406              0.05743
1  3              1178              0.04812
2  1              2807              0.11465
2  2              2946              0.12033
2  3              1260              0.05146
3  1              2745              0.11212
3  2              5315              0.21709
3  3              2564              0.10473
```

FINAL CLASS COUNTS AND PROPORTIONS FOR EACH LATENT CLASS VARIABLE
BASED ON THEIR MOST LIKELY LATENT CLASS PATTERN

Latent Class
Variable   Class

```
CB            1              6846              0.27962
              2              7013              0.28644
              3             10624              0.43393
CW            1              9814              0.40085
              2              9667              0.39485
              3              5002              0.20431
```

CLASSIFICATION QUALITY

```
Entropy                       0.630
```

Average Latent Class Probabilities for Most Likely Latent Class Pattern (Row)
by Latent Class Pattern (Column)

Latent Class Variable Patterns

Latent Class Pattern No.	CB Class	CW Class
1	1	1
2	1	2
3	1	3
4	2	1
5	2	2
6	2	3
7	3	1
8	3	2
9	3	3

	1	2	3	4	5	6	7	8	9
1	0.720	0.091	0.073	0.016	0.032	0.004	0.005	0.033	0.025
2	0.183	0.609	0.098	0.005	0.002	0.030	0.040	0.005	0.027
3	0.211	0.084	0.629	0.008	0.005	0.007	0.011	0.036	0.009
4	0.019	0.004	0.002	0.692	0.184	0.051	0.011	0.034	0.003
5	0.042	0.001	0.001	0.158	0.709	0.045	0.001	0.035	0.009
6	0.012	0.037	0.013	0.065	0.084	0.702	0.042	0.003	0.042
7	0.011	0.029	0.004	0.012	0.002	0.022	0.641	0.126	0.153
8	0.026	0.003	0.009	0.025	0.024	0.001	0.115	0.675	0.123
9	0.046	0.024	0.004	0.003	0.010	0.018	0.079	0.174	0.642

MODEL RESULTS

	Two-Tailed Estimate	S.E.	Est./S.E.	P-Value

Within Level

Latent Class Pattern 1 1

Thresholds				
H6_9$1	-0.718	0.218	-3.294	0.001
H6_9$2	0.973	0.299	3.258	0.001
H9_12$1	-2.516	0.463	-5.433	0.000
H9_12$2	0.675	0.132	5.118	0.000
H12_14$1	-1.025	0.145	-7.057	0.000
H12_14$2	1.240	0.116	10.725	0.000
H14_17$1	-1.566	0.149	-10.520	0.000
H14_17$2	1.090	0.100	10.909	0.000
H17_20$1	-1.998	0.125	-16.000	0.000
H17_20$2	1.549	0.100	15.556	0.000
H20_22$1	-0.933	0.085	-10.914	0.000
H20_22$2	1.829	0.103	17.770	0.000
H22_6$1	0.253	0.083	3.046	0.002
H22_6$2	2.308	0.117	19.691	0.000

```
Latent Class Pattern 1 2

Thresholds
H6_9$1              -4.021        1.788       -2.249        0.025
H6_9$2              -0.115        0.259       -0.445        0.656
H9_12$1              0.167        0.373        0.448        0.654
H9_12$2              2.142        0.586        3.657        0.000
H12_14$1            -3.210        1.518       -2.115        0.034
H12_14$2             0.858        0.167        5.124        0.000
H14_17$1             0.044        0.384        0.114        0.909
H14_17$2             1.617        0.293        5.509        0.000
H17_20$1            -2.109        0.390       -5.409        0.000
H17_20$2             1.399        0.196        7.126        0.000
H20_22$1            -0.367        0.174       -2.109        0.035
H20_22$2             2.347        0.382        6.151        0.000
H22_6$1              0.754        0.259        2.912        0.004
H22_6$2              2.542        0.264        9.646        0.000

Latent Class Pattern 1 3

Thresholds
H6_9$1             -15.000        0.000      999.000      999.000
H6_9$2               2.357        0.783        3.011        0.003
H9_12$1             -1.433        0.372       -3.850        0.000
H9_12$2             -0.604        0.279       -2.166        0.030
H12_14$1            -1.988        0.257       -7.749        0.000
H12_14$2             0.524        0.125        4.209        0.000
H14_17$1            -1.027        0.232       -4.436        0.000
H14_17$2             0.274        0.131        2.087        0.037
H17_20$1            -2.665        0.310       -8.605        0.000
H17_20$2             0.707        0.112        6.322        0.000
H20_22$1            -0.527        0.152       -3.462        0.001
H20_22$2             0.702        0.138        5.102        0.000
H22_6$1              1.119        0.185        6.062        0.000
H22_6$2              1.748        0.183        9.544        0.000

Latent Class Pattern 2 1

Thresholds
H6_9$1               1.663        0.199        8.370        0.000
H6_9$2               1.839        0.198        9.274        0.000
H9_12$1             -2.150        0.281       -7.643        0.000
H9_12$2             -0.869        0.140       -6.190        0.000
H12_14$1            -1.978        0.191      -10.349        0.000
H12_14$2             0.323        0.078        4.139        0.000
H14_17$1             0.237        0.183        1.293        0.196
H14_17$2             0.782        0.123        6.352        0.000
H17_20$1            -2.936        0.428       -6.853        0.000
H17_20$2             0.632        0.081        7.807        0.000
H20_22$1             0.028        0.142        0.194        0.846
H20_22$2             0.868        0.086       10.145        0.000
H22_6$1              0.658        0.109        6.010        0.000
H22_6$2              1.326        0.100       13.215        0.000

Latent Class Pattern 2 2

Thresholds
H6_9$1               1.640        0.171        9.619        0.000
H6_9$2               1.906        0.179       10.678        0.000
```

```
H9_12$1               -1.954       0.347      -5.636       0.000
H9_12$2               -0.360       0.127      -2.842       0.004
H12_14$1              -0.016       0.189      -0.084       0.933
H12_14$2               0.948       0.135       7.029       0.000
H14_17$1              -1.906       0.301      -6.327       0.000
H14_17$2               0.371       0.080       4.614       0.000
H17_20$1              -0.812       0.116      -7.030       0.000
H17_20$2               0.910       0.089      10.259       0.000
H20_22$1              -0.742       0.089      -8.318       0.000
H20_22$2               0.998       0.085      11.705       0.000
H22_6$1                0.298       0.083       3.608       0.000
H22_6$2                1.337       0.099      13.475       0.000

Latent Class Pattern 2 3

Thresholds
H6_9$1                -1.072       0.500      -2.144       0.032
H6_9$2                -0.309       0.346      -0.892       0.372
H9_12$1                2.441       1.044       2.339       0.019
H9_12$2                3.599       1.983       1.815       0.069
H12_14$1              -1.029       0.211      -4.880       0.000
H12_14$2               0.603       0.123       4.913       0.000
H14_17$1              -0.010       0.243      -0.041       0.967
H14_17$2               0.784       0.157       4.977       0.000
H17_20$1              -0.953       0.203      -4.684       0.000
H17_20$2               0.779       0.135       5.784       0.000
H20_22$1              -0.105       0.210      -0.500       0.617
H20_22$2               1.203       0.135       8.914       0.000
H22_6$1                0.582       0.299       1.950       0.051
H22_6$2                1.370       0.206       6.653       0.000

Latent Class Pattern 3 1

Thresholds
H6_9$1                -4.593       1.699      -2.703       0.007
H6_9$2                -2.975       0.428      -6.957       0.000
H9_12$1               -0.322       0.207      -1.553       0.120
H9_12$2                0.398       0.363       1.095       0.274
H12_14$1              -5.060       3.668      -1.380       0.168
H12_14$2               0.307       0.100       3.080       0.002
H14_17$1               0.186       0.530       0.351       0.726
H14_17$2               0.317       0.245       1.295       0.195
H17_20$1              -4.019       0.957      -4.199       0.000
H17_20$2               0.747       0.093       7.987       0.000
H20_22$1              -0.233       0.132      -1.767       0.077
H20_22$2               0.607       0.109       5.571       0.000
H22_6$1                1.304       0.146       8.918       0.000
H22_6$2                1.850       0.160      11.579       0.000

Latent Class Pattern 3 2

Thresholds
H6_9$1                -1.232       0.195      -6.305       0.000
H6_9$2                -0.858       0.169      -5.068       0.000
H9_12$1               -4.377       1.937      -2.260       0.024
H9_12$2               -1.488       0.316      -4.717       0.000
H12_14$1              -1.727       0.227      -7.611       0.000
H12_14$2               0.302       0.082       3.666       0.000
H14_17$1              -1.834       0.237      -7.730       0.000
```

```
H14_17$2          -0.294     0.186     -1.582     0.114
H17_20$1          -2.588     0.487     -5.313     0.000
H17_20$2           0.631     0.062     10.187     0.000
H20_22$1          -0.920     0.078    -11.852     0.000
H20_22$2           0.462     0.073      6.308     0.000
H22_6$1            0.640     0.119      5.361     0.000
H22_6$2            1.162     0.129      9.039     0.000


Latent Class Pattern 3 3

Thresholds
H6_9$1            -4.941     5.813     -0.850     0.395
H6_9$2            -2.680     0.887     -3.024     0.002
H9_12$1           -0.765     0.640     -1.195     0.232
H9_12$2            1.164     0.920      1.265     0.206
H12_14$1          -1.415     0.439     -3.226     0.001
H12_14$2           0.566     0.085      6.626     0.000
H14_17$1          -2.052     0.650     -3.158     0.002
H14_17$2           0.612     0.210      2.909     0.004
H17_20$1          -1.627     0.427     -3.810     0.000
H17_20$2           0.713     0.103      6.935     0.000
H20_22$1          -0.850     0.329     -2.585     0.010
H20_22$2           0.685     0.134      5.104     0.000
H22_6$1            1.237     0.195      6.349     0.000
H22_6$2            1.893     0.179     10.582     0.000


Between Level

Categorical Latent Variables

Within Level

Intercepts
CW#1              -0.076     0.366     -0.208     0.835
CW#2               0.475     0.309      1.539     0.124

Between Level

CW#1       ON
CB#1               1.223     0.473      2.585     0.010
CB#2               0.793     0.441      1.796     0.073

CW#2       ON
CB#1              -0.282     0.535     -0.526     0.599
CB#2               0.333     0.455      0.733     0.464

Means
CB#1              -0.417     0.100     -4.178     0.000
CB#2              -0.386     0.067     -5.770     0.000


QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix              0.428E-04
(ratio of smallest to largest eigenvalue)


