Statistical Modeling of Dietary Intake and Weight Gain During Pregnancy

By

Menglu Che

A thesis submitted in partial fulfillment of the

requirements for the degree of

Master of Science

in

Statistics

Department of Mathematical and Statistical Sciences

UNIVERSITY OF ALBERTA

Abstract

BACKGROUND: Healthy dietary intake and appropriate weight gain are two key components of an ideal pregnancy. The objective of this thesis was to investigate the weight gain pattern of a large cohort of pregnant women and its association with dietary intakes, which may provide valuable information for the clinical intervention of inappropriate gestational weight gain. Two instruments were used to capture the dietary intakes, the level of agreement and possibility of pooling the results need to be studied.

METHOD: For a validation sample of 58 child-bearing-age women, the total calories intakes captured by two instruments were compared by the Bland-Altman plot. The intakes of key nutrient captured by the interviewer-administered instrument version were predicted by the nutrient intakes from the web-based version with a regression model. Then we estimated the weight growth trajectories of each subject through functional principal component analysis techniques. The total weight gain predicted from the trajectory was then regressed on the prepregnancy body mass index, and dietary intakes and physical activities which were measured through pregnancy.

RESULTS: We found that the relative bias between the two instruments were small, yet the variances in individuals could be large. Energy-adjusted intakes of macronutrients showed reasonable correlations between the two instruments (0.56 for fat, 0.73 for protein, and 0.67 for carbohydrate). LASSO regularization based multiple regression greatly improved the cross-validated R^2 for folate from 0.0033 to 0.46. Our estimated weight growth trajectories showed good accuracy when compared to classic mixed-effect models with significant smaller root mean squared error. The predicted weight gain from trajectory had a strong correlation with prepregnancy body mass index, but

adding the dietary intake and physical activity information did not improve the R^2 of the model.

CONCLUSIONS: Direct pooling of the results from the two instruments may not be feasible. But when pooling is considered, energy-adjustment for macronutrients and the LASSO-based multiple regression for micronutrients are recommended. Functional principal component analysis has significant advantages of flexibility and robustness for the weight growth trajectory modeling. We found that the weight gain during pregnancy negatively correlated to prepregnancy BMI, but the dietary intake and physical activities measured in our study did not provide useful information in predicting the weight gain.

Acknowledgements

First and foremost, I would like to express my deep and sincere gratitude to my supervisors, Dr. Linglong Kong and Dr. Yan Yuan. I thank them for giving me the opportunity to conduct my research, where I learned so much about statistics in health practice. I admire their knowledge and greatly appreciate their guidance, support, and patience through my research and writing the thesis.

I would like to thank the examining committee, Dr. Ying Cui, Dr. Linglong Kong, Dr. Ivan Mizera, and Dr. Yan Yuan for their questions and comments.

I am also grateful to my colleagues and friends at the Department of Mathematical and Statistical Sciences, who helped and made useful comments on my research and thesis: Qian Shi, Sile Tao, Wei Tu, Dengdeng Yu.

Last and most importantly, I would thank my husband, Jian Deng's unconditional support for my life and study, without which the completion of my thesis would not be possible. To him I dedicate this thesis.

Contents

List of Tables

List of Figures

Chapter 1

Background and Motivation

1.1 Introduction

Maternal nutritional condition during pregnancy is significant as it plays a critical role in the health of both the mother and the fetus [6, 115]. In an ideal pregnancy, the mother maintans good health condition, and gives birth to a full-term, healthy infant [18]. Kaiser et al. identified the key components leading to a healthy pregnancy outcome: healthy prepregnancy weight, appropriate weight gain, diverse of dietary consumption, and appropriate prenatal vitamin supplementation [60]. However, many pregnant women do not get proper nutrition and they may suffer from undernutrition or improper weight gain [22]. These conditions are proved to be linked to an increased risk of various health issues of both the mother and her offspring [15, 17, 75, 96]. Protein-energy under-nutrition, which is associated with impaired fetal growth, is a major health issue in developing countries [78]. In developed countries, over-nutrition is more common [76], and it is associated with an increased risk of preeclampsia, gestational diabetes, and a higher cesarean section rate. A child born to an obese mother is more likely to be born large for gestational age, and to develop type II diabetes in later life. Some micronutrient deficiencies affect both the developing and developed countries. For example, the World Health Organization estimated that there is a high prevalence of iron deficiency in pregnant women around the world. Iron deficiency causes anemia and put women at a higher risk of perinatal mortality and morbidity, and is associated

with low birth weight and preterm birth [3].

Dietary intake accounts for the majority of nutrient intakes during pregnancy. It includes the intakes of macronutrients (carbohydrate, fat, and protein) which constitute the energy consumption, and the intakes of micronutrients which are crucial to the maternal metabolism and fetal development. Among the literatures on dietary intakes during pregnancy, most studies were conducted in developing countries [7, 57] or among low-income groups [110].

Weight gain during pregnancy is an important proxy of maternal nutrition. Health professionals have agreed on a uniform guideline of weight gain during pregnancy [123]. Pregnant women gaining weight within the recommended range has the best chance to have an optimal pregnancy outcome [84]. However, in developed countries such as the United States and Canada, excessive weight gain occurs to a large portion of the pregnant population [51, 68, 98], which calls for effective clinical interventions. Existing intervention practices focused on the counseling of dietary intake and physical activities based on the monitoring of weight gain [90, 86, 92], where the graph of weight growth of each individual was obtained and continually compared to a set weight gain goal. In both research and clinical communities, there is little work to statistically model the individual-specific weight growth curve (trajectory).

To fill in the gaps of the research work in developed countries, the Alberta Pregnancy Outcomes and Nutrition (APrON) study recruited a large cohort of pregnant women in Alberta, Canada. The study aims to investigate the relationship between maternal nutrient status and obstetric outcome as well as maternal and child health and development [61], so as to assist making appropriate dietary recommendations or food policy change.

1.2 Motivation

The APrON study used two versions of 24-hour food recall instruments to collect the nutrient intake data from participants, an interviewer-administered instrument and a self-administered, web-based instrument[61]. The two instruments are different by the administration type and the algorithm of extracting nutrient intakes from food items consumed. The interviewer-administered instrument has been developed and used for a long time [104], whereas the self-administered, web-based instrument is relatively new [73, 47]. Therefore, one of the objectives of my thesis is to determine whether or not the nutrient intakes extracted from the two instruments can be pooled together for further modeling, especially the "key nutrients" important for the fetal growth. Calibration by regression is considered where the nutrient intakes measured by the interviewer-administered instrument are responses and the nutrient intakes measured by the web-based instrument are predictors. I aim to develop a new calibration method for the key nutrients, which improves the performance of the simple linear regression based calibration.

To facilitate the clinical intervention of weight gain during pregnancy, another goal is to estimate the weight growth trajectory of each individual. The weight records and nutrient intakes were collected sparsely, up to five times for weight and up to three times for nutrient intake, and longitudinally through pregnancy. To model the weight growth trajectory as a smooth function, traditional longitudinal data analysis techniques, such as mixed-effect models, may not be suitable for such sparse data. I propose to apply functional principal component analysis by conditional expectation, which has been shown to work well for sparse data [124]. The estimated trajectories are compared with the estimation from mixed-effect model. Total weight gain is estimated from the trajectory, and the association between the estimated weight gain and prepregnancy

baseline information, nutrient intakes, physical activities are studied through multiple linear regressions.

1.3 Thesis Structure

In Chapter 2, I investigate the agreement between the two dietary intake assessment instruments employed in the APrON study, and examine the possibility of pooling the intake from the two instruments. I propose a new calibration method and report the improved R^2. In Chapter 3, I illustrate a new approach for weight growth trajectory modeling and compare the new approach and the traditional mixed-effect model. I then model the estimated total weight gain by the predictors including prepregnancy body mass index, nutrient intakes and physical activities. A predictive model is obtained for the total weight gain. A discussion and potential future work follow in Chapter 4.

Chapter 2

Machine Learning Based Calibration of a Web-based and Interviewer-administered 24-hour Food Recall

2.1 Introduction

2.1.1 Maternal Nutrition

Maternal nutrient intakes during pregnancy can significantly influence the mother's health, pregnancy outcomes, as well as the fetus/infant's well-being. For example, extra energy intake is necessary for pregnant women to support the development of the fetus as well as compensate increased maternal basal metabolism rate. Diet low in fiber and high in glycemic load during pregnancy has been linked to increased risk of gestational diabetes mellitus [125]. Adequate intake of micronutrients are important during pregnancy, too. Pregnant women with low folate intake have a greater risk of preterm delivery and low birth weight infant as reported in [100]. Folate intake is also confirmed to prevent neural tube defects [41]. Observational studies show that low dietary vitamin D intake during pregnancy is associated with increased risk of type

I diabetes and wheezing in offsprings [35, 14]. Therefore, efforts are taken by health professionals to improve maternal nutrition, including achieving appropriate energy intake and ensuring adequate intake of specific nutrients to meet maternal and fetal development requirements [82].

Nutrition intake mainly come from two sources: dietary intake and supplement usage. The former accounts for the overall energy and all the macronutrients (carbohydrate, protein, fat, and alcohol) intake. In general, a balanced diet provides adequate micronutrients (vitamins and minerals) for most people. However, some reports suggest that usual dietary intake of certain nutrients is inadequate to meet the needs of pregnant women; so the supplemental usage has been widely recommended during pregnancy [82]. However, research work showed that even with supplement usage, it may not fully compensate for dietary deficiencies [41, 65]. Therefore, studying the dietary pattern is necessary to help promote healthy dietary habits among pregnant women.

The majority of existing dietary pattern studies were either from developing countries [7, 57, 81, 97, 108, 116] or focused on low-income population [55, 109, 110]. Only a few were conducted in developed countries and targeted higher socioeconomic status (SES) population [26, 38, 77, 95, 102]. Moreover, the studies on the higher SES population did show deficiency in some particular nutrients, such as folate [95], calcium, iron, and vitamins D and E [38]. These studies suggested that the significance of nutrition intakes in these well-nourished population may be under-evaluated. In Canada, several large nutrition studies took place in the beginning of 21st century [26, 80]. Most of them studied specific nutrients, and similar conclusions were yielded, that certain nutrients intakes were lower than recommended among Canadian pregnant women, such as (n-3) fatty acid [26] and vitamin D [80].

2.1.2 Nutrient Intake Assessment

The use of an appropriate intake assessment instrument is critical in nutrition epidemiology. Commonly used instruments include food frequency questionnaires (FFQ) , food records, 24-hour food recall, which are designed to fit different situations [104].

FFQ asks participants to report the consumption of food items over a defined period of time, including consumption frequency as well as the portion sizes [104]. It is suitable for large-scale survey because its self-administered and machine-readable nature makes it cost-effective [91]. They are widely used to assess dietary intakes for a certain time span, such as three months, 12 months, etc. [34]. However, it is not detailed enough, and only captures the usual dietary intakes over a longer terms. It also relies on the participant's ability to form and report a generic tendency of diet, which can be difficult for many people. In fact, studies using reference recovery biomarkers indicated that FFQ is subject to substantial measurement errors which lead to biases [16].

The food record instrument asks the participants to record the consumed foods and beverages throughout of the reporting day. It is detailed, real-time, and captures current short-term diet. However, converting the reports which are usually handwritten into quantitative intakes, which is referred as coding of a questionnaire, can be quite time consuming and expensive. For the participants, the real-time recording can be a heavy burden and discourages them from responding. These drawbacks make it difficult to use on large scale studies [112].

In the classic, interviewer-administered 24-hour food recall, participant is interviewed by a trained professional, usually a dietitian, to report all food items consumed in the most recent 24 hour period. It is also detailed and short-term oriented as food record, but less expensive [34]. An improved version is the "multiple pass" 24-hour recall. It contains more than one steps, which are referred as "passes" of revisiting the dietary information: (i) a quick list of foods consumed; (ii) a detailed record; (iii)

reminder for food that might be forgotten, and (iv) a review of all the records and further details, such as portion sizes. This method has better precision than a general 24-hour food recall [19].

The 24-hour food recall has less report burden for the respondents and is more feasible for large-scale studies. However, the need for trained professionals who administer the interviews make it comparably expensive, especially for large cohort studies. It also takes a considerable amount of time and effort to code the recorded food and beverages (done by an algorithm). Due to these limitations, the web-based, self-administered 24-hour food recall was introduced. Compared to the interviewer-administered 24-hour food recall, the web-based instrument has lower study cost as trained professionals are no longer needed for the interview administration. Participants do not need to visit the clinic in person, giving them more convenience and freedom. Moreover, it has build-in nutrient "extraction" software that codes the responses. One example of web-based, self-administered 24-hour food recall is the widely used Automated Self-Administered 24-hour dietary recall (ASA24), the first version of which was launched in 2011 [107]. ASA24 has been validated in an adult group [63], showing good agreement with the interviewer-administered version. In Canada, a similar, yet much earlier version is the Food Behavior Questionnaire (FBQ), developed at the University of Waterloo in 2003 [73]. It has been further modified by researchers at the University of Alberta to assess meal behaviors in 2009 [105]. The collected results from FBQ instrument are comparable to the interviewer-administered one in a group of adolescents for total calories, and key nutrients (carbohydrates, protein, fat, calcium, iron, vitamin B6, B12, C, D, folate, zinc, sodium, and potassium) [47].

2.1.3 Instruments Used in the APrON Study

As a large-scale cohort study, APrON had to balance accuracy and cost for the food intake assessment. The dietary pattern of pregnant women often changes from time to time [6, 94]. So the instrument should capture short-term dietary intake with reasonable cost. Therefore, the 24-hour food recall questionnaires, which is more suitable for short-term studies than the FFQ and less expensive than the food record, was employed.

The interviewer-administered 24-hour food recall was used initially. Questionnaires was collected by professionally trained research assistants, using a multiple pass 24-hour recall questionnaire. Visual food models were shown to help women estimate portion sizes. Probes were provided to get more accurate information of food, including details such as cooking methods, location and time of eating, and food brand names. All information was reviewed back to the women to ensure information was correctly recorded. Nutrient intake information, including both macro and micronutrients, were calculated by the Food Processer software (version 10.6.0, 2010, esha Research, Salem, OR).

However, after interviewing the first cohort of 600 participants, it was observed that the respondence rate for in-person visits were lower than the self-reported questionnaire to be completed at home [61]. So a web-based instrument was introduced to increase the completion rate for the rest of cohort. The web-based instrument was based on a validated online FBQ as in [73]. All foods and beverages consumed the previous day were recorded by selecting items from a list of approximately 800 foods and beverages. Built in cues and response options, such as portion size images and beverage intake reminders, were provided throughout the online recall process. Additional visual cues (the virtual meal plate and meal summary) provided participants the opportunity to revise items in the recall. This online tool was also modified to be used with a cohort of

pregnant women by including additional food items that had been frequently reported in early interviews with pregnant women; as well as some ethnic foods and recipes. Subjects received brief instruction by a trained research assistant before completion of the web-based instrument.

2.1.4 Calibration Between Two Instruments

The dietary intakes measured through food questionnaires are well-known prone to various measurement errors, including errors from the memory and perception of the subjects, and the coding process of the questionnaires [33]. Though the two instruments employed in APrON have been validated before in Canadian adolescents [105], they have not been validated in pregnant women. A common issue in validation of dietary intake questionnaires is the lack of a gold standard, which measures the true intake [33]. So a related, but different approach, calibration is often considered. In the calibration between two instruments, the intakes measured by one instrument is estimated from the intakes measured by the other instrument, such that the effects of measurement bias can be are corrected. The main objective of this chapter is to study the agreement between the intakes of several key nutrients obtained from these two survey methods: the web-based survey and the interviewer-administered questionnaire, and investigate the possibility of calibrating and pooling the nutrients intake from the two instruments for further modeling.