SAVEDATA INFORMATION
```

Save file
H:\summer_project\Mplus\TimeSlots\Multilevel\NDNSslot_CW3CB3.txt

Order of variables

H6_9
H9_12
H12_14
H14_17
H17_20
H20_22
H22_6
ID_DY
AGE
SEX
CPROB1
CPROB2
CPROB3
CPROB4
CPROB5
CPROB6
CPROB7
CPROB8
CPROB9
CB
CW
MLCJOINT
ID

Save file format   Free

Save file record length   10000

DIAGRAM INFORMATION

Mplus diagrams are currently not available for Mixture analysis.
No diagram output was produced.

Beginning Time: 09:55:10
Ending Time: 10:02:01
Elapsed Time: 00:06:51

MUTHEN & MUTHEN
3463 Stoner Ave.
Los Angeles, CA 90066

Tel: (310) 391-9971
Fax: (310) 391-8971
Web: www.StatModel.com
Support: Support@StatModel.com

Copyright (c) 1998-2015 Muthen & Muthen

Appendix C

Other Solutions for MLCA

2 classes in day level

FIGURE C.1: 2 Classes solution in Day level.

FIGURE C.2: Multilevel Latent Class Solution, 2 classes in day level, 2 classes in individual level.

FIGURE C.3: Multilevel Latent Class Solution, 2 classes in day level, 3 classes in individual level.

FIGURE C.4: Multilevel Latent Class Solution, 2 classes in day level, 4 classes in individual level.

3 classes in day level, 4 classes in individual level

FIGURE C.5: Multilevel Latent Class Solution, 3 classes in day level, 4 classes in individual level.

4 classes in day level

FIGURE C.6: 4 Classes Solution in Day level.

FIGURE C.7: Multilevel Latent Class Solution, 4 classes in day level, 2 classes in individual level.

FIGURE C.8: Multilevel Latent Class Solution, 4 classes in day level, 3 classes in individual level.

Appendix D

R and STATA codes for processing individual level data analysis

```r
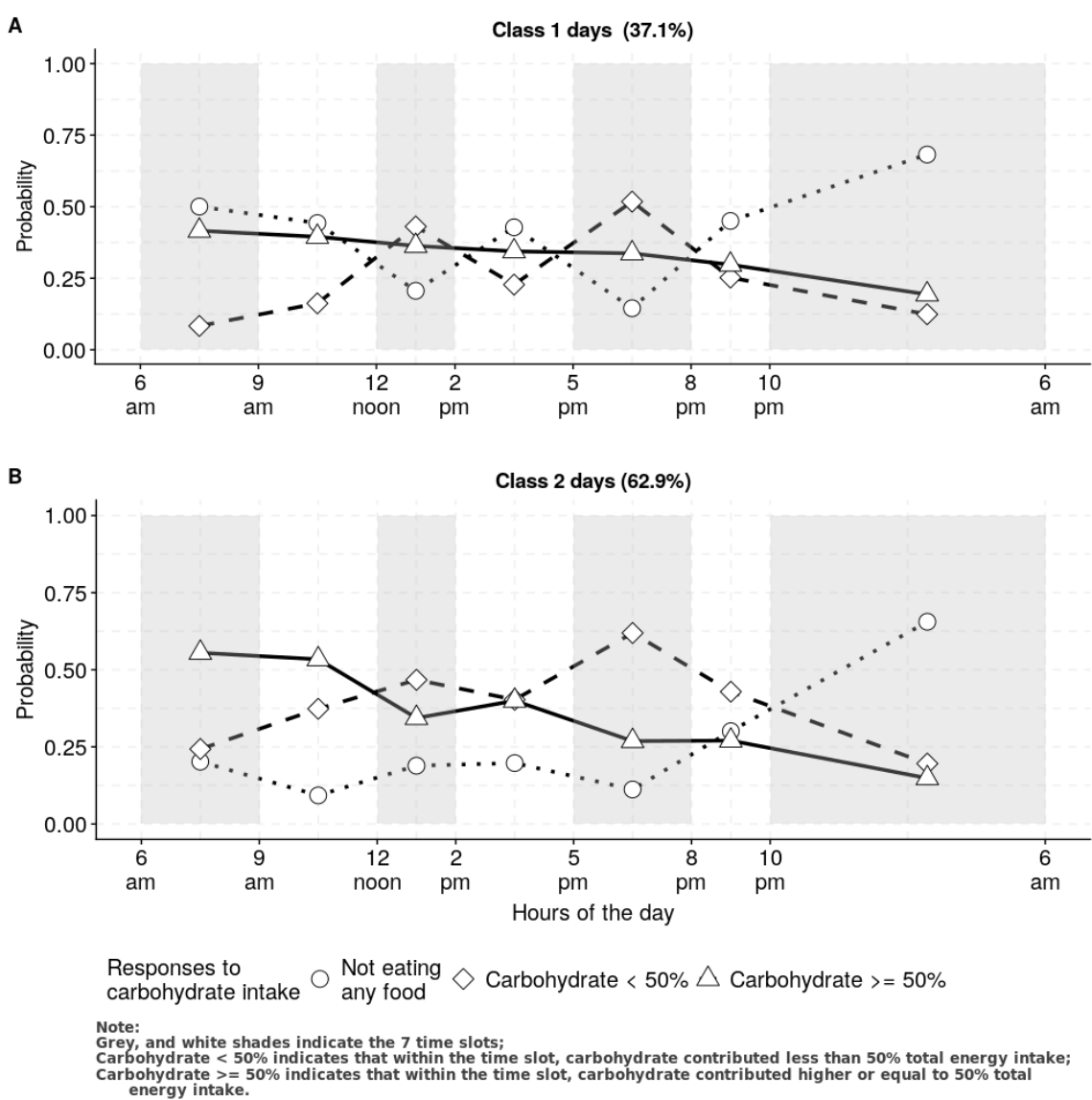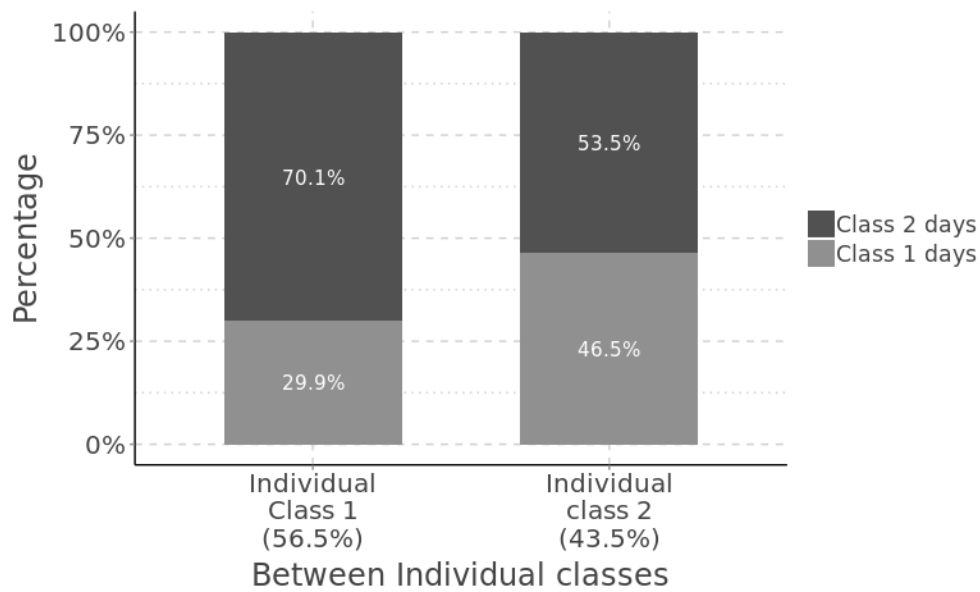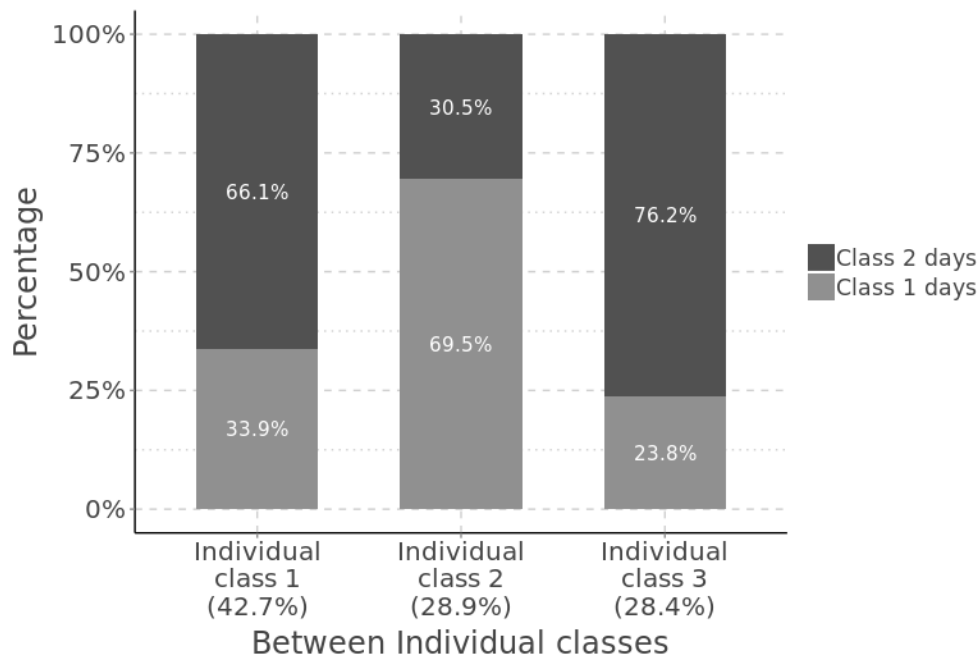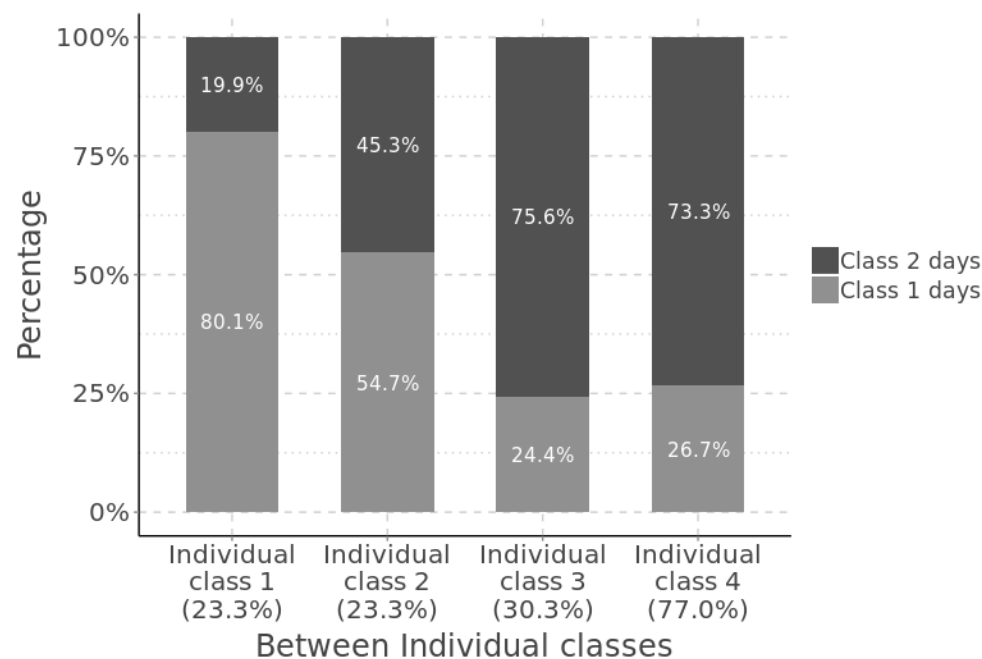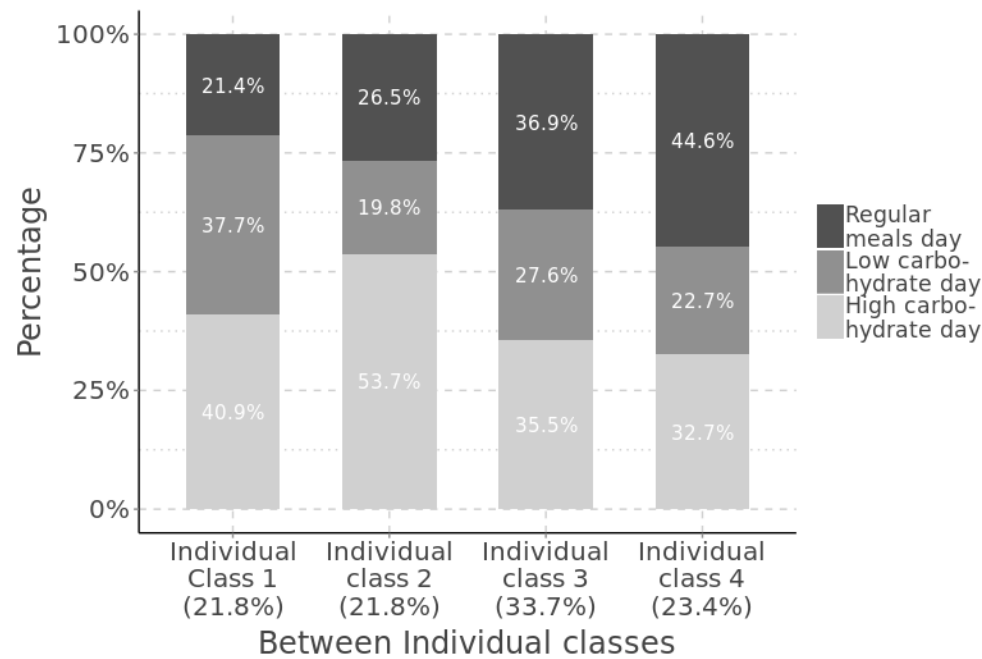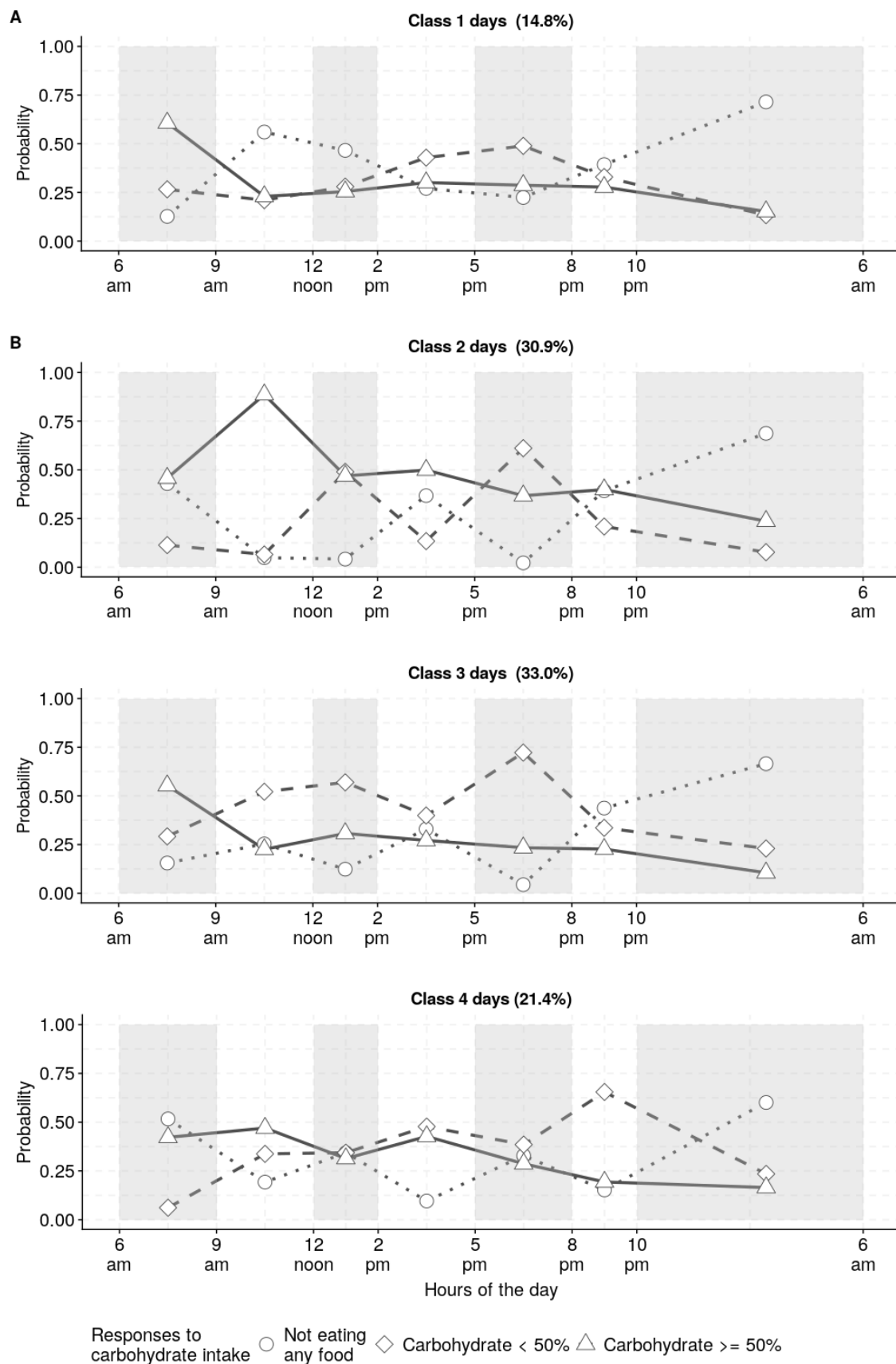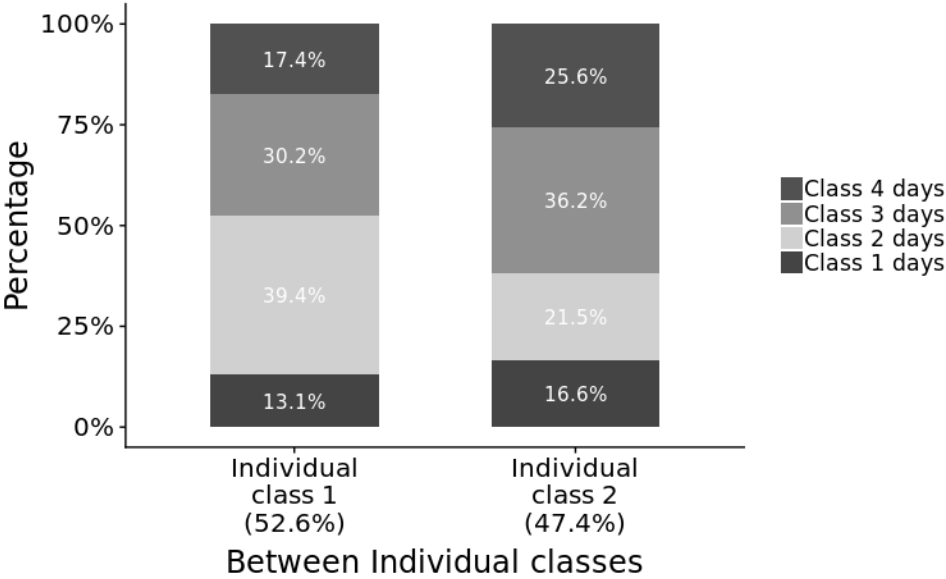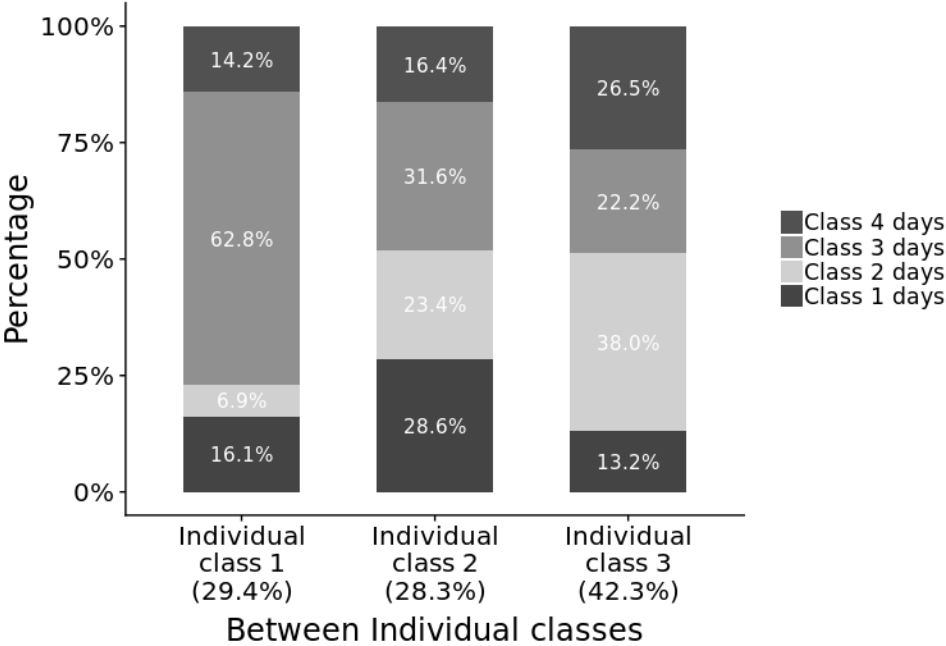library(plyr)
library(epiDisplay)
library(tidyverse)
library(dplyr)
library(readr)
library(haven)
library(naniar)