More specifically, I aim at calibrating the calorie contribution of the three macronutrients, protein, fat, and carbohydrate, as well as the micronutrients vitamin C, vitamin D, folate, iron and choline. These "key nutrients" play important roles in the fetus development, especially neurodevelopment [36, 41, 48, 106]. The macronutrients constitutes energy intake, and each of them has its unique metabolic values. Carbohydrate is the primary source of glucose, a major fuel for fetal growth [99]; protein intake is

needed for the placenta development, and is associated with both placenta and birth weight [39]; fat is the source of essential fatty acids, which are necessary for the fetus brain development [23]. Both vitamin C and folate are shown to be related with neural tube defects [103]. vitamin D is known to affect fetus brain development by influencing cell differentiation, neurotrophic factor expression, etc. [28]. Iron deficiency is a primary cause of anemia, which is believed to expose pregnant women to greater risk of perinatal mortality and morbidity, and can affect the cognitive development of offspring [3, 70]. Prenatal plasma choline is shown to be related to the early cognitive development of infants [121].

Previous studies on the validity and calibration of dietary intake instrument usually does not involve much complexity other than correlation coefficients or linear regression [2, 12, 58, 73]. Johansson et al. provided the calibration equation between a 24-hour food recall and an FFQ. The authors used Spearman's correlation coefficient to measure the agreement between these two instruments. The calibration equation was obtained by simple linear regression (SLR), with the intakes measured from the FFQ as predictor and the intakes measured from the 24-hour recall as response. The slope of this regression line is named "calibration factor". As a nonparametric measure of rank correlation, Spearman's correlation may overestimate the agreement between the two instruments, especially when the calibration equation is a parametric linear model. Pearson's correlation was employed in [2], which represents the performance of the SLR-based calibration. The work of Briggs et al. utilized more indicators of agreement such as Bland-Altman's plot [12]. In a Bland-Altman's plot, the mean difference between the two methods of measurement (the "bias") is calculated, and the 95% limits of agreement are defined as 1.96 standard deviations below or above the mean difference. Good agreement can be concluded if the 95% limits include 95% of differences between the two measurement results. The calibration was also based on

SLR.

SLR uses the intake of only one nutrient as predictor, and the correlation between different nutrients is not considered. Nevertheless, one food item usually contains more than one nutrients, and a nutrient is rarely consumed alone. So in food intake assessments, the correlation between the intakes of nutrients can be very complicated. The commonly existing reporting bias in either instrument, will probably result in that the intakes of more than one nutrients are affected. In this case, an SLR model may not capture the relationship well.

In common daily diets, certain nutrients tend to be consumed together. For example, dietary vitamin C often comes from dark green leafy vegetables and fruits, which are often rich in fiber and folate, too. So the intake of vitamin C, folate, and fiber are highly correlated [49]. Meanwhile, it is reported that people with high-fat or high-alcohol diet tend to consume less fruit and vegetables, making the intake of fat and alcohol negatively correlated with vitamin C, folate, and fiber. A food questionnaire does not evaluate the intake of nutrients independently, rather, it captures the entire dietary profile of the participant, where various correlation of nutrients are present. Therefore, it may be useful to include multiple predictors in the calibration. Better performance of the calibration equation, e.g., higher multiple correlation coefficients, may be achieved by considering the comprehensive dietary profile information.

Using multiple predictors can brings up new issues. Numerous nutrient intakes are computed from a food recall. Using the intake information of all nutrients in the estimation of a certain nutrient can result in overfitting, i.e., the model fits to the noise rather than the true signal. For example, when estimating the intake of vitamin C measured in instrument A, the variance of the model can be reduced by utilizing all nutrients' intake information measured in instrument B, including carbohydrate, protein, calcium and so on. However, estimation of the coefficients for these nutrients

could be driven by some random errors (noise) and result in bias in the model. The estimated model may fit current data set well, nonetheless when new data come in, the new predicted vitamin C may be not accurate as it is driven by the unrelated nutrients' intakes. Therefore, a variable selection procedure is needed in the multi-predictor modeling, to achieve a balance in the bias-variance tradeoff. This will also provide the advantage of easier interpretation. A well-studied and popular tool for variable selection in linear regression models is the Least Absolute Shrinkage and Selection Operator (LASSO) [114]. It is easy to implement and also provides selection flexibility through parameter tuning [114]. Details are to be provided in Section 2.2.2.

2.1.5 Our Contribution

We study the agreement between the intakes of nutrients measured from web-based and the interviewer-administered 24-hour food recall instruments employed in the APrON study. The nutrient intakes from the web-based instrument are treated as predictor, and the nutrient intakes from the interviewer-administered instrument are used as response. Different techniques are applied to improve calibration using the classic SLR method as a benchmark. Our work can be summarized as follows:

- Bland-Altman plot of the total calories are given, showing acceptable agreement between the two instruments.

- For macronutrients, I propose the energy-adjustment based calibration. Two energy adjustment approaches are investigated, and one of them improves the R^2 significantly.

- For several micronutrients of interest, namely vitamin C, D, iron, folate, and choline, the calibration equations with multiple predictors performed significantly

better than SLR. The R^2 for folate is remarkably improved when LASSO was used.

2.2 Materials and Methods

2.2.1 Materials

To study the agreement between the two instruments, 58 female volunteers were recruited by a friend-to-friend referring method. The inclusion criteria were: child-bearing age (16-40 years old), able to comprehend written/spoken English, and computer-literate. They were all non-pregnant. Each volunteer was asked to finish the web-based instrument and interviewer-administered instrument for the same 24-hour period, and the entire protocol took less than 90 minutes. The order in which each volunteer finished the two instruments was randomized.

The intake measurements from the two instruments has 25 nutrients in common. Alcohol is included in the interviewer-administered instrument but not in the web-based. To account for this, I subtract the calories attributed to alcohol from the total calories intake in the interviewer-administered instrument. After coding the food intakes for the two instruments, I add another inclusion criteria, that the macronutrient intakes, including carbohydrate, protein and fat, are consistent with the total calories, i.e., the total calories calculated from all macronutrients are within 90-110% of the directly calculated total calories. Applying the inclusion criteria, we have 55 valid records. The nutrients and their units are listed in Table 1.

Table 1: List of the nutrients and their units in the web-based and interviewer-administered instruments.

Nutrient	Unit	Nutrient	Unit
Fat	kCal	Vitamin B3	mg
Saturated fat	kCal	Vitamin B6	mg
Protein	kCal	Vitamin B12	mg
Carbohydrate	kCal	Vitamin C	mcg
Fiber	g	Vitamin D	IU
Sugar	g	Folate	mcg
MonoFat	g	Calcium	mg
PolyFat	g	Iron	mg
Transfat	g	Potassium	mg
Cholesterol	g	Sodium	mg
Vitamin A	IU	Caffeine	mg
Vitamin B1	mg	Choline	mg
Vitamin B2	mg		

2.2.2 Method

To assess the agreement between the two versions of questionnaires, I first obtain the Bland-Altman plot of the total calories captured from the two instruments. Bland-Altman plot was proposed by Altman and Bland in [4] to compare two methods of measurement. More specifically, the method is to decide whether one method could replace the other without loss of much accuracy. The difference and average is calculated for each pair of measurements, and plotted on the y and x coordinate respectively. For a sample S with two measurements S_1 and S_2, the coordinate is $\frac{S_1+S_2}{2}$ on x, and S_1-S_2 on y. The mean value of the differences, $\overline{S_1-S_2}$ is then the "relative bias", plotted as a dashed line. The standard deviation of the differences, $s = SD(S_1 - S_2)$ is the estimate of standard deviation of bias. Then the hypothesis that the relative bias is zero can be tested using a one-sample t-test. Confidence band is given on the plot (above and below the "bias"). The 95% reference interval $[\overline{S_1-S_2} - 1.96s, \overline{S_1-S_2} + 1.96s]$ is termed "limits of agreement". If the width of the reference interval is not acceptable

in the clinical goals, then the two method cannot be interchanged [37].

For the calibration procedure, the regression model is referred as calibration equation. In this chapter, I use the nutrient intakes measured in the web-based instrument and in the interviewer-administered instrument as predictors and responses respectively. The R^2 of the model is the proportion of variance in the response explained by the predictors. The square root of R^2 is the Pearson's correlation coefficient. They are indicators of the model performance: higher R^2 or Pearson's correlation coefficient suggests better calibration. Specifically, we are interested in the calibration of total calories intake, and the "key nutrients" as mentioned in Section 2.1.4, carbohydrates, fat, protein, vitamin C, vitamin D, folate, iron, and choline as response.

The improvement of the calibration is represented by an increase in the R^2 or Pearson's correlation coefficient. One possible approach to increase R^2 is to apply energy adjustment. Most nutrient intakes tend to be positively correlated with total energy intake. Thus part of the variation of nutrient intakes can be explained by the total energy intake variation instead of the diet composition. That is, the nutrient intake is confounded by total energy intake [40]. This confounding effect is especially strong for macronutrients, which contributes to total energy directly [118]. Therefore, I employ the energy adjustment approach for the macronutrients. There are two popular ways for energy adjustment [119]. One is the proportional adjustment, or nutrient density method. For each subject, I consider only the proportion of energy obtained from a macronutrient, i.e., all subjects has the macronutrients intake proportion summing up to 1. The other way is residual adjustment [62]. For a simple linear regression model with a macronutrient intake as response, and the total calories intake as predictor, the adjusted intake is defined to be the residual intake. Suppose we have subject S, with two measurements for each macronutrient: fat intakes F_1, F_2, protein intakes P_1, P_2, carbohydrate intakes C_1, C_2 and total calories T_1, T_2, where subscript 1 indicates that

the intake was measured in the interviewer-administered instrument, and subscript 2 indicates the intake was measured in the web-based instrument. The measurements approximately satisfy

$$9F_i + 4P_i + 4C_i = T_i, \text{ for } i = 1, 2.$$

Then in the proportional adjustment method, the adjusted intakes are now $F_{\text{adj},i} = F_i/T_i$, $P_{\text{adj},i} = P_i/T_i$, $C_{\text{adj},i} = C_i/T_i$ respectively. Taking fat as an example, the calibration equation is

$$\hat{F}_{\text{adj},1} = \hat{\beta}_0 + \hat{\beta}_1 F_{\text{adj},2} = \hat{\beta}_0 + \hat{\beta}_1 \frac{F_2}{T_2}.$$

In the residual adjustment method, the adjusted fat intake is according to the regression model

$$\tilde{F}_{\text{adj},i} = F_i - \hat{F}_i = F_i - (T_i\hat{\gamma_1} + \hat{\gamma_0}),$$

where $\hat{\gamma_1}$ and $\hat{\gamma_0}$ are the estimated coefficients in the regression model

$$F \sim \gamma_0 + \gamma_1 T.$$

Estimated adjusted intake of fat are similarly defined:

$$\hat{\tilde{F}}_{\text{adj},1} = \hat{\tilde{\beta}}_0 + \hat{\tilde{\beta}}_1 \tilde{F}_{\text{adj},2}$$

Using multiple predictors is another possible approach to improve the calibration. I regress the nutrient intake measured in the interviewer-administered instrument onto multiple nutrient intakes measured in the web-based instrument. By extending the range of predictors, I am essentially modeling the whole diet profile measured in the interviewer-administered instrument by the diet profile measured in the web-based one.

As discussed in Section 2.1.4, one food item may contain various nutrients, and the nutrient intakes correlates with each other in a complicated manner. Therefore using the whole profile as predictor may be more suitable, as some of the correlations may be useful in the modeling of a certain nutrient.

A popular tool to associate one response variable (in our scenario, the intake of certain nutrient measured in the interviewer-administered instrument) and multiple predictor variables (the nutrient intakes measured in the web-based instrument) is the multiple linear regression (MLR) through ordinary least square (OLS). MLR works properly when there are no redundant predictors, or no multi-collinearity. As shown in Table 1, 25 nutrient intakes are captured in both instruments. Some of intakes of the nutrients are mostly likely to be correlated, as discussed in Section 2.1.4. The large number of correlated nutrients could induce high redundancy. Meanwhile, including too many predictors may lead to overfitting and affect the prediction performance of the model. Regression based on regularization (penalization) are widely used to prevent overfitting. For a sample set $(X_i, y_i) = (x_{1,i}, ..., x_{p,i}, y_i)$ with $i = 1, ..., n$, unlike usual OLS regression which solves the coefficient $\beta = (\beta_0, \beta_1, ..., \beta_p)$ by minimizing the quadratic loss:

$$L(\beta) = \|Y - X\beta\|_2^2 = \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{j,i}\beta_j \right)^2,$$

the regression with regularization adds a penalty term in the loss function:

$$L_{\text{reg}}(\beta) = \|Y - X\beta\|_2^2 + \lambda P(\beta), \tag{2.2.1}$$

where λ is a tuning parameter to be chosen by user, and $P(\beta)$ is the penalty term which is a function of β. By adding the penalty term we are shrinking P, so the minimization

is called a shrinkage procedure.

When $P(\beta) = \|\beta\|_2^2 = \sum_{j=1}^{p} \beta_j^2$, the shrinkage is called ridge regression (RR) [53]. RR is easy to implement as a quadratic optimization, and shrinks each coefficient towards 0 [83]. The developed model is not sparse, where sparse means that some coefficient β_j's are zero, i.e., some predictors are not included in the developed model. If variable selection is needed, the LASSO method is more suitable, which uses $P(\beta) = \|\beta\|_1 = \sum_{j=1}^{p} |\beta_p|$. Tibshirani explained in [114] that this penalty term results in zero coefficients for some covariates, namely the solution $\hat{\beta} = (\hat{\beta}_0, ..., \hat{\beta}_p)$ has $\hat{\beta}_j = 0$ for some $j \in \{0, ..., p\}$.

The penalty coefficient λ, which is often referred as a "tuning parameter", decides the level of shrinkage of the estimates. A large λ value corresponds to more shrinkage of $\|\beta\|_1$ towards zero, and more coefficients β_j's are forced to be zero. Thus a simple linear model with fewer predictors will be obtained. In the bias-variance tradeoff, it results in an estimate with less model complexity or bias, but larger variance. Vice versa when using a small λ, the developed model has smaller square error, and more predictors (corresponding to larger $\|\beta\|_1$) may be selected, corresponding to an estimate with smaller variance but larger bias. By choosing different λ values, we get different numbers of variables.

The usual practice for choosing the tuning parameter λ is through a $k-$fold cross-validation (CV) method which minimizes the prediction error. The whole data set is divided into k "folds" of data, and each time we use $k-1$ of them to make a training set to estimate the model, and the k-th fold as validation set to calculate the error of the estimated model. After k folds of validation, the residual is calculated for each instance, and the mean the squared residuals across all folds are defined as the CV error. Naturally we may want to choose the λ with the least mean CV error $E_{\min}$, denoted by $\lambda_{\min}$. However, $\lambda_{\min}$ contains uncertainty resulted from the randomization

of the folds. There can be different choices of λ with close CV errors. An empirically widely used choice is the "one standard error" rule.

In this "one standard error" rule, we first average the CV errors in each fold, denoting them as $E_1, ..., E_k$. Then we compute the standard deviation, $sd(E)$ of these K averaged CV errors. and the standard error, $se(E)$ is estimated as $sd(E)/K$. We choose the largest λ such that the CV error is within $(E_{\min}, E_{\min} + se(E))$, so that we have more regularization and a simpler model is achieved. This rule was first proposed by Breiman et al. to define an optimal tree size for a classification tree [11]. It has been suggested by Hastie et al. for cross-validation [50] and used by many other model select ion practices, e.g., Guo et al. used it to improve the performance of classification of cancers using a reasonably reduced subset of genes [45]. I adopt this parameter tuning rule, and then compare the CV-R^2 of simple linear regression and the CV-R^2 of LASSO.

2.3 Results

The Bland-Altman plot of the total calories is shown in Figure.1. There is a slight trend of under-reporting for the web-based instrument, as the mean of the difference is above zero, but the under-reporting tendency is not severe. The points are randomly scattered horizontally, and the difference is not influenced by the mean of total calories. Most points (52 out of 55, 94.5%) are within the 95% limits of agreement. The rest three points are outside the limits of agreement, showing weak agreement for these subjects. The order in which the subject finishing these two instruments do not seem to play a role in under/over-reporting, since the triangles and diamonds are randomly scattered around the mean difference line.

Traditional questionnaire calibration employs the Pearson correlation coefficient,

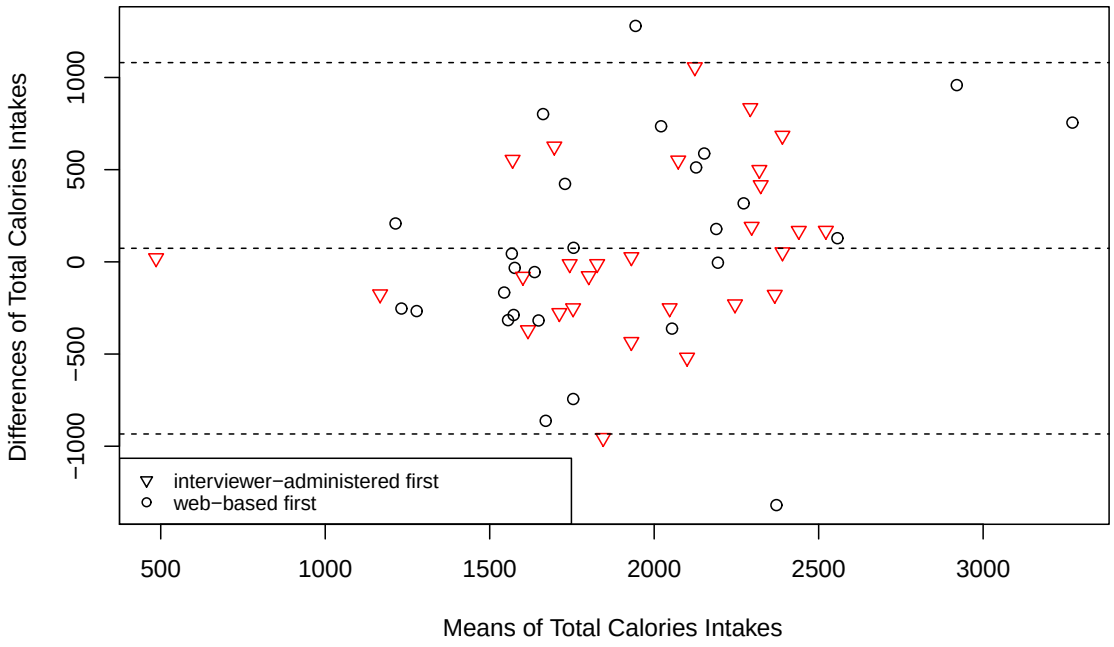


Figure 1: Bland-Altman plot of the total calories intake from the two instruments, interviewer-administered minus web-based. Middle dashed line is the mean difference, showing a light under-reporting trend in the web-based instrument. Upper and lower dashed line represents the 95% limits of agreement. Points outside the lines represents weak agreement.

Table 2: R^2 and Pearson's correlation coefficient r of the calibration based on simple linear regression (SLR)

Nutrient		R^2	r
Macronutrient	Fat	0.22	0.47
	Protein	0.28	0.52
	Carbohydrate	0.39	0.63
Micronutrient	Vitamin C	0.58	0.76
	Vitamin D	0.48	0.70
	Folate	0.03	0.16
	Iron	0.49	0.70
	Choline	0.19	0.44

which can be obtained from a simple linear regression. The R^2 and Pearson's correlation coefficient r are listed in Table 2. For the macronutrients fat, carbohydrates and protein, the Pearson's correlation coefficients (square roots of the R^2's) are acceptable ranging from 0.47 to 0.63. However, none of them has $r > 0.65$, which means a direct pooling of the two instruments is probably not appropriate. Fat has the least correlation, likely due to well-known selective underreporting of fat intake in self-administered web-based food intake questionnaires [71]. For the micronutrients, calibration based on SLR shows correlation coefficient for vitamin C (0.76), vitamin D(0.70), iron (0.70), folate (0.16), and choline (0.44). The agreement is good for vitamin C, D and iron, with $r > 0.7$, suggesting that pooling the results is possible. However the correlation is not ideal for other nutrients. The correlation for folate is especially low, indicating a probable issue in the measurement of folate from one of the instruments. I thus seek for amendments of the regression models.

Calibration of Macronutrients with Energy Adjustment

While the intakes as proportion of total calories describes the dietary composition, the nutrient residuals provide a measure of nutrient intake uncorrelated with total

Table 3: Comparison of Pearson's correlation coefficients r of method 1: SLR with no energy adjustment, method 2: SLR with proportional adjustment, method 3: SLR with residual adjustment. Method 2 greatly improves r, but improvement from method 3 is limited.

Nutrient	r		
	Method 1	**Method 2**	Method 3
Fat	0.47	**0.56**	0.52
Protein	0.52	**0.73**	0.54
Carbohydrate	0.63	**0.67**	0.57

Table 4: Performance comparison of the simple linear regression (SLR) and LASSO

Nutrient	SLR	LASSO	Variables Selected from the to Estimate
Vitamin C	0.50	0.52	Vitamin C, Potassium
Vitamin D	0.43	0.49	Saturated fat, Vitamin D, Choline
Folate	0.0033	0.46	Fiber, Vitamin B12, Iron, Potassium
Iron	0.40	0.43	Fiber, Iron
Choline	0.083	0.22	Vitamin B6, Choline

energy intake. In our study, the proportional adjustment greatly improved the Pearson's correlation coefficients of macronutrients' intakes from the two instruments, as shown in Table 3. The residual adjustment does not improve the correlation except for fat. The energy distribution of the macronutrients are more consistently reported, suggesting possible biases in the portion sizes in reporting.

Calibration of Micronutrients with Multiple Predictors

For the micronutrients, I use LASSO for variable selection. The R^2's are shown in Table 4. Note that the SLR R^2's are also 5-fold cross-validated on the same fold cuts as the CV-LASSO. So their values are different from the R^2's from direct SLR reported in Table 2.

For all the micronutrients considered, the CV-R^2's for LASSO are generally better than those of SLR. Vitamin C and iron have the least improvement, probably because their already high correlation by SLR, so the room of improvement is limited. There

are only one more variable selected other than the nutrient itself. Potassium is included to estimated vitamin C intake, accounting for the leafy greens and some fruits rich in both. Most non-heme (plant-sourced) iron rich foods are rich in fiber (e.g., spinach, asparagus), which explained the inclusion of fiber in the model that estimates iron. R^2 for folate has improved dramatically from 0.0033 to 0.46, and the selected predictors do not include folate from the web-based instrument. This suggests that, if there is a problem in capturing folate is in the web-based instrument, it is possible to discard the measurement and get a reasonable estimate from other nutrients. The selected variables (fiber, vitamin B12, iron, and potassium) are also consistent with the well-known nutrition facts: dietary folate are mainly from dark green leafy vegetables and enriched flour; the former is typically rich in fiber and potassium, and the latter is usually also enriched in vitamin B's. There is some improvement in R^2 of vitamin D and choline. The selected variables also reflects some common sources of the nutrients involved. Typical examples include fatty fish and eggs, which are rich in fat, vitamin D, vitamin B6 and choline. It is also noted that choline has a low correlation for SLR (0.083 vs 0.19). After experiments on different fold cuts, we found that it was accounted for by the randomization in the fold cuts, considering the SLR without CV in Table 2. Experiments with different fold cuts for 5-fold CV had variable R^2s.

2.4 Discussion

Data from food intake assessment instruments are often subject to report and measurement errors [16, 34]. Studying the magnitude and relative direction of the errors between two measurement instrument is essential to assess their impact for the feasibility of pooling these measurements. In our targeted study, APrON, two 24-hour food recall instruments were used - an interviewer-administered one and a web-based

one. The aim of this chapter is to explore the agreement between these two instruments, and to estimate the calibration equation which has the nutrient intakes from the interviewer-administered instrument as response, and the web-based one as predictor, and then assess the possibility to pool the nutrient intakes from different instruments.

The Bland-Altman plot in Figure 1 of these two instruments' total calories measurements show that there is a slight tendency of under-reporting the total calories intakes in the web-based instrument. The 95% limits of agreement is about 1000 (Cals) above and below the mean difference, indicating the difference between two measurements for the same individual is on the scale of hundreds of calories. Most points (52 out of 55, 94.5%) are within the limits of agreement. This suggests that under-reporting in the self-administered web-based questionnaire is not severe among the sample of participants, but the variances in individuals can be relatively large. There could be various reasons. Our participants in the study were recruited on campus who were likely to be more educated, and thus less likely to under-report when self-reporting, as noted in previous studies [21]. However, the instructions they received before or during completing the self-administered web-based instrument may not be detailed or thorough enough. For an individual participant, remembering and reporting the full dietary profile can be therefore difficult, and the report or measurement error could be large. These results calls for a further improvement of the web-based instrument and instructions of the instruments, to help the participants remember and report correctly and consistently. Of equal importance is the development of validation and calibration tools to minimize the relative error between two instruments.

For the calibration, I first use SLR to estimate the interviewer-administered intake by the same nutrient intake from the web-based instrument. It gives reasonable correlations for macronutrients including fat, protein, carbohydrate and micronutrients vitamin C, D, and iron, with $r > 0.45$, weaker correlation for choline ($r = 0.44$), and

poor correlation for folate ($r = 0.16$). The result of folate indicates that there could exist a problem of the folate intake coding process in one of the instruments. This indication helped our collaborators getting further insight into the instruments, and it was revealed that there existed an algorithm problem in the folate intake calculation in the web-based instrument.

After using SLR, I seek to improve this traditional calibration approach. One of the attempt for improvement is the energy-adjustment approach for macronutrients, as the macronutrients are confounded with total energy intake. By using the proportional adjustment, the Pearson's correlation coefficients of macronutrients fat, protein, carbohydrate increased from 0.47, 0.52, 0.63 to 0.56, 0.73, 0.67. This suggests that the energy components' distribution reported by participants are more consistent than the absolute energy intake of each macronutrient. Similar issues were also reported by other questionnaire validation studies [5]. Possible reason is the errors in evaluation of the portion sizes of food items.

The other attempt for improvement is the linear regression with LASSO regularization for micronutrients. When building the calibration equation for a certain nutrient, I considered all 25 nutrients from the web-based instrument, so that the modeling of the entire dietary profile is possible. To avoid overfitting and allow easier interpretation, a variable selection process by LASSO was hereby proposed. I improved the CV-R^2's of vitamin C by 0.02, vitamin D by 0.06, folate by 0.46, iron by 0.03, and choline by 0.14. Compared to the R^2 of choline without CV (0.19), the weak correlation of choline in SLR was due to the uncertainty in fold cuts. LASSO method shows an clear advantage for the folate intake calibration, which was poorly captured in the web-based instrument. For those nutrients whose intakes are more consistent across the two instruments, such as vitamin C, D and iron, the improvement of LASSO method is limited.

We conclude that, for the dietary intake measurement of child bearing age women, the agreement between the two instruments, web-based and interviewer-administered, is acceptable for the nutrients vitamin C,D, and iron. Choline intakes need further investigation. If the interviewer-administered instrument is used as a gold standard, the web-based instrument can be adopted. Especially in large cohort studies, employing of the web-based instrument can reduce the time commitment for the participants, cost of data collection and processing, and may increase the completion rate. However, it may not be feasible to directly pool the results from the two instruments together for other analysis, as the differences between the intakes are not negligible. For macronutrients, proportionally energy-adjusted intakes have stronger correlation between the two instruments, and would be recommended if pooling is considered. The web-based instrument was proved to be problematic in capturing folate intake, so it is suggested to use the calibration equation obtained by LASSO to estimate the folate intake.

Despite the promising results and significant improvements of our proposed calibration methods, our study materials and methods may have some limitations, which can be cast into the following two aspects. (i) The difference between the validation population and the targeted study population. APrON is a study of pregnant women, but subjects in this validation study were all non-pregnant. The diet of a generic child bearing aged woman can be different from a pregnant women [6, 94]. Especially for women with higher SES, who constituted a large part of the APrON cohort, they were more prone to make a change toward a healthier diet (e.g., higher folate intake) and avoid some ingredients, such as alcohol or unpasteurized food, which are possibly harmful to the fetus [67, 20]. If an instrument has a problem capturing related nutrients, the deviation between the diets of pregnant and nonpregnant women can be magnified or neglected. (ii) Converting the food items into nutrient intakes in two instruments revealed the issues in the algorithm for extracting nutrient intakes from diet.

In my inclusion criteria, I require the energy intake calculation to be consistent, and this criterion excluded three subjects, whose total calories did not agreeing with the macronutrients intakes by the corresponding instrument. Two of them are due to the inconsistency from the web-based instrument, and one of them from the interviewer-administered instrument. The extremely weak correlation of the folate intakes from these two instruments also suggested a probable coding problem. Our work helped our collaborators identify the error in the database used to calculate the intakes in the web-based instrument. Further revision of the calibration models should be performed. More research is needed in the study of the agreement of the two instruments for the usage among pregnant women.s

Chapter 3

Trajectory Modeling of Gestational Weight by Functional Principal Component Analysis

3.1 Introduction

3.1.1 Gestational Weight Gain

Gestational weight gain (GWG) refers to the weight a pregnant woman gained during pregnancy. It is a fundamental indicator of both the health for the mother and fetus [9]. Achieving appropriate GWG is essential during pregnancy. Inadequate and excessive GWG have both been linked to various negative pregnancy outcomes as well as future health issues for both the mother and child [17, 75, 96].

What is an appropriate GWG? The recommendations for GWG were evolving through the last century. In the 1940s, healthcare providers in the United States recommended a GWG less than 20lbs, because many birth complications were believed to be associated with high birth weight [15]. In 1967, the National Research Council increased the recommended GWG, suggesting a total weight gain of 24lbs during pregnancy [18]. Since the 1970s, healthcare providers started to be more aware of the

possible adverse outcomes of low birth weight [66, 74], which resulted in a further increase in the recommended GWG. It was also noted that different weight groups tend to have different weight gains during pregnancy, e.g., slimmer women usually have more weight gain during pregnancy than overweight women does [64]. In 1990, the Institute of Medicine (IOM) in the United States balanced the health risk of too large and too small GWG and put forward GWG recommendations based on pre-pregnancy body mass index (BMI) groups (Table 5) [82]. These recommendations of GWG have been widely adopted by healthcare providers [17, 85]. The 1990 GWG guideline has been re-examined in 2009 by IOM [123], as the constitution of population BMI groups in the US changed drastically during two decades- there are more obese women and less underweight women than in the 1990s. In this new guideline (Table 6), IOM redefined the pre-pregnancy BMI groups (underweight, normal weight, overweight, and obese) according to the World Health Organization (WHO) and made additional recommendations for rates of weight gain in the 2nd and 3rd trimester. In Canada, the two versions of guidelines have been adapted by Health Canada's 1999 and 2010 GWG guidelines respectively [52, 68].