# read the individual level data sets ----------------------------------------
# Change the data path accordingly --------------------------------------
setwd(".../UKDA-6533-stata11_se/stata11_se/")
blood78 <- read_dta("ndns_rp_yr7-8a_indiv.dta")
blood56 <- read_dta("ndns_rp_yr5-6a_indiv.dta")
blood14 <- read_dta("ndns_rp_yr1-4a_indiv_uk.dta")
food14 <- read_dta("ndns_rp_yr1-4a_personleveldietarydata_uk.dta")
food56 <- read_dta("ndns_rp_yr5-6a_personleveldietarydata.dta")
food78 <- read_dta("ndns_rp_yr7-8a_personleveldietarydata.dta")
names(blood78)[names(blood78)=="seriali"] <- "ID"
names(blood56)[names(blood56)=="seriali"] <- "ID"
names(blood14)[names(blood14)=="seriali"] <- "ID"
names(food78)[names(food78)=="seriali"] <- "ID"
names(food56)[names(food56)=="seriali"] <- "ID"
names(food14)[names(food14)=="seriali"] <- "ID"


# Loading the data from Mplus output
CW3CB3 <- read_table2("../NDNSslot_CW3CB3.txt",  # change the path accordingly
                      col_names = FALSE)

names(CW3CB3) <- c("Breakfast", "Morning.snack", "Lunch",
                   "Afternoon.snack", "Dinner", "Before.bedtime.snack",
                   "Midnight.food", "ID_DAY", "AGE", "SEX", "CPROB1",
                   "CPROB2", "CPROB3", "CPROB4", "CPROB5", "CPROB6",
                   "CPROB7", "CPROB8", "CPROB9",
                   "CB", "CW", "MLCJOINT", "ID")

CW3idday <- CW3CB3 %>%
  select(ID, ID_DAY, CW, CB, MLCJOINT, AGE, SEX)

Energy_slots <- read_csv("../Energy_slots.csv") # change the path accordingly
Energy_slots <- Energy_slots %>%
  rename(ID = id)
```

```r
# Recode day level classification to keep consistency
CW3idday$CW_new <- 0
CW3idday$CW_new[CW3idday$CW == 1] <- 3
CW3idday$CW_new[CW3idday$CW == 2] <- 1
CW3idday$CW_new[CW3idday$CW == 3] <- 2

Energy_slots <- Energy_slots %>%
  left_join(CW3idday, by = c("ID", "DayofWeek"))


Energy_slots$TimeSlot <- factor(Energy_slots$TimeSlot,
                                levels = c("[6,9)", "[9,12)",
                                           "[12,14)", "[14,17)",
                                           "[17,20)", "[20,22)", "[22, 6)"))

FoodbyCB_7slots <- Energy_slots %>%
  group_by(ID, TimeSlot) %>%
  summarise(sumEnergy = sum(Tot_Energ),
            sumCarb = sum(Tot_Carb),
            sumSugar = sum(Tot_Sugar),
            sumStarch = sum(Tot_Starch),
            sumFibre = sum(Tot_Fibre),
            sumNMES = sum(Tot_NMES),
            sumFat = sum(Tot_Fat),
            sumProt = sum(Tot_Prot),
            sumAlc = sum(Tot_Alc))

# Calculate the sum of each nutrients for each time slots
Carbsum <- FoodbyCB_7slots %>%
  select(ID, TimeSlot, sumCarb) %>%
  spread(key = TimeSlot,
         value = sumCarb)

Carbsum[is.na(Carbsum)] <- 0

names(Carbsum) <- c("ID", "Carb6_9", "Carb9_12", "Carb12_14", "Carb14_17",
                    "Carb17_20", "Carb20_22", "Carb22_6")


Energysum <- FoodbyCB_7slots %>%
  select(ID, TimeSlot, sumEnergy) %>%
  spread(key = TimeSlot,
         value = sumEnergy)
Energysum[is.na(Energysum)] <- 0


names(Energysum) <- c("ID", "Energy6_9", "Energy9_12", "Energy12_14",
"Energy14_17", "Energy17_20", "Energy20_22", "Energy22_6")

Starchsum <- FoodbyCB_7slots %>%
  select(ID, TimeSlot, sumStarch) %>%
  spread(key = TimeSlot,
         value = sumStarch)
Starchsum[is.na(Starchsum)] <- 0

names(Starchsum) <- c("ID", "Starch6_9", "Starch9_12", "Starch12_14",
"Starch14_17", "Starch17_20", "Starch20_22", "Starch22_6")
```

```r
Sugarsum <- FoodbyCB_7slots %>%
  select(ID, TimeSlot, sumSugar) %>%
  spread(key = TimeSlot,
         value = sumSugar)
Sugarsum[is.na(Sugarsum)] <- 0

names(Sugarsum) <- c("ID", "Sugar6_9", "Sugar9_12", "Sugar12_14",
"Sugar14_17", "Sugar17_20", "Sugar20_22", "Sugar22_6")

Fibresum <- FoodbyCB_7slots %>%
  select(ID, TimeSlot, sumFibre) %>%
  spread(key = TimeSlot,
         value = sumFibre)
Fibresum[is.na(Fibresum)] <- 0

names(Fibresum) <- c("ID", "Fibre6_9", "Fibre9_12", "Fibre12_14",
"Fibre14_17", "Fibre17_20", "Fibre20_22", "Fibre22_6")

NMESsum <- FoodbyCB_7slots %>%
  select(ID, TimeSlot, sumNMES) %>%
  spread(key = TimeSlot,
         value = sumNMES)
NMESsum[is.na(NMESsum)] <- 0

names(NMESsum) <- c("ID", "NMES6_9", "NMES9_12", "NMES12_14",
"NMES14_17", "NMES17_20", "NMES20_22", "NMES22_6")

Fatsum <- FoodbyCB_7slots %>%
  select(ID, TimeSlot, sumFat) %>%
  spread(key = TimeSlot,
         value = sumFat)
Fatsum[is.na(Fatsum)] <- 0

names(Fatsum) <- c("ID", "Fat6_9", "Fat9_12", "Fat12_14", "Fat14_17",
                "Fat17_20", "Fat20_22", "Fat22_6")

Protsum <- FoodbyCB_7slots %>%
  select(ID, TimeSlot, sumProt) %>%
  spread(key = TimeSlot,
         value = sumProt)
Protsum[is.na(Protsum)] <- 0

names(Protsum) <- c("ID", "Prot6_9", "Prot9_12", "Prot12_14", "Prot14_17",
                "Prot17_20", "Prot20_22", "Prot22_6")

Alcsum <- FoodbyCB_7slots %>%
  select(ID, TimeSlot, sumAlc) %>%
  spread(key = TimeSlot,
         value = sumAlc)
Alcsum[is.na(Alcsum)] <- 0

names(Alcsum) <- c("ID", "Alc6_9", "Alc9_12", "Alc12_14", "Alc14_17",
                "Alc17_20", "Alc20_22", "Alc22_6")

# Extract number of days of diary completed

NDAYS14 <- blood14 %>%
  select(ID, Ndays)
```

```r
NDAYS56 <- blood56 %>%
  select(ID, Ndays)

NDAYS78 <- blood78 %>%
  select(ID, NDays) %>%
  rename(Ndays = NDays)

NDAYS <- rbind(NDAYS14, NDAYS56, NDAYS78)
NDAYS$ID <- as.numeric(NDAYS$ID)

IntakeSlots <- Energysum %>%
  left_join(Carbsum, by = "ID") %>%
  left_join(Sugarsum, by = "ID") %>%
  left_join(Starchsum, by = "ID") %>%
  left_join(Fibresum, by = "ID") %>%
  left_join(Fatsum, by = "ID") %>%
  left_join(Protsum, by = "ID") %>%
  left_join(NMESsum, by = "ID") %>%
  left_join(Alcsum, by = "ID") %>%
  left_join(NDAYS, by = "ID")

IntakeSlots$ID <- as.numeric(IntakeSlots$ID)

IntakeSlots$Energy6_9 <- IntakeSlots$Energy6_9/(IntakeSlots$Ndays)

for (i in 3:57){
  IntakeSlots[, i] <- IntakeSlots[, i]/(IntakeSlots$Ndays)
}

# select the variables needed, recoding
# NAs, renaming to the same ----------------------------------------

BMI78 <- blood78 %>%
  select(ID, Sex, age, bmival, wstval, Diabetes, bpmedc2, bpmedd2,
         hyper140_2, hibp140_2,  Glucose, A1C, cigsta3, dnoft3,
         dnnow, wti_Y78, wtn_Y78, wtb_Y78, ethgrp5, ethgrp2,
         cluster1, cluster2, cluster3, nssec8, paidemployment,
         qual7, eqvinc, MarSt2, cluster4, cluster5, area, gor,
         LDL, HDL, Chol, Trig) %>%
  rename(wti = wti_Y78, wtn = wtn_Y78, wtb = wtb_Y78,
         drink = dnoft3) %>%
  mutate(Years = "7-8", MVPAtime = NA, MarStat = NA) %>%
  replace_with_na(replace = list(bmival = -1, qual7 = -8,
                                 wstval = -1, eqvinc = -1,
                                 bpmedd2 = -1, MarSt2 = -1,
                                 bpmedc2 = -1, hyper140_2 = -7,
                                 hibp140_2 = -7, Glucose = -1,
                                 A1C  = -1,  LDL = -1, HDL = -1,
                                 Chol = -1, Trig = -1, dnnow = -1,
                                 drink = -1, ethgrp5 = -4,
                                 ethgrp2 = -4, cigsta3 = -1,
                                 nssec8 = -9,  paidemployment = -1)) %>%
  replace_with_na(replace = list(hyper140_2 = -1, hibp140_2 = -1,
                                 nssec8 = 99, qual7 = -1, drink = -8,
                                 ethgrp2 = -9, ethgrp5 = -9)) %>%
  replace_with_na(replace = list(drink = -9, ethgrp2 = -8, ethgrp5 = -8,
                                 cigsta3 = -8)) %>%
  replace_with_na(replace = list(ethgrp2 = -1, ethgrp5 = -1))
```

```r
BMI56 <- blood56 %>%
  select(ID, Sex, age, area, bmival, wstval, Diabetes, bpmedc2, bpmedd2,
         hyper140_2, hibp140_2, MVPAtime, Glucose, A1C, cigsta3, dnoft3,
         dnnow, wti_Y56, wtn_Y56, wtb_Y56, ethgrp5, ethgrp2, qual7,
         cluster1, cluster2, cluster3, nssec8, paidemployment, eqvinc,
         MarSt2, cluster4, cluster5, area, gor, LDL, HDL, Chol, Trig) %>%
  mutate(Years = "5-6", MarStat = NA) %>%
  rename(wti = wti_Y56, wtn = wtn_Y56, wtb = wtb_Y56, drink = dnoft3) %>%
  replace_with_na(replace = list(bmival = -1, paidemployment = -9,
                                 wstval = -1, qual7 = -8,
                                 bpmedd2 = -1, eqvinc = -1,
                                 bpmedc2 = -1, MarSt2 = -1,
                                 hyper140_2 = -7, hibp140_2 = -7,
                                 Glucose = -1, A1C  = -1,
                                 dnnow = -1, drink = -1, ethgrp5 = -4,
                                 ethgrp2 = -4, LDL = -1, HDL = -1,
                                 Chol = -1, Trig = -1, cigsta3 = -1,
                                 MVPAtime = -1, nssec8 = -9)) %>%
  replace_with_na(replace = list(hyper140_2 = -1, hibp140_2 = -1,
                                 paidemployment = -8,
                                 drink = -8, nssec8 = 99)) %>%
  replace_with_na(replace = list(drink = -9, paidemployment = -1,
                                 qual7 = -1, cigsta3 = -8))

BMI14 <- blood14 %>%
  select(ID, Sex, age, bmival, wstval, Diabetes, bpmedc, bpmedd, hyper140,
         hibp140, MVPAtime, Glucose, A1C, cigsta3, dnoft3, dnnow,
         wti_CY1234, wtn_CY1234, wtb_CY1234, ethgr5, ethgr2, cluster,
         area, gor, nssec8, paidemployment, qual7, eqvinc, MarSt2,
         MarStat, LDL, HDL, Chol, Trig) %>%
  rename(hyper140_2 = hyper140, hibp140_2 = hibp140, bpmedd2 = bpmedd,
         bpmedc2 = bpmedc, cluster1 = cluster, ethgrp5 = ethgr5,
         ethgrp2 = ethgr2, wti = wti_CY1234, wtn = wtn_CY1234,
         wtb =  wtb_CY1234, drink = dnoft3) %>%
  mutate(cluster2 = NA, cluster3 = NA, cluster4 = NA, cluster5 = NA,
         Years = "1-4") %>%
  replace_with_na(replace = list(bmival = -1, paidemployment = -9,
                                 wstval = -1, qual7 = -8,
                                 bpmedd2 = -1, eqvinc = -1,
                                 bpmedc2 = -1, MarSt2 = -4,
                                 hyper140_2 = -7, MarStat = -4,
                                 hibp140_2 = -7,  Glucose = -1,
                                 A1C  = -1,  dnnow = -1, LDL = -1,
                                 HDL = -1, Chol = -1, Trig = -1,
                                 drink = -1, ethgrp5 = -4, ethgrp2 = -4,
                                 cigsta3 = -1, MVPAtime = -4,
                                 nssec8 = -8)) %>%
  replace_with_na(replace = list(hyper140_2 = -1, hibp140_2 = -1,
                                 MVPAtime = -1, paidemployment = -8,
                                 drink = -8, nssec8 = -1, qual7 = -1,
                                 MarSt2 = -1)) %>%
  replace_with_na(replace = list(drink = -9, paidemployment = -4,
                                 cigsta3 = -8, nssec8 = 99)) %>%
  replace_with_na(replace = list(paidemployment = -1))