Table 5: Recommended total weight gain ranges for pregnant women, by prepregnancy body mass index (BMI), Institute of Medicine, 1990

Prepregnancy BMI (kg/m^2)	Recommended Total GWG (kg)
Low (< 19.8)	12.5-18
Normal (19.8-26.0)	11.5-16
High (26.0-29.0)	7-11.5
Obese (> 29.0)	>6.8

Research conducted after the publication of the GWG guidelines supported the recommendations. Pregnant women who gained within the recommended range are more likely to bear an infant of optimal birth weight (between 3000 and 4000 g) [84]. Whereas suboptimal, including both inadequate and excessive GWG, are shown to

Table 6: Recommended total weight gain ranges for pregnant women, by prepregnancy body mass index (BMI), Institute of Medicine, 2009

Prepregnancy BMI (kg/m^2)	Recommended total GWG (kg)	Rate of weight gain in the 2nd and 3rd trimester (kg/week)
Underweight (< 18.5)	12.5-18	0.44-0.58
Normal weight (18.5-24.9)	11.5-16	0.35-0.50
Overweight (25.0-29.9)	7-11.5	0.23-0.33
Obese (> 30.0)	5-9	0.17-0.27

have negative impacts on the pregnancy outcomes, and the health of both women and their infants. Inadequate GWG increases the risk of preterm delivery [101], and small for gestational age (SGA) ($\leq$ 3rd percentile) infants [75] or low birth weight (LBW) ($\leq$ 2500g) infants [17, 68]. SGA and LBW infants have increased risk of mortality and morbidity [75], higher first-year-death rate, and in the long run, impaired growth and development [17]. On the other hand, excessive GWG is a recognized risk factor for many adverse pregnancy outcomes such as gestational diabetes, pre-eclampsia, hypertension [56], and increases the risk of Cesarean delivery [113]. Excessive GWG also contributes to mother's obesity later in life, and post-partum weight retention [88, 96]. For the child, excessive GWG may cause the infant to be large for gestational age (LGA) or having large birth weight ($\geq$ 4000g), which are known to increase the chance of childhood obesity and type 2 diabetes in later life [25, 117].

However, despite the various unfavorable outcomes, unhealthy GWG is prevalent, especially in developed countries, with excessive GWG more common than inadequate GWG [84]. Two GWG surveys in the United States show that less than half of the pregnant women achieve recommended GWG [51, 98], and a significantly larger proportion exceeded the GWG guideline. In Canada, a 2006 Maternity Experience Survey by Statistics Canada involving more than 6000 pregnant women showed that only 36% of them gained weight within the guideline range, while 42% of them gained more than recommended [68]. As a result, a large portion of pregnant women and their infants are

at higher risk of pregnancy and delivery complications as well as later adverse health outcomes. There is an urgent need to promote healthy weight gain during pregnancy.

Although the current situation of GWG in developed countries is not optimistic, fortunately the literature indicates that GWG can be altered through prenatal interventions [84, 111]. In the clinical intervention studies on promoting healthy GWG, researchers found several interventions to be promising. Polley et al. conducted a randomized clinical trial, where pregnant women received education on weight gain, healthy eating and exercise, and were provided the graph for charting weight gain. Those exceeding weight gain goals were given more intensive intervention. The percentage of excessive weight gain decreased from 58% to 33% among women with normal prepregnancy BMI [90]. Olson et al. conducted a prospective cohort study, finding that education through newsletter and a combination of healthcare provider monitoring and self-monitoring of weight gain significantly reduced the percentage of excessive GWG compared to a historical control group [86]. For women with normal BMI the percentage reduced from 45% to 29%; and for overweight women, the percentage reduced from 72% to 44%. Rauh et al. performed a cluster-randomized controlled intervention trial, in which two individual counseling sessions were delivered to each pregnant women, focusing on diet, physical activity and weight monitoring [92]. The intervention was shown to be effective, with a lower portion exceeding IOM guidelines (38% vs 60%).

The interventions above [86, 90, 92] focused on the progress of weight gain, i.e., the individual graph/trajectory of weight throughout pregnancy. In the studies mentioned above, each pregnant women received a personalized weight gain goal, usually a weight gain chart, and the counseling and interventions were customized according to the fulfilment of this goal. However, the individual weight gain chart only depends on the prepregnancy BMI and weight, and would not be adjusted by the weight measurements during pregnancy.

Although the individual weight gain trajectory is extremely important for intervention promoting a healthy weight gain, few studies focused on the modeling of such trajectories. Research and clinical communities have put considerable efforts on studies of the numerical aspects of weight gain, e.g., total weight gain (weight just before delivery minus weight just before conception), or rate per week (weight gained over a specified period divided by the duration of that period in weeks); all definitions were adapted from [82]. Yet obviously, these aspects can be derived from the trajectory once estimated. Most existing trajectory modeling works are for the population reference rather than individuals. For example, Abrams et al. used a piece-wise linear model to estimate the weight growth [1], where the weight change in each trimester is assumed to be a linear function. Linear regression may not be a good approximation, as the weight gain rate is unlikely a constant, even within a single trimester. Xu et al. proposed conditional (longitudinal) and unconditional (cross-sectional) standard references (centiles) for pregnant women in Malawi [122]. In setting up the references, the maternal weight was modeled as a continuous function with a paramter, where percentiles are obtained by plugging the appropriate standard normal value into the parameter. Customized weight growth usually only considers categorical background information, such as prepregnancy BMI [92].

3.1.2 Weight Trajectory of Pregnant Women: Functional Data Analysis in Longitudinal Setting

In the APrON study, each subject had weight measured by trained research assistant for up to three times during pregnancy. Self-reported pre-pregnancy weight (W_0) and highest weight during pregnancy (W_H) were also recorded. To our knowledge, no attempt has been made to model the gestational weight curve as a smooth function

of gestation age. Longitudinal data analysis (LDA) studies on this topic usually focus on post-partum outcomes, such as body weight change after pregnancy [42, 59], BMI, waist circumference, and blood pressure after pregnancy [32]. Functional data analysis (FDA) in longitudinal setting has drawn lots of attention in the statistics community [44, 46, 124], but has not been used in the modeling of GWG. In previous research of LDA, methods of fitting a smooth curve profiling for the mean function are well developed [27], such as lowess curve, kernel and spline estimation. For individual trajectory modeling, early works often used mix-effect models. Fitzmaurice presented a classic parametric non-linear mixed-effect model [30], and a nonparametric model is proposed by Rice et al. [13]. Guo et al. built a functional mixed-effect model [43]. Yet none of these models is suitable for sparse data, which is the case for the APrON study.

Yao et al. proposed the functional principal component analysis (FPCA) for irregularly spaced sparse longitudinal data [124]. In this method, the functional principal component (FPC) scores are framed as conditional expectations. By selecting the first K principal components (PCs), dimension reduction is achieved. The dominant modes of variation of the sample of random trajectories around an overall mean is characterized by the first several PCs. Through this approach, with a cohort of sparse longitudinal samples, we are able to estimate the individual continuous trajectory of each subject by borrowing information from the whole cohort. If one of the weight record is relatively far from the subjects' other records, the estimated trajectory will not be skewed too much by this point. This property makes the method more robust to measurement errors and thus more suitable in estimating the individual trajectories in APrON data.

3.1.3 Our Contributions

In this chapter, I utilize the functional principal component analysis (FPCA) techniques to estimate individual-specific weight growth trajectory during pregnancy for the participants of the APrON study. The total weight gain can be estimated from this trajectory. P repregnancy BMI, the dietary intake and physical activity records, are used as adjustment covariates to model this estimated total weight gain. My contributions are:

- Individual trajectories of weight growth are estimated through FPCA techniques. The rooted mean squared error (RMSE) is significantly improved compared to the longitudinal nonlinear mixed effect model (2.1 vs 2.6). For the subjects whose weight gain patterns are different than the population mean weight growth function, the FPCA estimated trajectory shows significant advantage in terms of adaptivity.

- Individual trajectories are estimated for nutrients intakes.

- Total GWG is estimated from the individual-specific trajectory, showing good agreement with the total GWG calculated from records.

- I studied the association betweetn total GWG estimated from trajectory and the prepregnancy BMI, nutrient intakes and PA. It is found that the simple linear regression model with prepregnancy BMI as the only covariate has satisfactory R^2. The self-reported nutrient intake/physical activity may be either not reliable or not providing useful information that further explains the GWG to improve the model fit

3.2 Materials and Methods

3.2.1 Materials

Data Collection

In the APrON study, pregnant women from Edmonton and Calgary were recruited during their pregnancy. Upon a woman's recruitment, her prepregnancy weight (W_0) and due date was reported. Women recruited before 13 weeks gestation were assessed in each trimester, labeled as gestation stage A, B, C. Those recruited in 14-27 weeks gestation were assessed in gestation B and C, and so on. Each assessment included a weight measurement; a 24-hour food recall questionnaire, either web-based or interviewer-administered; a self-administered Baecke's physical activity (PA) questionnaire to evaluate their physical activities during the last month.

Subjects' GA's were calculated based on due dates. Each women was asked for a last visit at three months after delivery, during which her highest weight during pregnancy (W_H) and GA at birth were reported. The procedure is shown in Figure 2.

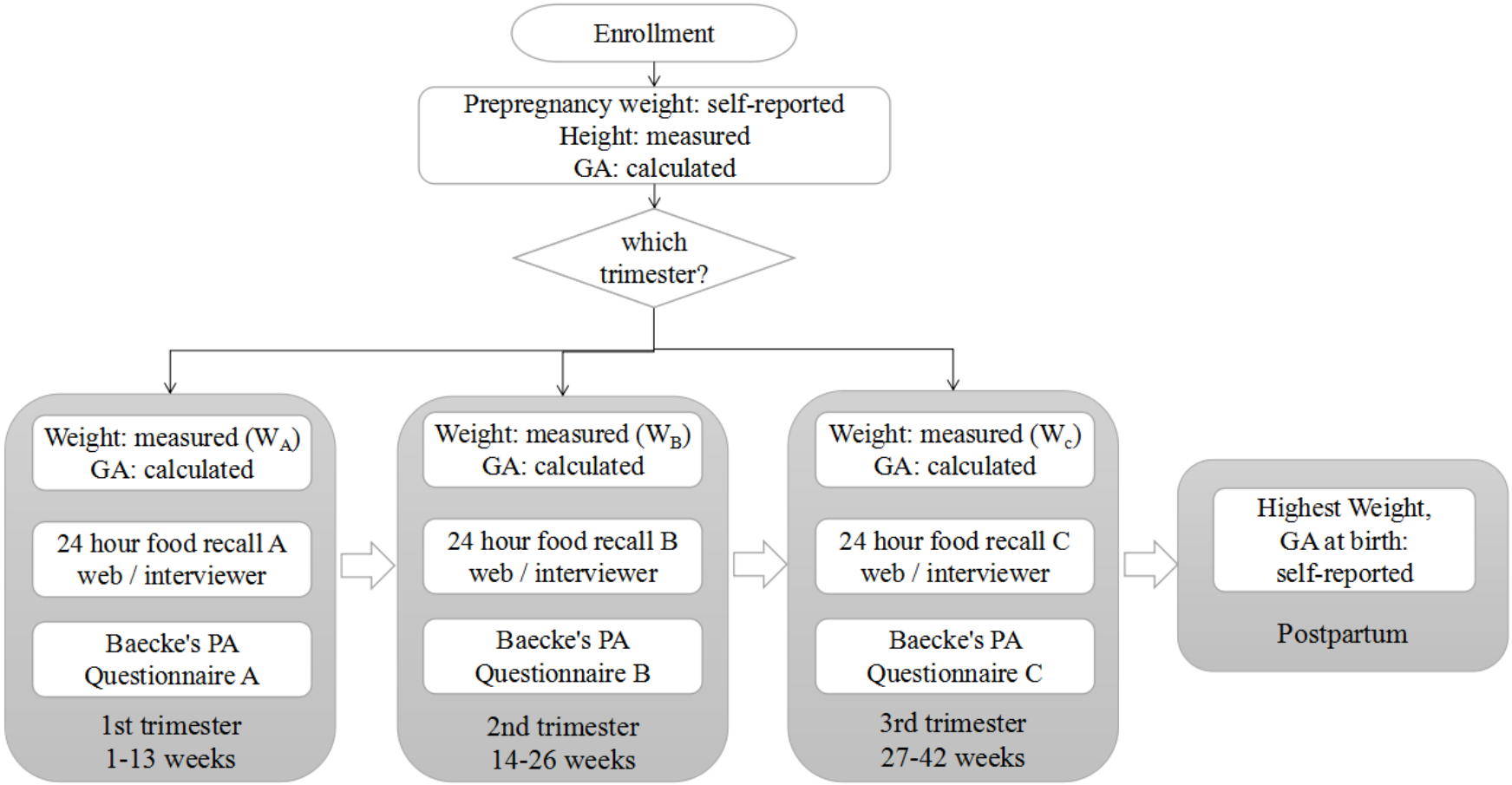


Figure 2: Flow Chart of the Data Collection

The weight measured in gestation stage A, B, C are denoted as W_A, W_B, and W_C respectively. In both the interviewer-administered and web-based food intake instruments, macro and micronutrients intake were calculated, and total calories intake were derived. Calibration study in Chapter 2 shows that direct pooling the results from the two instruments may not be feasible. In the Baecke's questionnaire, three types of activities were considered i) physical activity at work; ii) sport during leisure time; and iii) physical activity during leisure time excluding sport [8]. Activity levels of the three types were calculated as subindices-work index (WI), sport index (SI), and leisure index (LI) accordingly. The sum of the three subindices was defined as the total PA index (TI). The Baecke's questionnaire has been validated in various populations, including adult men [31, 89], adult women with hip disorders [87], and showed good validity and reliability. In summary, each subject has a maximum of 5 body weight data points for her trajectory, with GA varying from 0 to 42 in weeks: the self reported W_0, the measured weight W_A, W_B, W_C; and the self-reported W_H. One or more of these data points could be missing due to missing visit(s). We correspond W_0 to $t = 0$, and W_H is associated with the GA at birth.

Inclusion Criteria

Subjects who have a full-term(GA at birth $\geq$ 37 weeks), singleton live birth are considered in the analysis. For statistical modeling, we require at least one valid record during pregnancy, i.e., she has at least one record of physical activities or food intake and at least one weight record with corresponding GA . Due to the large amount of data collected, various issues were detected and data cleaning were performed. For example, the subjects who have apparently wrong GA at birth information, namely, the GA at birth is less than the GA of time point C, are excluded. For the food intake data, instances with total calorie intakes greater than 3500 kcals or less than 600 kcals

were excluded, as recommended by Csizmadi et al. [24]. In addition, women who have completed more than one versions of 24-hour food recall questionnaires are excluded, in order to model nutrient intake trajectories from data collected by the two instruments separately.

3.2.2 Method

Model set up and estimation

We consider $X_i(t),\ i = 1, ..., n$, which are n realizations of a smooth random function $X(t)$, with unknown mean $EX(t) = \mu(t)$ and smooth covariance function $\text{cov}(X(s), X(t)) = G(s, t)$. The j-th observation of $X_i(t)$ made at a random time T_{ij}, denoted as Y_{ij}, is subject to uncorrelated measurement errors with mean 0 and constant variance σ^2. Then the model is

$$Y_{ij} = X_i(T_{ij}) + \epsilon_{ij} \tag{3.2.1}$$

where $j = 1, ..., N_i$. N_i is the number of observations of X_i, and ϵ_{ij} is the measurement error. In our case, X can be the trajectory of body weight, intake of certain nutrient, or any PA index, and $X_i(t)$ is then the weight/nutrient intake/PA index of the i-th subject at time t. The time t is in a closed time interval $\mathcal{T} = [0, 42]$ (weeks).

In FPCA, the i-th realization can be expressed as

$$X_i(T_{ij}) = \mu(T_{ij}) + \sum_{k=1}^{\infty} \xi_{ik}\phi_k(T_{ij}) + \epsilon_{ij}, \tag{3.2.2}$$

where ϕ_k's are the PC functions, assumed to be smooth.