BMI <- bind_rows(BMI14, BMI56, BMI78)

Energy14 <- food14 %>%
```

```r
  select(ID, Country, SurveyYear, EnergykJ, Carbohydrateg,
         CHOpctotE, Proteing, ProteinpctotE, Alcoholg, AlcoholpctotE,
         Fatg, FatpctotE)
Energy14$Country[Energy14$Country == "Northern Ireland"] <- "NI"

Energy56 <- food56 %>%
  select(ID, Country, Surveyyear, EnergykJ, Carbohydrateg,
         CHOpctotE, Proteing, ProteinpctotE, Alcoholg, AlcoholpctotE,
         Fatg, FatpctotE) %>%
  rename(SurveyYear = Surveyyear)

Energy78 <- food78 %>%
  select(ID, Country, SurveyYear, EnergykJ, Carbohydrateg,
         CHOpctotE, Proteing, ProteinpctotE, Alcoholg, AlcoholpctotE,
         Fatg, FatpctotE)

Energy <- bind_rows(Energy14, Energy56, Energy78)

dta_NDNS <- read_csv("dta_NDNS_Tslots.csv") # extract the day of week data

CW3idday$DayNo <- ave(CW3idday$ID_DAY, CW3idday$ID,
                      FUN = seq_along) # adding the day no

dta_DayofWeek <- dta_NDNS %>%
  select(id, id_dy, DayofWeek)

dta_DayofWeek$DayNo <- unlist(strsplit(dta_DayofWeek$id_dy,
                                       "D"))[c(FALSE, TRUE)] # creating a day No
names(dta_DayofWeek)[1] <- "ID"
dta_DayofWeek$DayNo <- as.numeric(dta_DayofWeek$DayNo)

CW3idday <-   CW3idday %>%
  left_join(dta_DayofWeek, by= c("ID", "DayNo"))

## Manually check the day of week
CW3idday$DayofWeek[CW3idday$ID_DAY == 40714261000] <- "Saturday"
CW3idday$DayofWeek[CW3idday$ID_DAY == 112050710000] <- "Tuesday"
CW3idday$DayofWeek[CW3idday$ID_DAY == 310122510000] <- "Wednesday"
CW3idday$DayofWeek[CW3idday$ID_DAY == 5050616100] <- "Monday"
CW3idday$DayofWeek[CW3idday$ID_DAY == 50506161000] <- "Tuesday"
CW3idday$DayofWeek[CW3idday$ID_DAY == 505061610000] <- "Wednesday"
CW3idday$DayofWeek[CW3idday$ID_DAY == 7090824100] <- "Monday"
CW3idday$DayofWeek[CW3idday$ID_DAY == 70908241000] <- "Tuesday"
CW3idday$DayofWeek[CW3idday$ID_DAY == 709082410000] <- "Wednesday"
CW3idday$DayofWeek[CW3idday$ID_DAY == 807021910000] <- "Monday"
CW3idday$DayofWeek[CW3idday$ID_DAY == 812101310000] <- "Sunday"


CW3CB3_reg <- CW3CB3[!duplicated(CW3CB3$ID), ]

# extract only the CB variable (Between individual classes == 1 or 2)
CW3CB3_reg <- CW3CB3_reg %>%
  select(ID, AGE, SEX, CB)

tab1(CW3CB3_reg$CB, graph = FALSE)


## dataset for 3by3 multilevel latent classes
```

```
CW3CB3_regss <- CW3CB3_reg %>%
  left_join(BMI, by = "ID") %>%
  left_join(Energy, by = "ID")

# Rescale the weighting values
# individual weighting

a <- sum(CW3CB3_regss[CW3CB3_regss$Years == "1-4",]$wti)
b <- sum(CW3CB3_regss[CW3CB3_regss$Years == "5-6",]$wti)
c <- sum(CW3CB3_regss[CW3CB3_regss$Years == "7-8",]$wti)

CW3CB3_regss$wti1to8 <- CW3CB3_regss$wti

CW3CB3_regss[CW3CB3_regss$Years == "1-4",]$wti1to8 <-
  CW3CB3_regss[CW3CB3_regss$Years == "1-4",]$wti*(a+b+c)*(1/2)/a
CW3CB3_regss[CW3CB3_regss$Years == "5-6",]$wti1to8 <-
  CW3CB3_regss[CW3CB3_regss$Years == "5-6",]$wti*(a+b+c)*(1/4)/b
CW3CB3_regss[CW3CB3_regss$Years == "7-8",]$wti1to8 <-
  CW3CB3_regss[CW3CB3_regss$Years == "7-8",]$wti*(a+b+c)*(1/4)/c
mean(CW3CB3_regss$wti1to8)

CW3CB3_regss$wti1to8 <- CW3CB3_regss$wti1to8/1.209816814
summ(CW3CB3_regss$wti1to8, graph = FALSE)
#Check if the weighting sum up to the sample size we have
sum(CW3CB3_regss$wti1to8, graph = FALSE)

# Nurse weights

a <- sum(CW3CB3_regss[CW3CB3_regss$Years == "1-4",]$wtn)
b <- sum(CW3CB3_regss[CW3CB3_regss$Years == "5-6",]$wtn)
c <- sum(CW3CB3_regss[CW3CB3_regss$Years == "7-8",]$wtn)

CW3CB3_regss$wtn1to8 <- CW3CB3_regss$wtn

CW3CB3_regss[CW3CB3_regss$Years == "1-4",]$wtn1to8 <-
  CW3CB3_regss[CW3CB3_regss$Years == "1-4",]$wtn*(a+b+c)*(1/2)/a
CW3CB3_regss[CW3CB3_regss$Years == "5-6",]$wtn1to8 <-
  CW3CB3_regss[CW3CB3_regss$Years == "5-6",]$wtn*(a+b+c)*(1/4)/b
CW3CB3_regss[CW3CB3_regss$Years == "7-8",]$wtn1to8 <-
  CW3CB3_regss[CW3CB3_regss$Years == "7-8",]$wtn*(a+b+c)*(1/4)/c
mean(CW3CB3_regss$wtn1to8)

CW3CB3_regss$wtn1to8 <- CW3CB3_regss$wtn1to8/0.907003577
summ(CW3CB3_regss$wtn1to8, graph = FALSE)
#Check if the weighting sum up to the sample size we have
sum(CW3CB3_regss$wtn1to8, graph = FALSE)

# Blood weights
a <- sum(CW3CB3_regss[CW3CB3_regss$Years == "1-4",]$wtb)
b <- sum(CW3CB3_regss[CW3CB3_regss$Years == "5-6",]$wtb)
c <- sum(CW3CB3_regss[CW3CB3_regss$Years == "7-8",]$wtb)

CW3CB3_regss$wtb1to8 <- CW3CB3_regss$wtb

CW3CB3_regss[CW3CB3_regss$Years == "1-4",]$wtb1to8 <-
  CW3CB3_regss[CW3CB3_regss$Years == "1-4",]$wtb*(a+b+c)*(1/2)/a
CW3CB3_regss[CW3CB3_regss$Years == "5-6",]$wtb1to8 <-
  CW3CB3_regss[CW3CB3_regss$Years == "5-6",]$wtb*(a+b+c)*(1/4)/b
```

```
CW3CB3_regss[CW3CB3_regss$Years == "7-8",]$wtb1to8 <-
  CW3CB3_regss[CW3CB3_regss$Years == "7-8",]$wtb*(a+b+c)*(1/4)/c
mean(CW3CB3_regss$wtb1to8)

CW3CB3_regss$wtb1to8 <- CW3CB3_regss$wtb1to8/0.4817444505
summ(CW3CB3_regss$wtb1to8, graph = FALSE)
#Check if the weighting sum up to the sample size we have
sum(CW3CB3_regss$wtb1to8, graph = FALSE)

weightings <- CW3CB3_regss %>% select(ID, wti1to8, wtn1to8, wtb1to8)

# Combine the data from nutrient intake at each time slot
CW3CB3_7regss <- CW3CB3_7regss %>%
  left_join(IntakeSlots, by = "ID")

## dta ready to be analysed in STATA
 # change the path accordingly
write_dta(CW3CB3_7regss, "../CW3CB3_7regss.dta")


*************************************************
// Analysing NDNS survey data in stata
// for CW3CB3 survey data analysis
// date created: 2018-08-01
// manipulation of the data was done in R
// import data from CW3CB3_7sregss.dta
// change the path accordingly
*************************************************


use "../CW3CB3_7regss.dta", clear

label define smoking 1 "current" 2 "ex-smoker" 3 "Never"
label values cigsta3 smoking
label define gender 1 "Men" 2 "Women"
label values Sex gender
label define paid 1 "No" 2 "Yes"
label values paidemployment paid
label define ethnicity5 1 "White" 2 "Mixed" 3 "Black" 4 "Asian" 5 "Other"
label values ethgrp5 ethnicity5
label define ethnicity2 1 "White" 2 "non-White"
label values ethgrp2 ethnicity2

gen Married = 1 if MarStat == 2 | MarSt2 == 2
replace Married = 1 if MarSt2 == 3
replace Married = 0 if Married !=1
tab Married
tab MarSt2
tab MarStat

label define Partner 0 "No" 1 "Yes"
label values Married Partner

gen Education = qual7 == 1
label define Ed 0 "lower than Degree" 1 "Degree or higher"
label values Education Ed

replace Education = . if qual7 >100
tab Educ
```

```
egen BMIcat = cut(bmival), at(10, 25, 30, 40, 60)
tab BMIcat

************************************************************
// variables need to be log transformed                  //
************************************************************

gen logalc = ln(Alcoholg+1)
summ logalc, detail
gen logMVP = ln(MVPAtime+1)
summ logMVP, detail
gen logGlu = ln(Glucose)
summ logGlu, detail
gen logA1C = ln(A1C)
summ logA1C, detail
gen logChol = ln(Chol)
summ logChol, detail
gen logLDL = ln(LDL)
gen logHDL = ln(HDL)
gen logTG = ln(Trig)

************************************************************
//  weighting use wti to see the individual results     //
//                                                      //
************************************************************