The estimation of the trajectory is implemented in five steps: (1) Estimation of the mean function $\mu(t)$; (2) Smooth estimation of the covariance surface $G(s, t)$ and measurement error σ^2; (3) Estimation of the eigenfunctions (PCs) and eigenvalues; (4)

Estimation of the FPC scores, ξ_{ik}, through conditional expectation; (5) Choice of the number of PCs to be selected. The details are given as following:

(1) The mean function μ is estimated based on all the data points from all individuals, where local linear smoothers proposed by Fan et al. [29] are employed. Let $\kappa_1(u)$ be a univariate compact supported kernel function of order (ν, ω) with $0 \leqslant \nu < \omega$. That is, $\kappa_1(u)$ satisfies

$$\int u^l \kappa_1(u) du = \begin{cases} 0 & \text{if } 0 \leqslant l < \omega, l \neq \nu, \\ (-1)^{|\nu|} |\nu|! & \text{if } l = \nu, \\ \text{any nonzero value} & \text{if } l = \omega. \end{cases}$$

Then we define the local linear scatterplot smoother for $\mu(t)$ by minimizing

$$\sum_{i=1}^{n} \sum_{j=1}^{N_i} \kappa_1 \left(\frac{T_{ij} - t}{h_\mu} \right) [Y_{ij} - \beta_0(t - T_{ij}) - \beta_1(t - T_{ij})]^2 \tag{3.2.3}$$

with respect to β_0 and β_1. The parameter h_μ (kernel radius) determines the level of localization. The estimate of $\mu(t)$ is then $\hat{\mu}(t) = \hat{\beta}_0(t)$.

(2) Let $G_i(T_{ij}, T_{il}) = (Y_{ij} - \hat{\mu}(T_{ij}))(Y_{il} - \hat{\mu}(T_{il}))$ be the "raw" covariances. Then we have $E[G_i(T_{ij}, T_{il}) | T_{ij}, T_{il}] \approx \text{cov}(X(T_{ij}), X(T_{il})) + \sigma^2 \delta_{jl}$, where δ_{jl} is the Kronecker notation, i.e., $\delta_{jl} = 0$ for $j \neq l$ and $\delta_{jl} = 1$ for $j = l$. So the diagonal of the raw covariance matrix should be removed, and only those $G_i(T_{ij}, T_{il})$ with $j \neq l$ should be included as input data for the covariance surface smoothing step.

Then the surface smoothing also follows the local linear smoother in [29]. Let $\kappa_2(u, v)$ be a bivariate compact supported kernel function of order (ν, ω), where ν is a

multi-index (ν_1, ν_2) satisfying $0 \leqslant \nu_1 + \nu_2 \leqslant \omega$. The order index means

$$\int u^{l_1} v^{l_2} \kappa_2(u,v) du dv = \begin{cases} 0 & \text{if } 0 \leqslant l_1 + l_2 < \omega, l_1 \neq \nu_1, l_2 \neq \nu_2, \\ (-1)^{|\nu|}|\nu|! & \text{if } l_1 = \nu_1, l_2 = \nu_2, \\ \text{any nonzero value} & \text{if } l_1 + l_2 = \omega, \end{cases}$$

where $|\nu| = \nu_1 + \nu_2$. The local linear surface smoother for $G(s,t)$ is obtained by minimizing

$$\sum_{i=1}^{n} \sum_{1 \leqslant j \neq l \leqslant N_i} \kappa_2 \left(\frac{T_{ij} - s}{h_G} \frac{T_{ij} - t}{h_G} \right) \times [G_i(T_{ij}, T_{il}) - (\beta_0 + \beta_1(s - T_{ij}) + \beta_2(t - T_{il}))]^2 \quad (3.2.4)$$

with respect to β_0, β_1 and β_2. The estimate of $G(s,t)$ is then $\hat{G}(s,t) = \hat{\beta}_0(s,t)$.

Next we estimate σ^2. Let $\widehat{G}(s,t)$ be a smooth surface estimate of $G(s,t)$. As the covariance is maximal along diagonal, we fit a local linear component along the diagonal, and a quadratic component along the direction perpendicular to the diagonal. We denote the diagonal of the resulting surface estimate by $\widetilde{G}(t)$. A local linear smoother $\hat{V}(t)$ focusing on diagonal values $\{G(t,t) + \sigma^2\}$ is obtained as in step (1), by using $G(T_{ij}, T_{ij})$ as input. To minimize boundary effects, take only $\mathcal{T}_1 = [10.5, 31.5]$, the middle half of $\mathcal{T}$, we estimate σ^2 as

$$\hat{\sigma}^2 = \frac{2}{|\mathcal{T}|} \int_{\mathcal{T}_1} [\hat{V}(t) - \tilde{G}(t)] dt. \quad (3.2.5)$$

(3) Estimating the eigenfunctions (PCs) and eigenvalues is essentially solving for the eigenequations

$$\int_{\mathcal{T}} \widehat{G}(s,t) \hat{\phi}_k(s) ds = \hat{\lambda}_k \hat{\phi}_k(t)$$

with respect to $\hat{\phi}_k$'s and $\hat{\lambda}_k$'s, where the estimated eigenfunction $\hat{\phi}_k$'s are subject to a unit L^2 norm and are perpendicular to each other. These eigenfunctions are estimated

by discretizing the smoothed covariance as in [93]. The estimate of the eigenfunction corresponding to the largest eigenvalue, $\hat{\mathbf{u}} = \hat{\phi}_1(\mathbf{t})$ ($\mathbf{t}$ is a grid of t thus $\hat{\mathbf{u}}$ is a discretization of $\hat{\phi}_1$) is obtained by

$$\arg\max \mathbf{u}^T \widehat{\mathbf{G}} \mathbf{u}$$

$$\text{subject to } \|\mathbf{u}\| = 1 \text{ and } \mathbf{u}^T \mathbf{D} \mathbf{u} \leqslant \beta,$$

where $\widehat{\mathbf{G}}$ is the discretization of $\widehat{G}(s,t)$, β is a smoothing parameter and $\mathbf{D}$ is a roughening matrix, e.g., $\mathbf{D} = F^T F$ where F is a second-difference operator.

(4) Due to the sparse nature of our data, the traditional estimation of the FPC scores $\xi_{i,k} = \int (X_i(t) - \mu(t))\phi_k(t)dt$ through numerical integration does not work well, as the integration of a smooth function can not be estimated well by Riemann sums when there are only a few data points. Yao et al. proposed the FPCA through conditional expectation [124], which is more suitable for sparse data. Denote the i-th subject with observations $\tilde{Y} = (Y_{i1}, ..., Y_{iN_i}) = (X_i(T_{i1}) + \epsilon_{i1}, ..., X_i(T_{iN_i}) + \epsilon_{iN_i})$, where T_{ij} are the corresponding observation times and ϵ_{ij} are the measurement errors, the best prediction of the FPC scores for the subject, is the conditional expectation $E(\xi_{ik}|\tilde{Y}_i)$. If we further assume that ξ_{ik} and ϵ_{ij} are jointly normal, then Mardia et al. claimed that the conditional expectation is given as

$$\tilde{\xi}_{ik} = E(\xi_{ik}|\tilde{Y}_i) = \lambda_k \phi_{ik}^T \Sigma_{Y_i}^{-1} (\tilde{Y}_i - \mu_i), \tag{3.2.6}$$

where $\phi_{ik} = (\phi_k(T_{i1}), ..., \phi_k(T_{iN_i}))$, $\mu_i = (\mu(T_{i1}), ..., \mu(T_{iN_i}))$, and $\Sigma_{Y_i} = \text{cov}(\tilde{Y}_i, \tilde{Y}_i) = \text{cov}(\tilde{X}_i, \tilde{X}_i) + \sigma^2 I_{N_i}$, is a $N_i \times N_i$ matrix, the (j,l) entry of which is $G(T_{ij}, T_{il}) + \sigma^2 \delta_{jl}$, with δ_{jl} the Kronecker [72].

Plugging in the quantities on the right hand side in (3.2.6) by their estimates, we

get the "plugging in" estimator of the FPC score

$$\hat{\xi}_{ik} = \hat{E}(\xi_{ik}|\tilde{Y}_i) = \hat{\lambda}_k \hat{\phi}_{ik}^T \hat{\Sigma}_{Y_i}^{-1}(\tilde{Y}_i - \hat{\mu}_i), \tag{3.2.7}$$

where $\hat{\mu}_i$ is obtained in step (1), $(\hat{\Sigma}_{Y_i})_{j,l} = \widehat{G}(T_{ij}, T_{il}) + \hat{\sigma}^2\delta_{jl}$ estimated in step (2), and $\hat{\phi}_{ik} = (\phi_k(T_{i1}), ..., \phi_k(T_{iN_i}))$, λ_k obtained from step (3). A MATLAB-based package, Principal Analysis by Conditional Expectation (PACE), was developed by Yao et al. for the implementation of various Functional Data Analysis (FDA) [124]. We adapted this package for our data analysis .

Prediction of Individual Trajectories

Upon obtaining the PCs and FPC scores, the first several PCs explained the largest fraction of total variations in all the trajectories, and thus representing the dominant modes of variation. For each individual, the trajectory can be well approximated by a linear combination of these K smooth functions and is thus smooth. Dimensionality is then reduced from ∞ to K, or equivalently, we project any function on to the space spanned by the first K PCs. There are several different criteria to determine the choice of K. Popular choices include the Akaike Information Criterion (AIC) and Fraction of Variance Explained (FVE). For a model with the first K PCs selected, the AIC value is defined as AIC $= K - l$, where $l = \ln(L)$ is the maximum of the log-likelihood of the model. l can be estimated by summing the contributions from all subjects, conditional on the estimated FPC scores (3.2.6), assuming that ξ_{ik} and ϵ_{ij} are jointly normal:

$$\hat{l} = \sum_{i=1}^{n} \left\{ -\frac{N_1}{2}\lg(2\pi) - \frac{N_1}{2}\lg\hat{\sigma}^2 - \frac{1}{2\hat{\sigma}^2}\left(\tilde{Y}_i - \hat{\mu}_i - \sum_{k=1}^{K}\hat{\xi}_{ik}\hat{\phi}_{ik}\right)^T \left(\tilde{Y}_i - \hat{\mu}_i - \sum_{k=1}^{K}\hat{\xi}_{ik}\hat{\phi}_{ik}\right) \right\}.$$

Then the best model is of the smallest AIC value. The AIC method balances large likelihood (larger $\hat{l}$) and less model complexity (smaller K), however, is usually computationally expensive. The FVE method chooses the first K PCs that explain a pre-defined fraction (threshold) of total variance. So by choosing a considerably large threshold, the model chosen explains most of the variability. As the covariance is estimated in the FPCA process, the FVE method is more efficient. We usually employ the FVE method, with a threshold of 0.99 of the variance explained. Namely, K is the least positive number that the first K eigenfunctions explain more than 99% of the total variance. Once K is chosen we approximate the trajectory of the i-th subject by

$$\hat{X}_i^K(t) = \hat{\mu}(t) + \sum_{k=1}^{K} \hat{\xi}_{ik}\hat{\phi}_k(t), \tag{3.2.8}$$

where $\hat{\xi}_{ik}$'s are the individual-specific FPC scores. In (3.2.8), the terms $\hat{\mu}(t)$ and $\hat{\phi}_k(t)$ borrows information from the entire data set, while $\hat{\xi}_{ik} = \hat{E}(\xi_{ik}|\tilde{Y}_i) = \hat{\lambda}_k\hat{\phi}_{ik}^T\hat{\Sigma}_{Y_i}^{-1}(\tilde{Y}_i - \hat{\mu}_i)$ is driven by $\tilde{Y}_i$, the observations of the specific subject. We will show that this approach handles the longitudinal sparse data set extremely well when the subjects have a general common trend. As discussed in Section 3.2, for each subject in our study, we have up to 5 records of their weight and up to 3 points of nutrient intake and PA measurements, which are all longitudinal sparse data. Take weight as an example, $n > 1000$ subjects are included in the modeling, each has weights measured from $t = 0$ (W_0) to $t \geq 37$ (W_H at GA at birth). The mean function and covariance surface can first be estimated from the pooled data. The PCs are thus estimated. For each subject, the FPC scores can then be estimated, conditional on the up to 5 observations of her weight.

Confidence Bands of Individual Trajectories

Let K be the number chosen by FVE approach, and the vector of FPC scores $\xi_{K,i} = (\xi_{i1}, ..., \xi_{iK})^T, \tilde{\xi}_{K,i} = (\tilde{\xi}_{i1}, ..., \tilde{\xi}_{iK})^T$. Recall that $\tilde{\xi}_{ik} = E(\xi_{ik}|\tilde{Y}_i) = \lambda_k \phi_{ik}^T \Sigma_{Y_i}^{-1}(\tilde{Y}_i - \mu_i)$, then the covariance matrix of $\tilde{\xi}_{K,i}$ can be written as $H\Sigma_{Y_i}^{-1}H^T$, where H is the covariance matrix between $\xi_{K,i}$ and $\widetilde{Y}_i$. Note that for a fixed sample, λ_k, ϕ_{ik} and $(\hat{\Sigma}_{Y_i})_{j,l} = \widehat{G}(T_{ij}, T_{il}) + \hat{\sigma}^2\delta_{jl}$ are independent with $\widetilde{Y}_i$, so $\tilde{\xi}_{K,i}$ is a linear function of $\widetilde{Y}_i$. H can be rewritten as

$$H = (\lambda_1\phi_{i1}, \cdots, \lambda_K\phi_{iK})^T.$$

The estimation error of $\tilde{\xi}_{K,i}$ can be assessed by $\text{var}(\tilde{\xi}_{K,i} - \xi_{K,i})$. The conditional expectation $E(\xi_{K,i}|\widetilde{Y}_i)$ is the projection of $\xi_{K,i}$ on the space span $\left\{\tilde{Y}_i\right\}$, thus $E(\tilde{\xi}_{K,i}\xi_{K,i}^T) = E(\tilde{\xi}_{K,i}\tilde{\xi}_{K,i}^T)$, and

$$var(\tilde{\xi}_{K,i} - \xi_{K,i}) = \text{var}(\xi_{K,i}) - \text{var}(\tilde{\xi}_{K,i}) = \Omega_K,$$

where $\Omega_K = \Lambda - H\Sigma_{Y_i}^{-1}H^T = \text{diag}(\lambda_1, ..., \lambda_K) - H\Sigma_{Y_i}^{-1}H^T$. Under Gaussian assumptions, we have $(\tilde{\xi}_{K,i} - \xi_{K,i}) \sim \mathcal{N}(0, \Omega_K)$. From (3.2.8), the individual trajectory is estimated as $\hat{X}_i^K(t) = \hat{\mu}(t) + \sum_{k=1}^K \hat{\xi}_{ik}\hat{\phi}_k(t) = \hat{\mu}(t) + \hat{\phi}_{K,t}^T\xi_{K,i} = (\xi_{i1}, ..., \xi_{iK})^T$ where $\hat{\phi}_{K,t} = (\hat{\phi}_1(t), ..., \hat{\phi}_K(t))^T$. It was showed that $\hat{X}_i^K(t) - X(t)$ approximately follows normal distribution $\mathcal{N}(0, \hat{\phi}_{K,t}^T\hat{\Omega}_K\hat{\phi}_{K,t})$.

Therefore, the $(1-\alpha)$ asymptotic point-wise confidence intervals for individual trajectories are given as

$$\widehat{X}_i^K(t) \pm \Phi^{-1}(1-\alpha/2)\sqrt{\hat{\phi}_{K,t}^T\widehat{\Omega}_K\hat{\phi}_{K,t}},$$

where $\widehat{X}_i^K(t) = \hat{\mu}(t) + \hat{\phi}_{K,t}^T\hat{\xi}_{K,i}$, $\widehat{\Omega}_K = \widehat{\Lambda} - \widehat{H}\widehat{\Sigma}_{Y_i}^{-1}\widehat{H}^T$, with $\widehat{\Lambda} = \text{diag}(\hat{\lambda}_1, ..., \hat{\lambda}_K)$ and $\widehat{H} = (\hat{\lambda}_1\hat{\phi}_{i1}, \cdots, \hat{\lambda}_K\hat{\phi}_{iK})^T$.