// weighting with individual weights, area is primary sampling unit,
// gor is the cluster variable

svyset area [pweight = wti1to8], strata(gor)

svydescribe wti // describe the weighted data set

svy: tabulate Sex CB, row se ci format(%7.3f)
svy: tabulate Sex CB, col se ci format(%7.3f)
svy: tabulate Country CB, col se ci format(%7.3f)
svy: tabulate Country CB, row se ci format(%7.3f)
svy: tabulate SurveyYear CB, col se ci format(%7.3f)
svy: tabulate SurveyYear CB, row se ci format(%7.3f)
svy: tabulate paid CB, col se ci format(%7.3f)

svy: tabulate MarSt2 CB
svy: tabulate MarStat CB
svy: tabulate Married CB, row se ci format(%7.3f)
svy: tabulate Married CB, col se ci format(%7.3f)
svy: mean eqvinc, over(CB)
test [eqvinc]1 = [eqvinc]2 = [eqvinc]3, mtest(b)
// bonferroni-adjusted p-values for multiple groups using the mtest(b) option

svy: tabulate ethgrp2 CB, row se ci format(%7.3f)
svy: tabulate ethgrp2 CB, col se ci format(%7.3f)
svy: tabulate Education CB, row se ci format(%7.3f)
svy: tabulate Education CB, col se ci format(%7.3f)
```

```
****************************************************
//  nutritional distribution                      //
//                                                //
****************************************************
svy: mean EnergykJ, over(CB)
test [EnergykJ]1 = [EnergykJ]2 = [EnergykJ]3, mtest(b)

svy: mean Energy6, over(CB)
test [Energy6_9]1 = [Energy6_9]2 = [Energy6_9]3, mtest(b)

svy: mean Energy9, over(CB)
test [Energy9_12]1 = [Energy9_12]2 = [Energy9_12]3, mtest(b)

svy: mean Energy12, over(CB)
test [Energy12_14]1 = [Energy12_14]2 = [Energy12_14]3, mtest(b)

svy: mean Energy14, over(CB)
test [Energy14_17]1 = [Energy14_17]2 = [Energy14_17]3, mtest(b)

svy: mean Energy17, over(CB)
test [Energy17_20]1 = [Energy17_20]2 = [Energy17_20]3, mtest(b)

svy: mean Energy20, over(CB)
test [Energy20_22]1 = [Energy20_22]2 = [Energy20_22]3, mtest(b)

svy: mean Energy22, over(CB)
test [Energy22_6]1 = [Energy22_6]2 = [Energy22_6]3, mtest(b)


svy: mean Carbohydrateg, over(CB)
test [Carbohydrateg]1 = [Carbohydrateg]2 = [Carbohydrateg]3, mtest(b)

svy: mean Carb6, over(CB)
test [Carb6_9]1 = [Carb6_9]2 = [Carb6_9]3, mtest(b)

svy: mean Carb9, over(CB)
test [Carb9_12]1 = [Carb9_12]2 = [Carb9_12]3, mtest(b)

svy: mean Carb12, over(CB)
test [Carb12_14]1 = [Carb12_14]2 = [Carb12_14]3, mtest(b)

svy: mean Carb14, over(CB)
test [Carb14_17]1 = [Carb14_17]2 = [Carb14_17]3, mtest(b)

svy: mean Carb17, over(CB)
test [Carb17_20]1 = [Carb17_20]2 = [Carb17_20]3, mtest(b)


svy: mean Carb20, over(CB)
test [Carb20_22]1 = [Carb20_22]2 = [Carb20_22]3, mtest(b)

svy: mean Carb22, over(CB)
test [Carb22_6]1 = [Carb22_6]2 = [Carb22_6]3, mtest(b)

svy: mean Sugar6, over(CB)
test [Sugar6_9]1 = [Sugar6_9]2 = [Sugar6_9]3, mtest(b)
```

```
svy: mean Sugar9, over(CB)
test [Sugar9_12]1 = [Sugar9_12]2 = [Sugar9_12]3, mtest(b)

svy: mean Sugar12, over(CB)
test [Sugar9_12]1 = [Sugar9_12]2 = [Sugar9_12]3, mtest(b)

svy: mean Sugar14, over(CB)
test [Sugar14_17]1 = [Sugar14_17]2 = [Sugar14_17]3, mtest(b)

svy: mean Sugar17, over(CB)
test [Sugar17_20]1 = [Sugar17_20]2 = [Sugar17_20]3, mtest(b)

svy: mean Sugar20, over(CB)
test [Sugar20_22]1 = [Sugar20_22]2 = [Sugar20_22]3, mtest(b)

svy: mean Sugar22, over(CB)
test [Sugar22_6]1 = [Sugar22_6]2 = [Sugar22_6]3, mtest(b)

svy: mean Starch6, over(CB)
test [Starch6_9]1 = [Starch6_9]2 = [Starch6_9]3, mtest(b)

svy: mean Starch9, over(CB)
test [Sugar12_14]1 = [Sugar12_14]2 = [Sugar12_14]3, mtest(b)

svy: mean Starch12, over(CB)
test [Starch12_14]1 = [Starch12_14]2 = [Starch12_14]3, mtest(b)

svy: mean Starch14, over(CB)
test [Starch14_17]1 = [Starch14_17]2 = [Starch14_17]3, mtest(b)

svy: mean Starch17, over(CB)
test [Starch17_20]1 = [Starch17_20]2 = [Starch17_20]3, mtest(b)


svy: mean Starch20, over(CB)
test [Starch20_22]1 = [Starch20_22]2 = [Starch20_22]3, mtest(b)

svy: mean Starch22, over(CB)
test [Starch20_22]1 = [Starch20_22]2 = [Starch20_22]3, mtest(b)

svy: mean Fibre6, over(CB)
test [Starch22_6]1 = [Starch22_6]2 = [Starch22_6]3, mtest(b)

gen Fibreg = Fibre6 + Fibre9 + Fibre12 + Fibre14 + Fibre17 + Fibre20 + Fibre22
svy: mean Fibreg, over(CB)
test [Fibreg]1 = [Fibreg]2 = [Fibreg]3, mtest(b)


svy: mean Fibre9, over(CB)
test [Fibre9_12]1 = [Fibre9_12]2 = [Fibre9_12]3, mtest(b)


svy: mean Fibre12, over(CB)
test [Fibre12_14]1 = [Fibre12_14]2 = [Fibre12_14]3, mtest(b)

svy: mean Fibre14, over(CB)
test [Fibre14_17]1 = [Fibre14_17]2 = [Fibre14_17]3, mtest(b)
```

```
svy: mean Fibre17, over(CB)
test [Fibre17_20]1 = [Fibre17_20]2 = [Fibre17_20]3, mtest(b)

svy: mean Fibre20, over(CB)
test [Fibre20_22]1 = [Fibre20_22]2 = [Fibre20_22]3, mtest(b)

svy: mean Fibre22, over(CB)
test [Fibre22_6]1 = [Fibre22_6]2 = [Fibre22_6]3, mtest(b)

svy: mean NMES6, over(CB)
test [NMES6_9]1 = [NMES6_9]2 = [NMES6_9]3, mtest(b)

svy: mean NMES9, over(CB)
test [NMES9_12]1 = [NMES9_12]2 = [NMES9_12]3, mtest(b)


svy: mean NMES12, over(CB)
test [NMES12_14]1 = [NMES12_14]2 = [NMES12_14]3, mtest(b)

svy: mean NMES14, over(CB)
test [NMES14_17]1 = [NMES14_17]2 = [NMES14_17]3, mtest(b)

svy: mean NMES17, over(CB)
test [NMES17_20]1 = [NMES17_20]2 = [NMES17_20]3, mtest(b)


svy: mean NMES20, over(CB)
test [NMES20_22]1 = [NMES20_22]2 = [NMES20_22]3, mtest(b)

svy: mean NMES22, over(CB)


svy: mean CHO, over(CB)
test [CHOpctotE]1 = [CHOpctotE]2 = [CHOpctotE]3, mtest(b)


svy: mean Proteing, over(CB)
test [Proteing]1 = [Proteing]2 = [Proteing]3, mtest(b)

svy: mean Prot6, over(CB)
test [Prot6_9]1 = [Prot6_9]2 = [Prot6_9]3, mtest(b)

svy: mean Prot9, over(CB)
test [Prot9_12]1 = [Prot9_12]2 = [Prot9_12]3, mtest(b)

svy: mean Prot12, over(CB)
test [Prot12_14]1 = [Prot12_14]2 = [Prot12_14]3, mtest(b)

svy: mean Prot14, over(CB)
test [Prot14_17]1 = [Prot14_17]2 = [Prot14_17]3, mtest(b)

svy: mean Prot17, over(CB)
test [Prot17_20]1 = [Prot17_20]2 = [Prot17_20]3, mtest(b)

svy: mean Prot20, over(CB)
test [Prot20_22]1 = [Prot20_22]2 = [Prot20_22]3, mtest(b)

svy: mean Prot22, over(CB)
test [Prot22_6]1 = [Prot22_6]2 = [Prot22_6]3, mtest(b)
```

```
svy: mean Proteinp, over(CB)
test [ProteinpctotE]1 = [ProteinpctotE]2 = [ProteinpctotE]3, mtest(b)

svy: mean Fatg, over(CB)
test [Fatg]1 = [Fatg]2 = [Fatg]3, mtest(b)


svy: mean Fat6, over(CB)
test [Fat6_9]1 = [Fat6_9]2 = [Fat6_9]3, mtest(b)

svy: mean Fat9, over(CB)
test [Fat9_12]1 = [Fat9_12]2 = [Fat9_12]3, mtest(b)


svy: mean Fat12, over(CB)
test [Fat12_14]1 = [Fat12_14]2 = [Fat12_14]3, mtest(b)

svy: mean Fat14, over(CB)
test [Fat14_17]1 = [Fat14_17]2 = [Fat14_17]3, mtest(b)

svy: mean Fat17, over(CB)
test [Fat17_20]1 = [Fat17_20]2 = [Fat17_20]3, mtest(b)

svy: mean Fat20, over(CB)
test [Fat20_22]1 = [Fat20_22]2 = [Fat20_22]3, mtest(b)

svy: mean Fat22, over(CB)
test [Fat22_6]1 = [Fat22_6]2 = [Fat22_6]3, mtest(b)

svy: mean Fatp, over(CB)
test [FatpctotE]1 = [FatpctotE]2 = [FatpctotE]3, mtest(b)

svy: mean Alcoholg, over(CB)
test [Alcoholg]1 = [Alcoholg]2 = [Alcoholg]3, mtest(b)

svy: mean Alc6, over(CB)
test [Alc6_9]1 = [Alc6_9]2 = [Alc6_9]3, mtest(b)

svy: mean Alc9, over(CB)
test [Alc9_12]1 = [Alc9_12]2 = [Alc9_12]3, mtest(b)

svy: mean Alc12, over(CB)
test [Alc12_14]1 = [Alc12_14]2 = [Alc12_14]3, mtest(b)

svy: mean Alc14, over(CB)
test [Alc14_17]1 = [Alc14_17]2 = [Alc14_17]3, mtest(b)

svy: mean Alc17, over(CB)
test [Alc14_17]1 = [Alc14_17]2 = [Alc14_17]3, mtest(b)

svy: mean Alc20, over(CB)
test [Alc14_17]1 = [Alc14_17]2 = [Alc14_17]3, mtest(b)

svy: mean Alcoholp, over(CB)
test [AlcoholpctotE]1 = [AlcoholpctotE]2 = [AlcoholpctotE]3, mtest(b)

svy: tabulate cigsta3 CB, col se ci format(%7.3f)
```

```
svy: tabulate dnnow CB, col se ci format(%7.3f)
svy: tabulate hibp CB, col se ci format(%7.3f)

sum MVP [weight=wti1to8] if CB ==1 , det
sum MVP [weight=wti1to8] if CB ==2 , det
sum MVP [weight=wti1to8] if CB ==3 , det
svy: mean MVP, over(CB)

svy: mean logMVP, over(CB) eform
test [logMVP]1 = [logMVP]2 = [logMVP]3, mtest(b)

disp exp(.731059) - 1
dis exp(.6768489) -1
dis exp(.7852691) -1

disp exp(.6239265) - 1
dis exp(.571165) -1
dis exp( .6766879) -1

disp exp(.7273621) - 1
dis exp(.684545) -1
dis exp(.7701791) -1

svy: mean logalc, over(CB)


disp exp( 2.035795) - 1
dis exp(1.933326) -1
dis exp(2.138264) -1


*******************************************************
//  re-weighting use wtn to see the BMI,wc measurements //
//                                                    //
*******************************************************

svyset area [pweight = wtn1to8], strata(gor)
svy: mean wst, over(CB)
test [wstval]1 = [wstval]2 = [wstval]3, mtest(b)

gen Men = Sex == 1
svy, subpop(Men): mean wst, over(CB)
test [wstval]1 = [wstval]2 = [wstval]3, mtest(b)

gen Women = Sex == 2
svy, subpop(Women): mean wst, over(CB)
test [wstval]1 = [wstval]2 = [wstval]3, mtest(b)

svy: mean bmi, over(CB)

test [bmival]1 = [bmival]2 = [bmival]3, mtest(b)