The $(1-\alpha)$ asymptotic simultaneous confidence intervals for individual trajectories

are given as

$$\widehat{X}_i^K(t) \pm \sqrt{\chi^2_{K,1-\alpha}\hat{\phi}_{K,t}^T \widehat{\Omega}_K \hat{\phi}_{K,t}}.$$

Total GWG Prediction

The total GWG can be estimated from the estimated weight growth trajectory. Then its association with the background information (prepregnancy BMI), nutrient intakes, and PA can be studied by multivariate linear regression models, where the total GWG is the response to be predicted, and the other information are covariates.

For a multivariate linear regression model, both the response and predictors should be scalars. In APrON study, the nutrient intakes and PA indices were collected longitudinally. To fit a multivariate regression model, we can transform the longitudinal data into scalar ones. A natural idea is to use the average of all the measurements for a covariate. For example,

$$G = \beta_0 + \beta_1 \mathrm{BMI} + \beta_2 \mathrm{T} + \beta_3 \mathrm{TI} + \epsilon, \tag{3.2.9}$$

where G is the total GWG, BMI is the prepregnancy BMI, T is the average of all the total calories measurements obtained from the prenatal 24-hour food recalls, and TI is the average of all the total PA index obtained from the prenatal Baecke's questionnaires. The covariates can also be the macronutrients (fat, protein, carbohydrate) intake and the subindices of PA, where the average is similarly obtained. The only issue is to avoid taking total calories and all the three macronutrients together, or the total PA index and all the three subindices as covariates, which will cause colinearity.

However, note that the trajectories of these longitudinal covariates can be estimated. The estimated trajectories can also be utilized in the modeling. Müller et al.

proposed a functional additive model (FAM) for such scenarios. Let $X(t)$ be the trajectory of certain covariates with sparse longitudinal observations. Then the trajectory of a subject $X_i(t)$ is estimated by FPCA

$$\hat{X}_i(t) = \hat{\mu}(t) + \sum_{k=1}^{K} \hat{\xi}_{ik}\hat{\phi}_k(t),$$

where $\hat{\mu}(t)$ is the estimated mean function of the entire sample; $\hat{\xi}_{ik}$'s are the estimated FPC scores; $\hat{\phi}_k(t)$'s are the PC functions, and K is the number of PCs been selected.

In the functional regression model $Y_i = \int X_i(t)\beta(t)dt + \epsilon$, if the estimated trajectory of $X(t)$ is used, it can be re-written as

$$\begin{aligned} Y_i &\approx \int (\hat{\mu}(t) + \sum_{k=1}^{K} \hat{\xi}_{ik}\hat{\phi}_k(t)\beta(t)) + \epsilon \\ &= \int \hat{\mu}(t) + \sum_{k=1}^{K} \hat{\xi}_{ik} \int \beta(t)\hat{\phi}_k(t) + \epsilon \\ &= \int \hat{\mu}(t) + \sum_{k=1}^{K} \hat{\xi}_{ik}\beta_k + \epsilon \end{aligned} \tag{3.2.10}$$

where β_k is the coefficient when projecting the functional coefficient $\beta(t)$ onto the functional space spanned by $\hat{\phi}_k(t)$. Now (3.2.10) converts the functional regression to a multivariate linear regression, where $\hat{\xi}_{ik}$'s are the covariates, and β_k's are the coefficient to be estimated. Once we get the estimated $\hat{\beta}_k$, the functional coefficient $\beta(t)$ can be (partially) recovered by

$$\beta(t) \approx \sum_{i=1}^{K} \hat{\beta}_k\hat{\phi}_k(t).$$

This additive model can be easily extended to multiple functional covariates and a mixture of scalar and functional covariates, by adding additional terms on the right

hand side of (3.2.10). For example, if the model includes the following predictors: prepregnancy BMI's $\mathbf{x} = [x_1, x_2, ..., x_n]$, the estimated trajectories of total calories $\tilde{\mathbf{T}} = [\hat{T}_1(t), \hat{T}_2(t), ..., \hat{T}_n(t)]$ and the SI's $\tilde{\mathbf{S}} = [\hat{S}_1(t), \hat{S}_2(t), ..., \hat{S}_n(t)]$, the model can be written as

$$\begin{aligned} Y &= \beta_0 + \mathbf{x}\beta_1 + \int T(t)\beta_2(t)dt + \int\int S(t)\beta_3(t)dt + \epsilon \\ &\approx \beta_0 + \mathbf{x}\beta_1 + \int \hat{\mu_T}(t)dt + \sum_{k=1}^{K_T} \beta_{2,k}\hat{\xi}_{2,k} + \int \hat{\mu_S}(t)dt \sum_{k=1}^{K_S} \beta_{3k}\hat{\xi}_{3,k}), \quad (3.2.11) \end{aligned}$$

where K_T and K_S are the number of PCs selected by the FPCA of $T(t)$ and $S(t)$; $\beta_{2,1}, ..., \beta 2, K_T, \beta_{3,1}, ..., \beta_{3,K_S}$ are the coefficients to be estimated. The functional coefficients $\beta_2(t)$ and $\beta_3(t)$ can be then approximated as:

$$\beta_2(t) \approx \sum_{i=1}^{K_T} \hat{\beta}_{2,k}\hat{\phi}_{T,k}(t), \quad \beta_3(t) \approx \sum_{i=1}^{K_S} \hat{\beta}_{3,k}\hat{\phi}_{S,k}(t),$$

where $\hat{\phi}_{T,k}(t), k = 1, ..., K_T$ and $\hat{\phi}_{S,k}(t), k = 1, ..., K_S$ are the PCs selected by the FPCA of $T(t)$ and $S(t)$.

3.3 Results

3.3.1 Data Summary

The APrON study recruited 2191 pregnant women. Each subject completed up to three prenatal 24-hour food recall questionnaires, corresponding to the three trimesters (denoted as time point A, B, and C) respectively, as shown in Table 7. In total, 4453 prenatal 24-hour food recall records, including both interviewer-administered and the web-based version, have been collected. Among the 2191 subjects, 21 subjects did

Table 7: Completion Pattern of the 24-hour Food Recall

Interview(s) Per Subject	0	1	2	3
Number of Subjects (n)	21	309	1437	423
Interview Pattern		A: 59 B: 227 C: 23	AB: 39 AC: 21 BC: 1377	ABC

not complete any questionnaires. There were 309 subjects who completed one questionnaire, including 59 subjects completing the questionnaire at A (1st trimester), 227 subjects completing at B (2nd trimester), 23 subjects completing at C (3rd trimester). The majority participants, 1437, completed two questionnaires, mostly at B and C (1377 subjects); some missed C (39 subjects) and some missed B (21 subjects). There are 423 subjects completed all three prenatal questionnaires.

We have 1540 subjects with consistent weight and food intake/PA records satisfying the given inclusion criteria in Section 3, i.e., at least one valid prenatal record, valid W_0 and prepregnancy BMI, GA at birth greater than 37 weeks and greater than GA at time point C. Among these 1540 subjects, 725 completed the interviewer-administered 24-hour food recall questionnaires, 797 completed the web-based 24-hour food recall questionnaires, 6 completed both versions of questionnaires, and 12 had no 24-hour food recall, but only Baecke's PA questionnaire completed.

The inclusion procedure is shown in Figure 3.

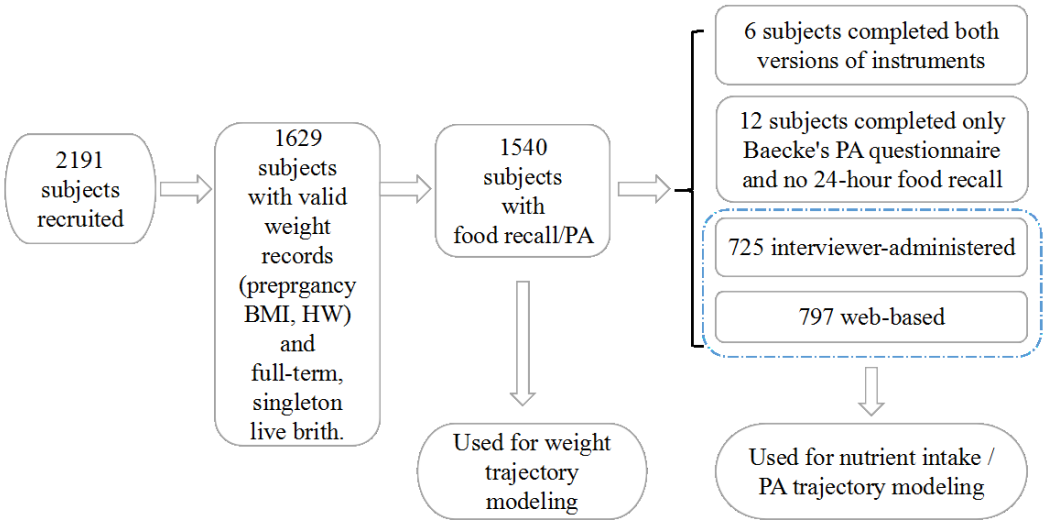


Figure 3: Flow chart of data inclusion

3.3.2 Modeling the Trajectories of Weight and Nutrient Intakes

Modeling of Weight Trajectories

A discrepancy between the nutrient intakes measured in the two 24-hour food recall instruments was observed; see Chapter 2. Therefore, we model the web-based and interviewer-administered nutrient intakes separately. On the contrary, women's weights were measured consistently using a standard equipment, thus weight trajectories were modeled using the entire analysis cohort of 1540 subjects, as seen in Figure 3.

For a sparse longitudinal data set, the distribution of the time points when the observations were made can be revealed from the assembled pairs (T_{ij}, T_{ik}). An assembled pair represents two time points T_{ij} and T_{ik} for the i-th subject. The pair is ordered, thus each pair is counted twice, (T_{ij}, T_{ik}) and (T_{ik}, T_{ij}) respectively, which are

symmetric about the 45° slanted line $T_{ij} = T_{ik}$. For example, the assembled pairs of all the 24-hour food recall records of the 2169 subjects are shown in Figure 4. Considering the part below the slanted line only, two clusters of data are present: the smaller one on the lower left corner corresponds to those subjects who have records in the 1st and 2nd trimester. According to Table 7, this cluster includes the 423 subjects who completed the questionnaire in all three trimesters (ABC) and the 39 subjects who completed the questionnaire only in the first two trimesters (AB). The larger one on the right corresponds to those subjects with records in the 3rd trimester and either the 1st (lower half with $0 < T_{ik} < 13$) or 2nd trimester (upper half with $13 < T_{ik} < 27$). The densest part is the upper half, which represents the subjects who have completed the questionnaires in both the 2nd and 3rd trimester. This part includes the 423 subjects who completed the questionnaires at ABC and the 1377 subjects who completed the questionnaires at BC. Subjects will not complete two questionnaires in a single trimester, so there are three empty areas along the 45° line $T_{ij} = T_{ik}$: $(0, 13) \times (0, 13)$, $(16, 29) \times (16, 29)$ and $(30, 37) \times (30, 37)$. After gestation of 37 weeks, women have much higher chance to be in labor than before, so the subjects will probably not schedule to complete the questionnaire after 37 weeks. Thus points with $T_{ij} > 37$ or $T_{jk} > 37$ are relatively rare.

The assembled pairs of all the weight records of the 1540 subjects are shown in Figure 5. Note that by adding W_H to GA at birth, a larger cluster occurs with $T_{ij} \geq 37$. The points in this cluster represent the subjects who have valid GA at birth, and at least one weight record during pregnancy. The self-reported prepregnancy weight is added to $t = 0$, resulting in the horizontal line $T_{ik} = 0$. The coverage of this line is approximately 5 to 42 weeks, indicating that at almost any time point after 5 weeks, we have a weight record for certain subject.

Using the method described in Section 3.2.2, we estimated a smoothed mean weight

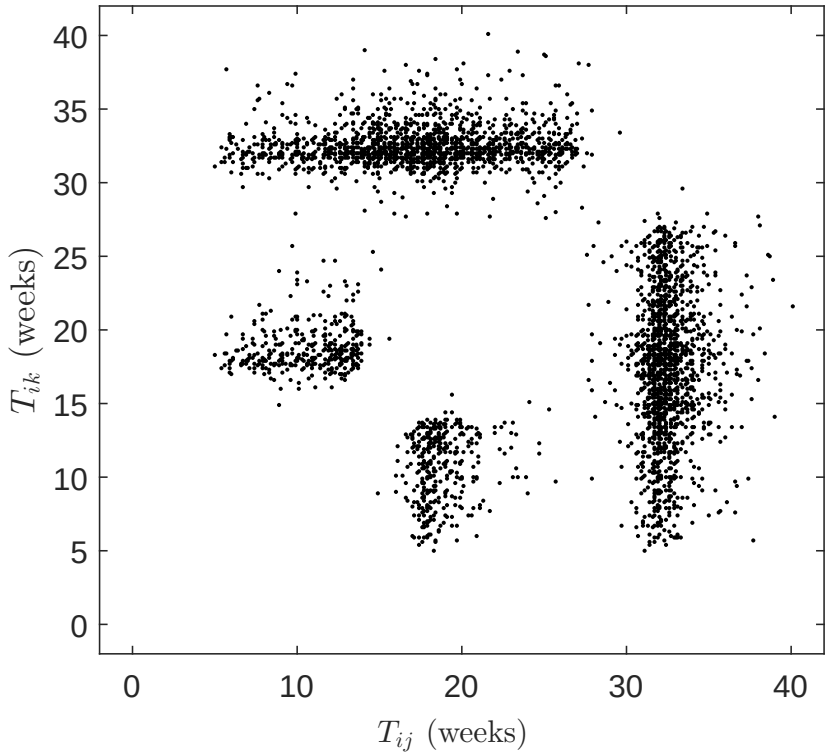


Figure 4: Assembled pairs (T_{ij}, T_{ik}) of all the 24-hour food recall records of the 2169 subjects

trajectory all subjects as in Figure 6. The smooth estimate of the variance function for weight data is shown in Figure 7. It is decreasing before 25 weeks, suggesting that the weights of subjects vary the most at the beginning of pregnancy. A fluctuation occurs around 27-35 weeks. It is easy to see from Figure 6 that there are only a few data points around $t = 29$. So the estimate of variance around $t = 29$ is not as reliable as elsewhere with more data points. The fluctuation is probably due to the lack of data. After 35 weeks, the variance decreases more rapidly, which indicates a smaller variance among the W_H records of all subjects.