*******************************************************
//  re-weighting use wtb to see the blood test results //
//                                                    //
*******************************************************

svyset area [pweight = wtb1to8], strata(gor)
```

```
svy: mean HDL, over(CB)
test [HDL]1 = [HDL]2 = [HDL]3, mtest(b)

svy: mean Chol, over(CB)
test [Chol]1 = [Chol]2 = [Chol]3, mtest(b)

svy: mean LDL, over(CB)
test [LDL]1 = [LDL]2 = [LDL]3, mtest(b)

svy: mean Trig, over(CB)
test [Trig]1 = [Trig]2 = [Trig]3, mtest(b)

gen DM = A1C <= 6.5 if !missing(A1C)

svy, subpop(DM): mean Glucose, over(CB)
test [Glucose]1 = [Glucose]2
test [Glucose]1 = [Glucose]2 = [Glucose]3, mtest(b)

svy, subpop(DM): mean A1C, over(CB)
test [A1C]1 = [A1C]2
test [A1C]1 = [A1C]2 = [A1C]3, mtest(b)

svy: tabulate DM CB, col se ci format(%7.3f)

svy, subpop(DM): mean Glucose, over(CB)
test [Glucose]1 = [Glucose]2
test [Glucose]1 = [Glucose]2 = [Glucose]3, mtest(b)
svy, subpop(DM): mean A1C, over(C)
test [A1C]1 = [A1C]2
test [A1C]1 = [A1C]2 = [A1C]3, mtest(b)

svy: tabulate DM C, col se ci format(%7.3f)

svy, subpop(DM): mean logGlu, over(CB)
test [logGlu]1 = [logGlu]2 = [logGlu]3, mtest(b)

dis exp(1.642848)
dis exp(1.632226)
dis exp(1.653471)

dis exp(1.620347)
dis exp(1.606447)
dis exp(1.634246)

dis exp(1.629356)
dis exp(1.620271)
dis exp(1.63844)

svy, subpop(DM): mean logA1C, over(CB)
test [logA1C]1 = [logA1C]2 = [logA1C]3, mtest(b)

dis exp(1.699581)
dis exp(1.69296)
dis exp(1.706203)

dis exp(1.691608)
dis exp(1.683897)
dis exp(1.699318)
```

```
dis exp(1.705623)
dis exp(1.700665)
dis exp(1.710581)

svy: mean logChol, over(CB)
dis exp(1.598818)
dis exp(1.577698)
dis exp(1.619939)

dis exp(1.55251)
dis exp(1.530408)
dis exp(1.574613)

dis exp(1.599389)
dis exp(1.583391)
dis exp(1.615388)

test [logChol]1 = [logChol]2 = [logChol]3, mtest(b)

svy: mean logHDL, over(CB)
dis exp(.3293169)
dis exp(.3026793)
dis exp(.3559545)

dis exp(.2749379)
dis exp(.2476816)
dis exp(.3021941)

dis exp(.3269002)
dis exp(.3062623)
dis exp(.3475381)

test [logHDL]1 = [logHDL]2 = [logHDL]3, mtest(b)

svy: mean logLDL, over(CB)
dis exp(1.058635)
dis exp(1.028391)
dis exp(1.08888)

dis exp(1.018181)
dis exp(.984431)
dis exp(1.051931)

dis exp(1.075229)
dis exp(1.051369)
dis exp(1.09909)

test [logLDL]1 = [logLDL]2 = [logLDL]3, mtest(b)

svy: mean logTG, over(CB)
dis exp(.1273876)
dis exp(.0777152)
dis exp(.17706)

dis exp(.1012169)
dis exp(.0460972)
dis exp(.1563366)

dis exp(.0983298)
```

```
dis exp(.0607423)
dis exp(.1359172)


test [logTG]1 = [logTG]2 = [logTG]3, mtest(b)

****************************************************
// Analysing NDNS survey data in stata
// for CW3CB3 survey data analysis on hypertension
// date created: 2018-08-06
// manipulation of the data was done in R
// import data from CW3CB3_7sregss.dta
// change the path accordingly
****************************************************

use "../CW3CB3_7regss.dta", clear

label define smoking 1 "current" 2 "ex-smoker" 3 "Never"
label values cigsta3 smoking
label define gender 1 "Men" 2 "Women"
label values Sex gender
label define paid 1 "No" 2 "Yes"
label values paidemployment paid
label define ethnicity5 1 "White" 2 "Mixed" 3 "Black" 4 "Asian" 5 "Other"
label values ethgrp5 ethnicity5
label define ethnicity2 1 "White" 2 "non-White"
label values ethgrp2 ethnicity2

gen Married = 1 if MarStat == 2 | MarSt2 == 2
replace Married = 1 if MarSt2 == 3
replace Married = 0 if Married !=1
tab Married
tab MarSt2
tab MarStat

label define Partner 0 "No" 1 "Yes"
label values Married Partner

gen Education = qual7 == 1
label define Ed 0 "lower than Degree" 1 "Degree or higher"
label values Education Ed

replace Education = . if qual7 >100
tab Educ

egen BMIcat = cut(bmival), at(10, 25, 30, 40, 60)
tab BMIcat

***********************************************************
// variables need to be log transformed              //
***********************************************************

gen logalc = ln(Alcoholg+1)
summ logalc, detail
gen logMVP = ln(MVPAtime+1)
summ logMVP, detail
gen logGlu = ln(Glucose)
summ logGlu, detail
gen logA1C = ln(A1C)
```

```
summ logA1C, detail
gen logChol = ln(Chol)
summ logChol, detail
gen logLDL = ln(LDL)
gen logHDL = ln(HDL)
gen logTG = ln(Trig)


************************************************************
//  weighting use wti to see the individual results    //
************************************************************


// weighting with individual weights, area is primary sampling unit,
// gor is the cluster variable
svyset area [pweight = wti1to8], strata(gor)

svydescribe wti // describe the weighted data set


************************************************************
//  re-weighting use wtn to see the BMI,wc measurements //
************************************************************
svyset area [pweight = wtn1to8], strata(gor)


gen Men = Sex == 1  //  n of men = 2537
gen Women = Sex == 2 // n of women = 3618

svy, subpop(Men): tab hibp, se ci format(%7.3f)
svy, subpop(Women): tab hibp, se ci format(%7.3f)

svy, subpop(Men): mean age, over(hibp)
test [age]1 = [age]0
svy, subpop(Women): mean age, over(hibp)
test [age]1 = [age]0

svy, subpop(Men): mean wst, over(hibp)
test [wstval]1 = [wstval]0
svy, subpop(Women): mean wst, over(hibp)
test [wstval]1 = [wstval]0

svy, subpop(Men): tabulate CB hibp, col se ci format(%7.3f)

svy, subpop(Women): tabulate CB hibp, col se ci format(%7.3f)

svy, subpop(Men): tabulate Country hibp, col se ci format(%7.3f)

svy, subpop(Women): tabulate Country hibp, col se ci format(%7.3f)

svy, subpop(Men): tabulate SurveyYear hibp, col se ci format(%7.3f)

svy, subpop(Women): tabulate SurveyYear hibp, col se ci format(%7.3f)

svy, subpop(Men): tabulate ethgrp2 hibp, col se ci format(%7.3f)

svy, subpop(Women): tabulate ethgrp2 hibp, col se ci format(%7.3f)

svy, subpop(Men): tabulate Edu hibp, col se ci format(%7.3f)
```

```
svy, subpop(Women): tabulate Edu hibp, col se ci format(%7.3f)

svy, subpop(Men): tabulate cigsta3 hibp, col se ci format(%7.3f)

svy, subpop(Women): tabulate cigsta3 hibp, col se ci format(%7.3f)

svy, subpop(Men): tabulate Married hibp, col  se ci format(%7.3f)

svy, subpop(Men): mean logMVP, over(hibp)
test [logMVP]1 = [logMVP]0

disp exp(.9234363 ) - 1
dis exp(.8457101) -1
dis exp(1.001163) -1

disp exp(.828635) - 1
dis exp(.730244) -1
dis exp( .9270261) -1

svy, subpop(Women): mean logMVP, over(hibp)
test [logMVP]1 = [logMVP]0

disp exp(.5916676) - 1
dis exp(.5473043) -1
dis exp(.6360309) -1

disp exp(.4231103) - 1
dis exp(.3536885) -1
dis exp(.4925322) -1

svy: mean logalc, over(CB)

disp exp( 2.035795) - 1
dis exp(1.933326) -1
dis exp(2.138264) -1

svy, subpop(Men): mean bmi, over(hibp)
test [bmival]1 = [bmival]0
svy, subpop(Women): mean bmi, over(hibp)
test [bmival]1 = [bmival]0

svy, subpop(Men): mean EnergykJ, over(hibp)
test [EnergykJkJ]1 = [EnergykJkJ]0
svy, subpop(Women): mean EnergykJ, over(hibp)
test [EnergykJkJ]1 = [EnergykJkJ]0

svy, subpop(Men): mean Carbo, over(hibp)
test [Carbohydrateg]1 = [Carbohydrateg]0

svy, subpop(Women): mean Carbohydrateg, over(hibp)
test [Carbohydrateg]1 = [Carbohydrateg]0

svy, subpop(Men): mean Proteing, over(hibp)
test [Proteing]1 = [Proteing]0

svy, subpop(Women): mean Carbohydrateg, over(hibp)
test [Carbohydrateg]1 = [Carbohydrateg]0
```

```
svy: tabulate Sex hibp, col se ci format(%7.3f)

svy: tabulate paid hibp, col se ci format(%7.3f)

gen DM = A1C > 6.5 if !missing(A1C)

svy, subpop(Men): tabulate Married hibp, col se ci format(%7.3f)
svy, subpop(Women): tabulate Married hibp, col se ci format(%7.3f)

svy, subpop(Men): mean eqvinc, over(hibp)
test [eqvinc]1 = [eqvinc]0

svy, subpop(Women): mean eqvinc, over(hibp)
test [eqvinc]1 = [eqvinc]0

********************************************************
**    Building the GLM model
**    date: 07/08/2018
********************************************************

svyset area [pweight = wtn1to8], strata(gor)

// crude association between CB and hypertension

svy, subpop(Men): logistic hibp i.CB

svy, subpop(Women): logistic hibp i.CB

// in non DM
svy, subpop(Men if DM != 1): logistic hibp i.CB
svy, subpop(Women if DM != 1): logistic hibp i.CB

// looking for confounders one by one
// Age: -> confounder
svy, subpop(Men): logistic hibp i.CB age
test age
svy, subpop(Women): logistic hibp i.CB age
test age

svy, subpop(Men): logistic hibp i.CB##c.age
svy, subpop(Women): logistic hibp i.CB##c.age
test 2.CB#c.age 3.CB#c.age // no interaction

// Partner -> confounder
svy, subpop(Men): logistic hibp i.CB i.Married
test 1.Married
svy, subpop(Women): logistic hibp i.CB i.Married
test 1.Married

svy, subpop(Men): logistic hibp i.CB##i.Married
svy, subpop(Women): logistic hibp i.CB##i.Married

test 2.CB#1.Married 3.CB#1.Married // -> no interaction

// Income -> not confounder for men but confounder for women
svy, subpop(Men): logistic hibp i.CB eqvinc
test eqvinc
svy, subpop(Women): logistic hibp i.CB eqvinc
test eqvinc
```

```
svy, subpop(Men): logistic hibp i.CB##c.eqvinc
svy, subpop(Women): logistic hibp i.CB##c.eqvinc

test 2.CB#c.eqvinc 3.CB#c.eqvinc // -> (probably) no interaction

// Education -> confounder
svy, subpop(Men): logistic hibp i.CB i.Edu
test 1.Edu
svy, subpop(Women): logistic hibp i.CB i.Edu
test 1.Edu

svy, subpop(Men): logistic hibp i.CB##i.Edu
svy, subpop(Women): logistic hibp i.CB##i.Edu

test 2.CB#1.Edu 3.CB#1.Edu // no interaction

// BMI -> confounder
svy, subpop(Men): logistic hibp i.CB bmi
test bmi
svy, subpop(Women): logistic hibp i.CB bmi
test bmi

svy, subpop(Men): logistic hibp i.CB##c.bmi
svy, subpop(Women): logistic hibp i.CB##c.bmi

test 2.CB#c.bmival 3.CB#c.bmival // no ineraction

// paid employment -> not confounder
svy, subpop(Men): logistic hibp i.CB i.paid
test 2.paid
svy, subpop(Women): logistic hibp i.CB i.paid
test 2.paid

svy, subpop(Men): logistic hibp i.CB##i.paid
svy, subpop(Women): logistic hibp i.CB##i.paid

test 2.CB#2.paid 3.CB#2.paid

// Smoking -> confounder
svy, subpop(Men): logistic hibp i.CB i.cigsta3
test 2.cigsta3 3.cigsta3
svy, subpop(Women): logistic hibp i.CB i.cigsta3
test 2.cigsta3 3.cigsta3

svy, subpop(Men): logistic hibp i.CB##i.cigsta3
svy, subpop(Women): logistic hibp i.CB##i.cigsta3

test 2.CB#2.cigsta3 2.CB#3.cigsta3 3.CB#2.cigsta3 3.CB#3.cigsta3 // no interaction

// Total energy intake -> confounder
svy, subpop(Men): logistic hibp i.CB EnergykJ
test EnergykJ
svy, subpop(Women): logistic hibp i.CB EnergykJ
test EnergykJ

svy, subpop(Men): logistic hibp i.CB##c.EnergykJ
svy, subpop(Women): logistic hibp i.CB##c.EnergykJ
test 2.CB#c.EnergykJkJ 3.CB#c.EnergykJkJ // no interaction
```

```
// ethnicity -> not confounder
svy, subpop(Men): logistic hibp i.CB i.ethgrp2
test 2.eth
svy, subpop(Women): logistic hibp i.CB i.ethgrp2
test 2.eth

svy, subpop(Men): logistic hibp i.CB##i.ethgrp2
svy, subpop(Women): logistic hibp i.CB##i.ethgrp2
test 2.CB#2.ethgrp2 // no interaction

// Alcohol -> not confounder for men but confounder for women
svy, subpop(Men): logistic hibp i.CB Alcoholg
test Alcoholg
svy, subpop(Women): logistic hibp i.CB Alcoholg
test Alcoholg

svy, subpop(Men): logistic hibp i.CB##c.Alcoholg
svy, subpop(Women): logistic hibp i.CB##c.Alcoholg
test 2.CB#c.Alcoholg 3.CB#c.Alcoholg // no interaction

// logMVP -> not confounder
svy, subpop(Men): logistic hibp i.CB logMVP
test logMVP
svy, subpop(Women): logistic hibp i.CB logMVP
test logMVP

svy, subpop(Men): logistic hibp i.CB##c.logMVP
svy, subpop(Women): logistic hibp i.CB##c.logMVP
test 2.CB#c.logMVP 3.CB#c.logMVP // no interaction

//  Model includes all possible confounders in Men
svy, subpop(Men): logistic hibp i.CB age i.Married i.Edu bmi i.cig EnergykJ
linktest
svy, subpop(Men): logistic hibp i.CB age i.Married i.Edu wst i.cig EnergykJ
linktest
svy, subpop(if Men & DM != 1): logistic hibp i.CB age i.Married ///
 i.Edu bmi i.cig EnergykJ
linktest
svy, subpop(if Men & DM != 1): logistic hibp i.CB age i.Married ///
 i.Edu wst i.cig EnergykJ
linktest

// use the model above to see whether any other factors are confounders
// conditional on the other variables

// income -> not confounder
svy, subpop(Men): logistic hibp i.CB age i.Married i.Edu bmi i.cig ///
 EnergykJ eqvinc
test eqvinc

// ethnicity -> not confounder
svy, subpop(Men): logistic hibp i.CB age i.Married i.Edu bmi i.cig ///
 EnergykJ i.ethgrp2
test 2.ethgrp2

// alcohol -> not confounder
svy, subpop(Men): logistic hibp i.CB age i.Married i.Edu bmi i.cig ///
 EnergykJ Alcoholg
```

```
test Alcoholg

// MVP(physical activity) -> not confounder
svy, subpop(Men): logistic hibp i.CB age i.Married i.Edu bmi i.cig ///
 EnergykJ logMVP
test logMVP

//  Model includes all possible confounders in Women
svy, subpop(Women): logistic hibp i.CB age i.Married eqvinc i.Edu ///
 bmi i.cig EnergykJ Alcoholg
linktest
svy, subpop(Women): logistic hibp i.CB age i.Married eqvinc i.Edu ///
 wst i.cig EnergykJ Alcoholg
linktest


// use the model above to see whether any other factors are confounders
// conditional on the other variables

// ethnicity -> not confounder
svy, subpop(Women): logistic hibp i.CB age i.Married eqvinc i.Edu ///
 bmi i.cig EnergykJ Alcoholg i.ethgrp2
test 2.ethgrp2

// MVP (physical activity) -> not confounder
svy, subpop(Women): logistic hibp i.CB age i.Married eqvinc i.Edu ///
 bmi i.cig EnergykJ Alcoholg logMVP
test logMVP


svy, subpop(if Women & DM != 1): logistic hibp i.CB age i.Married ///
 eqvinc i.Edu bmi i.cig EnergykJ Alcoholg
linktest

svy, subpop(if Women & DM != 1): logistic hibp i.CB age i.Married ///
 eqvinc i.Edu wst i.cig EnergykJ Alcoholg
linktest
```

Appendix E

Example of a food diary for one day

Day	EXAMPLE		Day: Thursday	Date: March 31st
Time	Where? With whom? TV on? Table?	What	Brand Name	Amount eaten
		How to describe what you had and how much you had can be found on pages 20–25		
		6am to 9am		
7.30am	Kitchen Family No TV At table	Orange juice, unsweetened, UHT Tea Milk, fresh semi skimmed Sugar white Weetabix Milk as above Sugar as above Toast wholemeal, large loaf Butter unsalted Strawberry Jam	Tesco Tesco Tesco Silverspoon Hovis Anchor Co-op	Large glass Mug A little 2 level teaspoons 2 Drowned 2 heaped teaspoons 2 thin slices thick spread on both 1 teaspoon on one slice
		9am to 12 noon		
11am	School playground With friends	Coca cola diet Potato crisps, Salt and Vinegar	Coca Cola Walkers	330ml can 25g packet from a multipack
12noon	School corridor Alone	Water from water cooler Mars Bar		small plastic cup 1 kingsize
		12 noon to 2pm		
12.45pm	School canteen With friends At table	Sandwich, from home White bread, large loaf Spread Ham unsmoked Cheddar cheese Branston Pickle Apple with skin from home Ribena Light, Ready to Drink, Blackcurrant, from canteen Kitkat from home	Kingsmill Flora Light Tescos	2 med slices thin spread on both slices 1 slice 2 medium slices 1 teaspoon 1 (left core) 220ml carton 2 fingers
1.50pm	School corridor Alone	Chewing gum	Orbit Sugar Free	1 piece

FIGURE E.1: NATIONAL DIET AND NUTRITION SURVEY – Food and Drink Diary Example, from 6 am to 2 pm.

Day EXAMPLE			Day: Thursday	Date: March 31st
Time	Where? With whom? TV on? Table?	what	Brand Name	Amount eaten
		2pm to 5pm		
3.45pm	Bus Alone	Wine gums	Maynards	140g packet
4.30pm	Home, sitting room, With family TV on Not at table	Tea (as above) Chocolate Hob Nobs	Mcvitites	mug 3
		5pm to 8pm		
6.30pm	Friend's kitchen With friends No TV At table	Chicken in tomato sauce made by friend's mum Tomato fresh Sweetcorn tinned Peach yoghurt low fat Lemon squash No Added Sugar	See recipe Mullerlight Sainsbury's	3 tablespoons 3 slices 1 dessertspoon 200g pot medium glass
		8pm to 10pm		
8pm	Home, sitting room Alone TV on, Not at table	Satsuma Cream Crackers (no spread)	Jacob's	1 4
9.30pm	Kitchen Alone No TV, At table	Thick cut, frozen chips fried in vegetable oil Brown sauce	McCains HP	small portion 1 dessertspoon
		10pm to 6am		
10.30pm	Bedroom Alone TV on Not at table	Hot chocolate drink made with water	Cadbury's	Mug (made with 4 tsp powder)
2am	Bedroom (in bed) Alone No TV	Water tap		½ small glass

FIGURE E.2: NATIONAL DIET AND NUTRITION SURVEY – Food
and Drink Diary Example, from 2 pm to 6 am.

Write in recipes or ingredients of made up dishes or take-away dishes				
NAME OF DISH: Chicken in tomato Sauce		*Serves: 4 people*		
Ingredients	*Amount*	*Ingredients*		*Amount*
Pieces of chicken	3 pieces	Olive oil		2 tbsp
Sauce made with:				
Tinned tomatoes	1 tin			
Green pepper	1 medium			
Onion	1 small			
Brief description of cooking method				
Chicken pieces fried in olive oil, then mixed in with tomato and vegetable sauce.				

FIGURE E.3: NATIONAL DIET AND NUTRITION SURVEY – Food
and Drink Diary Example, home made food recipes.