The smooth estimate of the correlation surface of weight data is shown in Figure 8. The entire surface is above 0.7, indicating that weight records of the same subject are highly correlated at all times. However, the correlation between the very early

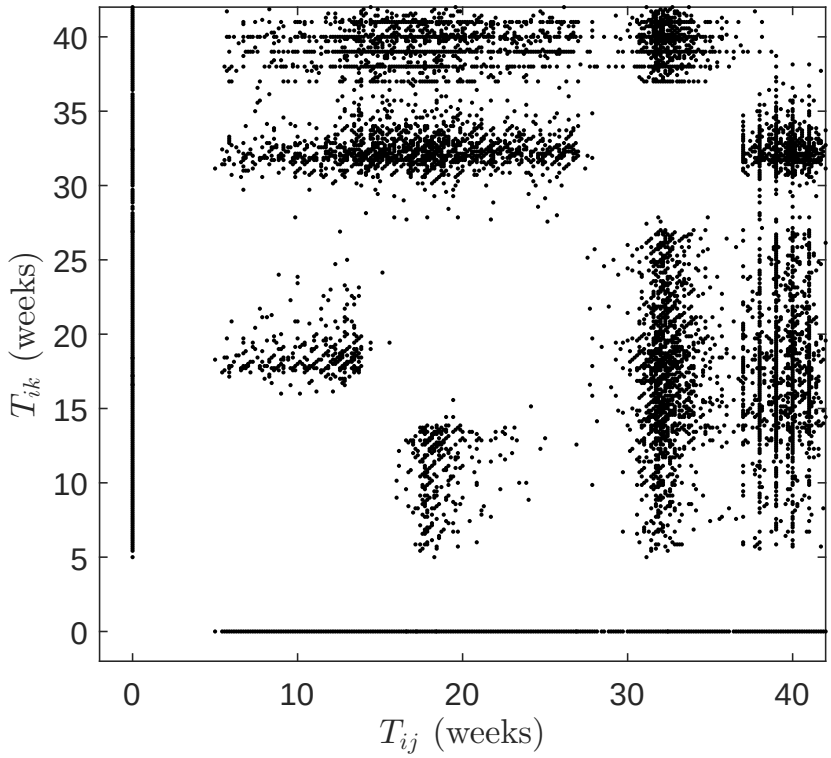
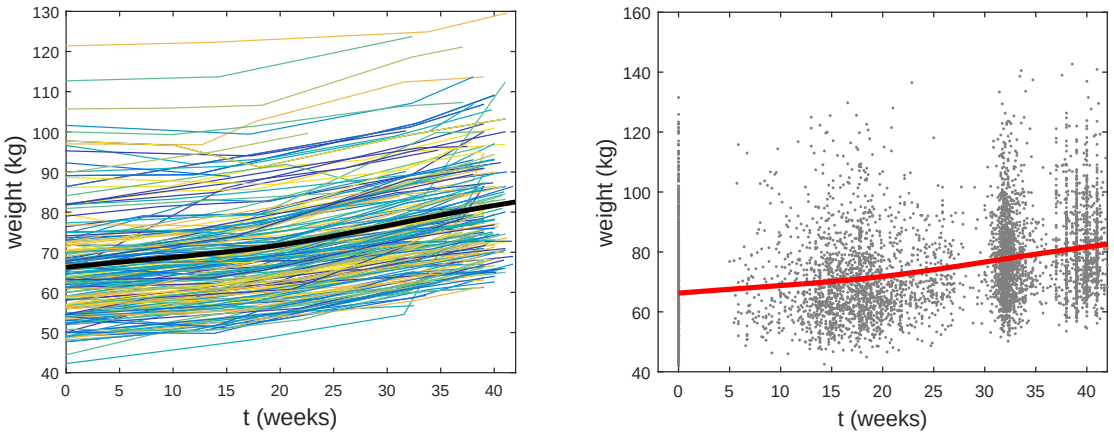


Figure 5: Assembled pairs (T_{ij}, T_{ik}) of all the weight records of the 1540 subjects



(a) (b)

Figure 6: (a) Observed individual weight trajectories of randomly selected 100 subjects, overlaid with the smooth estimate of the mean function and (b) All the weight records overlaid with the smooth estimate of the mean function

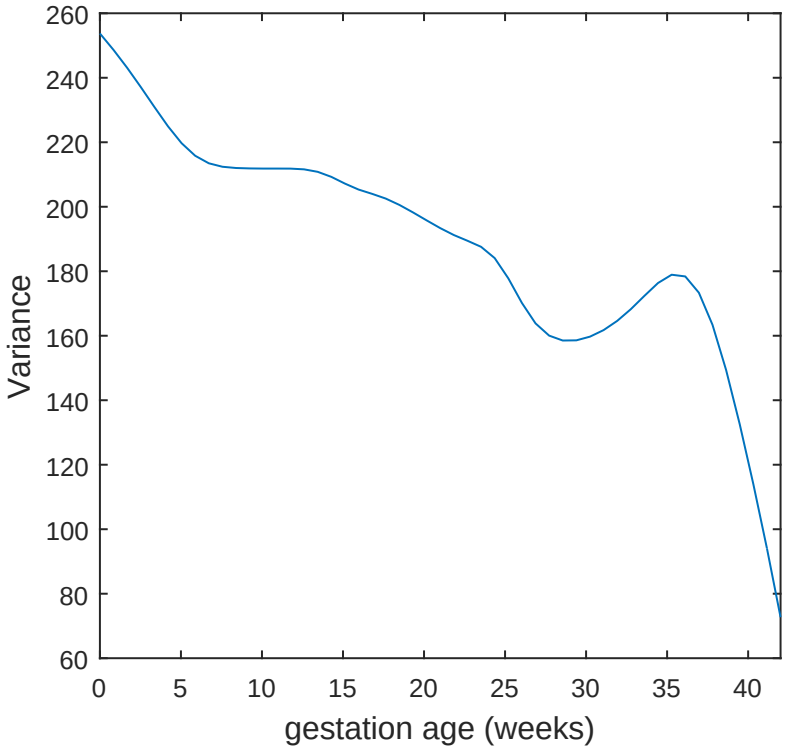


Figure 7: Smooth estimate of the variance function of weight data. Variance is non-stationary, following a general decreasing trend except for the fluctuation around 30 weeks.

weight and weight measured at later times decreases drastically in the first 20 weeks. Correlation between the pre-pregnancy weight and the weight after 20 weeks is weaker than weight before 20 weeks. It suggests that a subject's weight begins to develop some new pattern after gestation 20 weeks. This lowered correlation does not increase much in later pregnancy, except around 30 weeks. Similarly as the variance in Figure 7, the fluctuation may be due to a random error in the estimation of covariance. However, the dependence within later times in pregnancy is strong after 10 weeks, which can be interpreted as the weight of a pregnant woman at any two time points in the 2nd and 3rd trimester are highly correlated with each other.

The scree plot is shown in Figure 9(a). The first three PCs account for 95.7%,

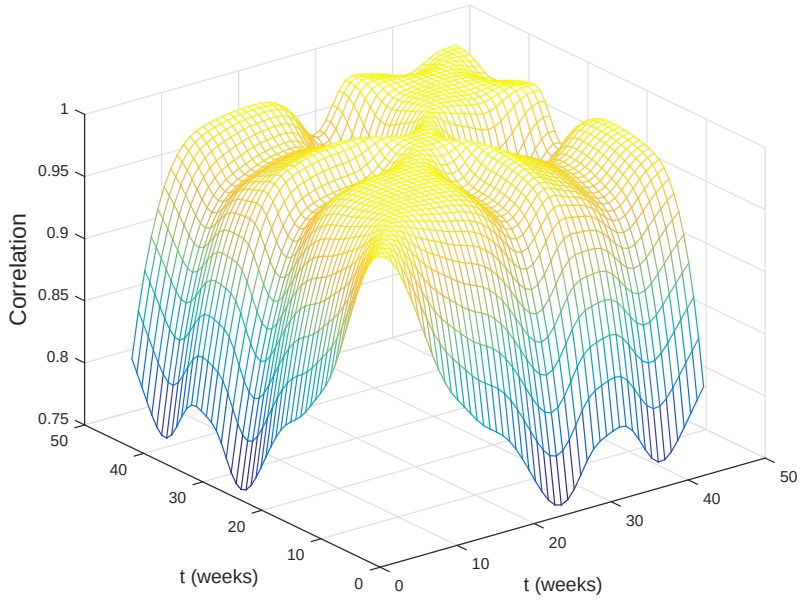


Figure 8: Smooth estimate of the correlation surface. Correlation surface is always above 0.7, indicating high correlation at all times. However, the correlation between the weight at very early times and later times decreases rapidly in the first 20 weeks.

2.8%, and 1.1% of the total variation respectively, totally more than 99%. Thus they are selected in the estimation. Their graphs are shown in Figure 9(b-d). The first PC is flat during the first trimester, and decreases rapidly in the second and third trimester, similar to the smooth estimate of the variance. The second PC increases most rapidly in the first trimester, corresponding to a contrast between prepregnancy weight or very early times and the second/third trimester. The third corresponds to a contrast between the second trimester and the third trimester. So these PCs correspond to weight gain during different times. The first explains most of the variation, suggesting that the majority of weight is gained during the last two trimesters.

The performance of the FPCA-based individual trajectory is compared with the traditional parametric mixed effect model. As seen from Figure 6, the weight growth curve is nonlinear, and convex in the first half of pregnancy. After 30 weeks, the growth

rate begins to slow down. So part of a logistic curve is reasonable to model the weight trajectory, which is in the form of

$$W(t) = \frac{L}{1 + exp(-k(t - t_0))}$$

While a logistic curve always starts from zero, we put a baseline parameter c to this model, and it becomes

$$W(t) = \frac{L}{1 + exp(-k(t - t_0))} + c. \tag{3.3.1}$$

In (3.3.1), L is the maximum of of the curve, but we use only part of the curve, so it acts more as a magnifying factor, which is correlated with the total weight gain. k is the steepness of the curve, which can be interpreted as the weight growth rate. t_0 is an inflection point where the curve changes from convex to concave. c is the starting value of the response W and can be interpreted as the prepregnancy weight.

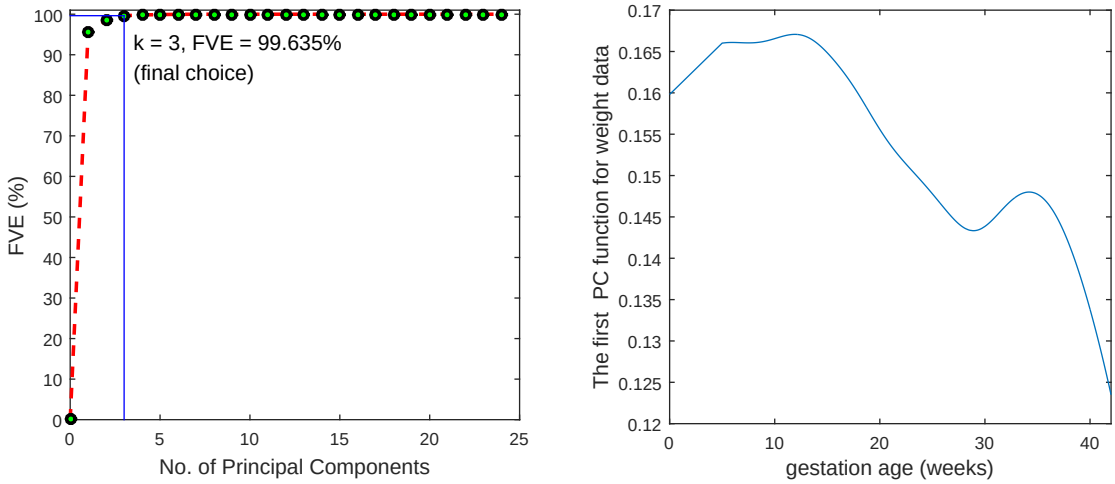
It is observed that on average, the growth rate begins to slow down at 30 weeks. So theinflection point is set to $t_0 = 30$. L,k, and c are to be estimated. Since the total weight gain and prepregnancy weight are the main sources of variation of individual trajectories, we set these two as random effects.

By estimating the weight trajectory of each subject, we have a fitted value for each observation, and the mean squared error (MSE) and root mean squared error (RMSE) can be thus calculated:

$$\text{MSE} = \frac{1}{\sum_{i=1}^{n} N_i} \sum_{i=1}^{n} \sum_{j=1}^{N_i} (\hat{W}(T_{ij}) - W(T_{ij}))^2, \text{RMSE} = \sqrt{\text{MSE}}$$

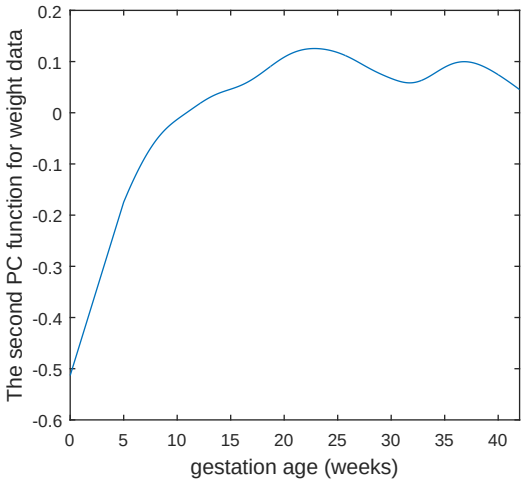
We use the root mean squared error (RMSE) to compare the two methods. The nonlinear mixed effect (NLME) model has an RMSE of 2.6 (kg) and the FPCA approach has

an RMSE of 2.1(kg), which can be seen as a significant improvement. We randomly select one subject in each pre-pregnancy BMI category as defined in Table 6, and their weight trajectories estimated from these two approaches are shown in Figure 10. For subject A, her weight steadily increases during pregnancy, which is of similar pattern to the mean function of all samples; see Figure 6. Both FPCA and NLME yield an estimated trajectory of good agreement with the weight records. However, for subject B, from her weight records, she lost weight at the beginning of her pregnancy and did not regain to her prepregnancy weight even in the 3rd trimester. So her weight gain pattern has large deviation from the common trend as the mean function of all samples. The NLME approach still uses a curve shape similar to subject A (and in fact, similar to the mean function of the whole sample), so the prediction errors are large for three weight records. The trajectory estimated by FPCA is much more adapted to the weight records, and has much smaller residuals. The reason of this discrepancy boil down to two aspects. (i) For NLME, the function has been specified, so the flexibility is very limited. The differences between subjects are all accounted by the random effects, L and c. So for different subjects, the estimated trajectories only differ by a dilation and shift, yet the shape is maintained. When the trajectory of a certain subject is not following the general trend, it is difficult to accurately estimate the trajectory by NLME. (ii) For each subject we have only up to five weight records. When only three or four data points are present, the estimate of NLME by either restricted maximum likelihood or maximum likelihood is prone to be biased. In the FPCA approach, the estimated trajectory is in the form of $\hat{X}_i^K(t) = \hat{\mu}(t) + \sum_{k=1}^{K} \hat{\xi}_{ik}\hat{\phi}_k(t)$, which borrows strength from the entire sample (estimation of $\phi_k(t)$), yet also adapt to the individual observations (estimation of ξ_{ik}). As long as the total covariance structure is consistently estimated, the estimate of trajectory works reasonably well.

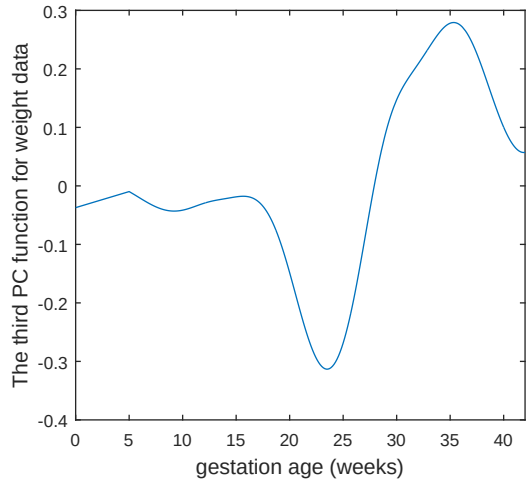


(a)

(b)



(c)



(d)

Figure 9: (a) Scree plot of the weight data and (b-d) The first, second, and third PC functions for weight data. The three PCs account for 95.7%, 2.8%, and 1.1% of the total variation respectively.

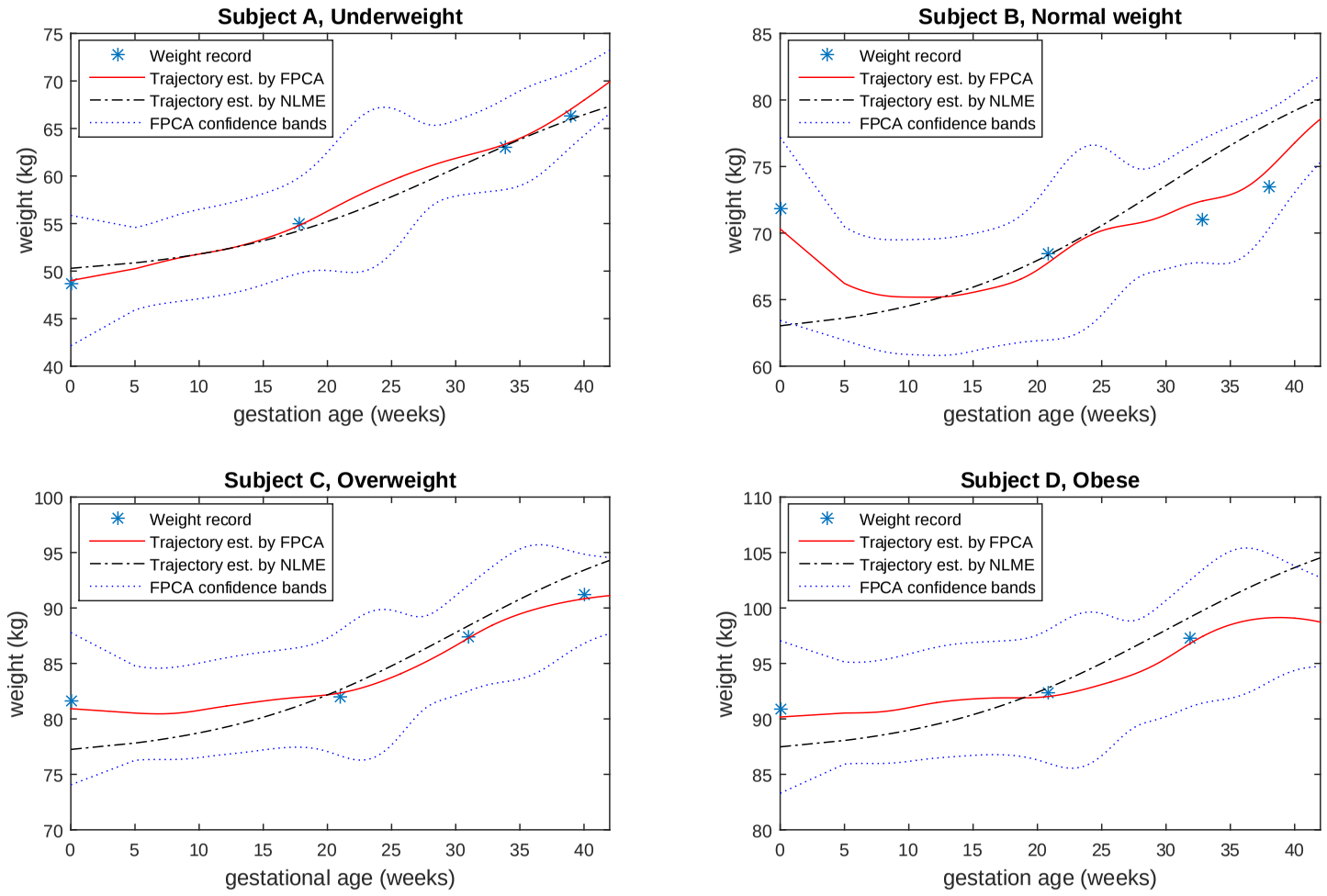


Figure 10: Predicted trajectories and confidence bands of the weight measurements of 4 random subjects

Modeling of Nutrient Intakes-Example of Fat Intake

For the nutrient intakes collected from 24-hour food recalls, the variance is much larger than that of the weight. In addition, we have less data (up to 3 instead of up to 5 time points) for each subject. Taking fat as an example, for the interviewer-administered data, Figure 11 (a) shows the observed individual trajectories of 100 randomly selected subjects and the smooth estimate of the mean function. All the fat intake measurements and the estimated mean are shown in Figure 11(b). From the previous Figure 4 we can see that there was no questionnaire completed before 5 weeks or after 38 weeks. Therefore, unlike the weight data, we are only able to estimate the mean function, and all the other functions thereafter, within $\mathcal{T} = [5, 38]$. Only one PC is selected, explaining 99.9% of the total variance; see Figure 11(c-d). The PC function is monotonically decreasing since 5 weeks, corresponding to a contrast between very early and very late times. This can be due to the prevalent change of dietary pattern during pregnancy which influences the fat intake [6]. Because only one PC $\hat{\phi}(t)$ is selected, the estimated trajectories of all subjects are in the form

$$\hat{X}(t) = \hat{\mu}(t) + \xi_{i1}\hat{\phi}(t),$$

and differ by the FPC score ξ_{i1} only. Correlation estimation is thus trivial as $\text{cor}(\hat{X}(s), \hat{X}(t)) = \text{E}(\xi_1\phi(s)\xi_1\phi(t)) = \xi_1^2 =$ constant always hold. We show only the fitted covariance surface; see Figure 12. From the diagonal we can see that the variance of fat intake is rapidly decreasing.

Four subjects who completed the interviewer-administered 24-hour food recall are randomly selected. The predicted trajectory and point-wise confidence bands of their fat intakes are displayed in Figure 13. The trajectories are all of similar shape as the mean function shown in Figure 11 (a-b), which is almost stationary for the first 30

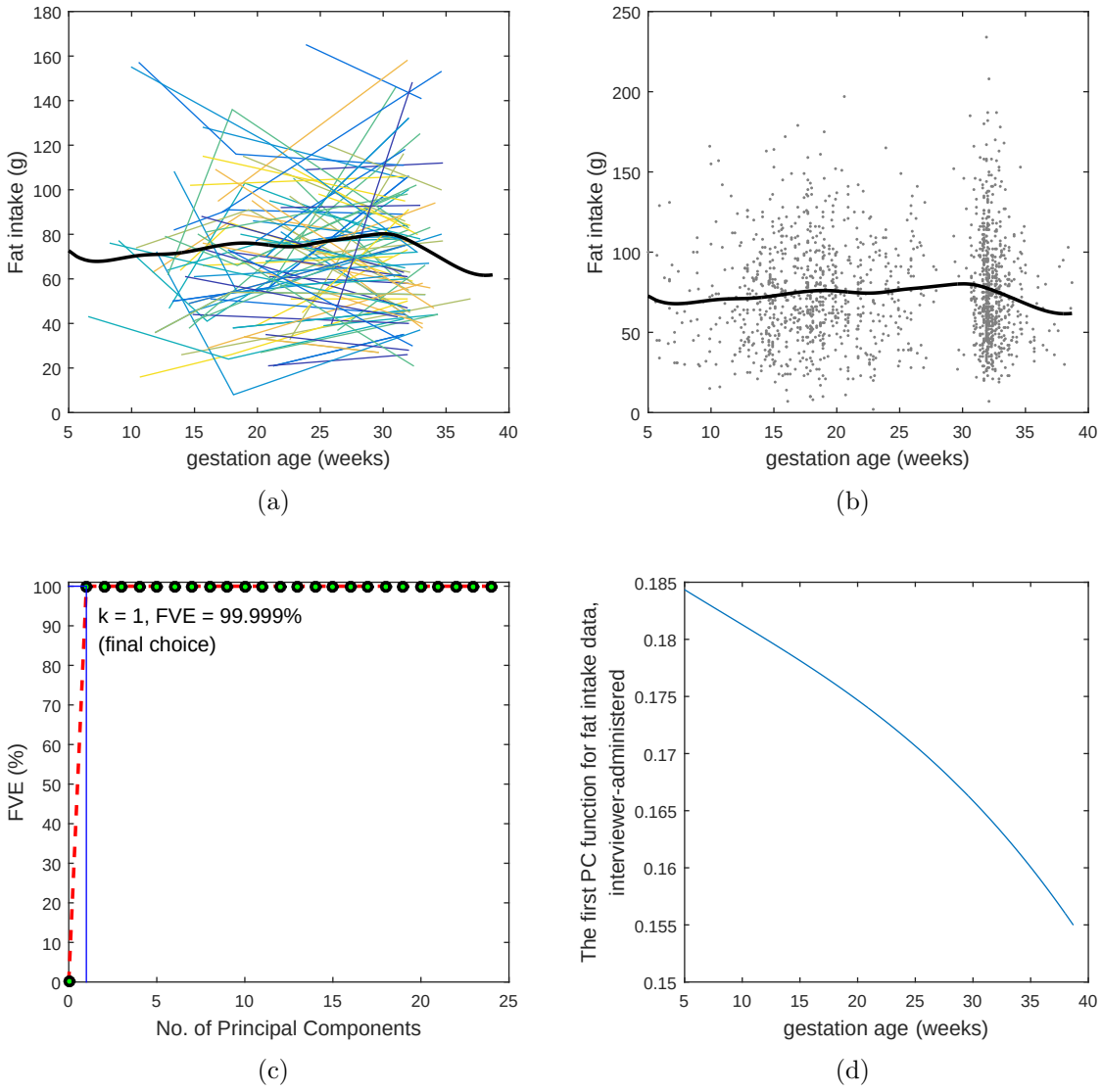


Figure 11: (a) Observed individual trajectories of the fat intake of 100 randomly selected subjects completing the interviewer-administered 24 hour food recall, overlaid with the smooth estimate of the mean function of the all the fat intake data measured from the interviewer-administered 24-hour food recall; (b) All the records of fat intakes measured from the interviewer-administered 24-hour food recall, overlaid with the smooth estimate of the mean function; (c) Scree plot of the fat intake data measured from the interviewer-administered 24-hour food recall; (d) First PC function, indicating the dominant mode of variation.

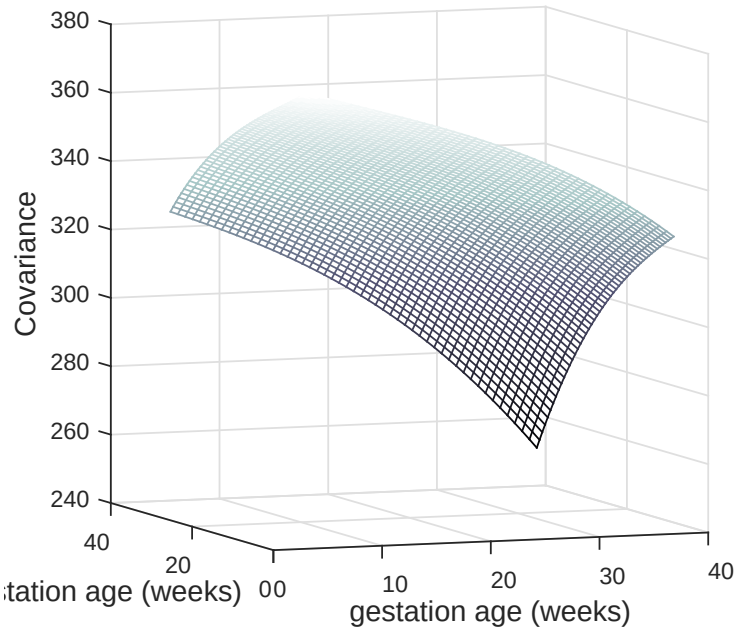


Figure 12: Smooth estimate of the covariance surface of the fat intakes data measured from the interviewer-administered 24-hour food recall. The diagonal shows a variance decreasing with respect to gestation age.

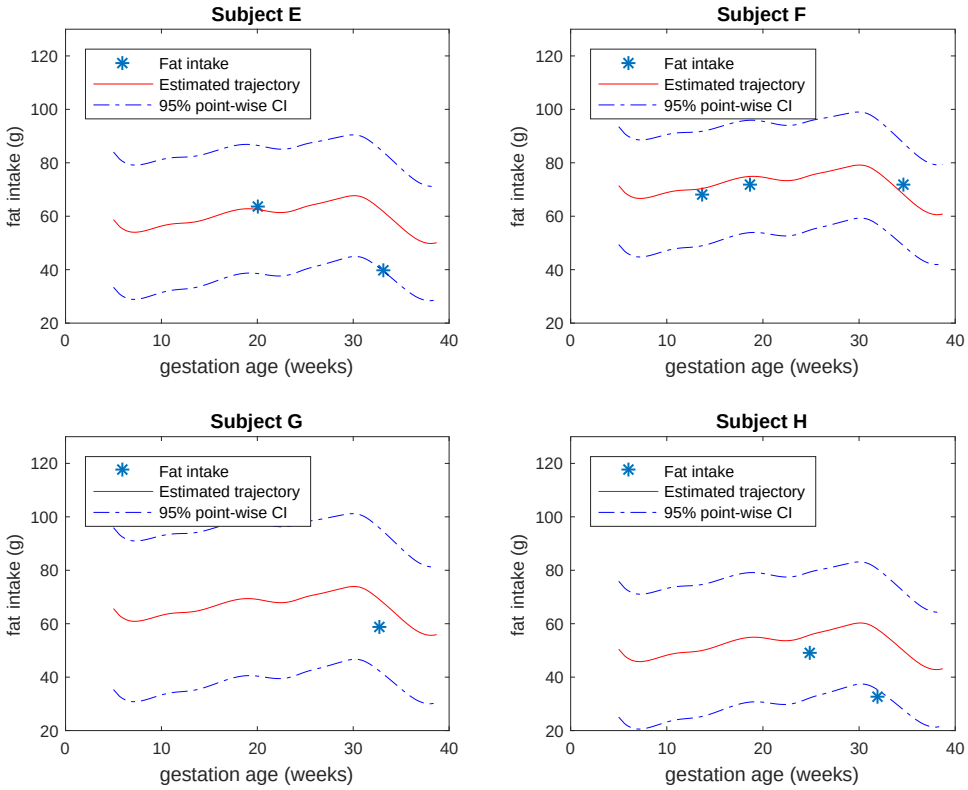


Figure 13: Predicted trajectories and confidence bands of the fat intakes of 4 random subjects who completed the interviewer-administered 24-hour food recall. Some points have considerable distances to the predicted trajectories, or even outside the confidence bands.

weeks, and then decreases. However, from Figure 11 (b), only a few observations of fat intakes were made after 34 weeks, so this decreasing trend may not be reliable. Compared to the weight data, the predicted trajectories are not as close to the observed values. Even though the confidence bands are much wider (compared to the weight data), there are observations outside the confidence bands. The accuracy of the predicted trajectory is not as optimal as the weight data. The corresponding plots of the fat intake data collected using the web-based instrument are shown in Figure 14. We find that the patterns of the mean function and dominant modes of variation (first PC function) are different in two versions of questionnaires. Unlike the smooth mean

function estimated from interviewer-administered instrument, the fat intakes measured from the web-based instrument show that on average, the subjects first increased their fat intake from the beginning of pregnancy, and at around 25 weeks, the fat intakes became more stable. From the assembled pairs of time points in Figure 4, the data points from 21-30 weeks and beyond 36 weeks are quite sparse on the whole data set. So the estimated mean function can be driven by a few data points and may not be as reliable as the estimates from 15-21 weeks and 30-35 weeks. The differences between the estimated trajectories could also be due to of the discrepancy between the two instruments.

There are two PC functions selected for the web-based fat intake data. The first is similar to the mean function to some extends, and the second is a contrast between the early time and late times. The smooth estimate of the correlation surface is shown in Figure 15. The entire surface is above 0.75, indicating that the correlation within one subject is very strong. The dependence between the fat intake at early time and later times dies off relatively rapidly, consistent with the second PC, where a contrast between different times is present. Four subjects who completed the interviewer-administered 24-hour food recall are randomly selected. The predicted trajectory and point-wise confidence bands of their fat intakes are displayed in Figure 13. As the interviewer-administered fat intake data, the predicted trajectories are in the similar shape as the estimated mean function. The width of confidence band gets reduced when more observations are provided. Some points are outside the confidence band, indicating the reliability of the prediction may not be good enough.

3.3.3 Modeling the Weight Gain

In the previous parts of this section, we successfully estimated the trajectories of weight from sparse measurements. During pregnancy, the trajectory can be used to predict

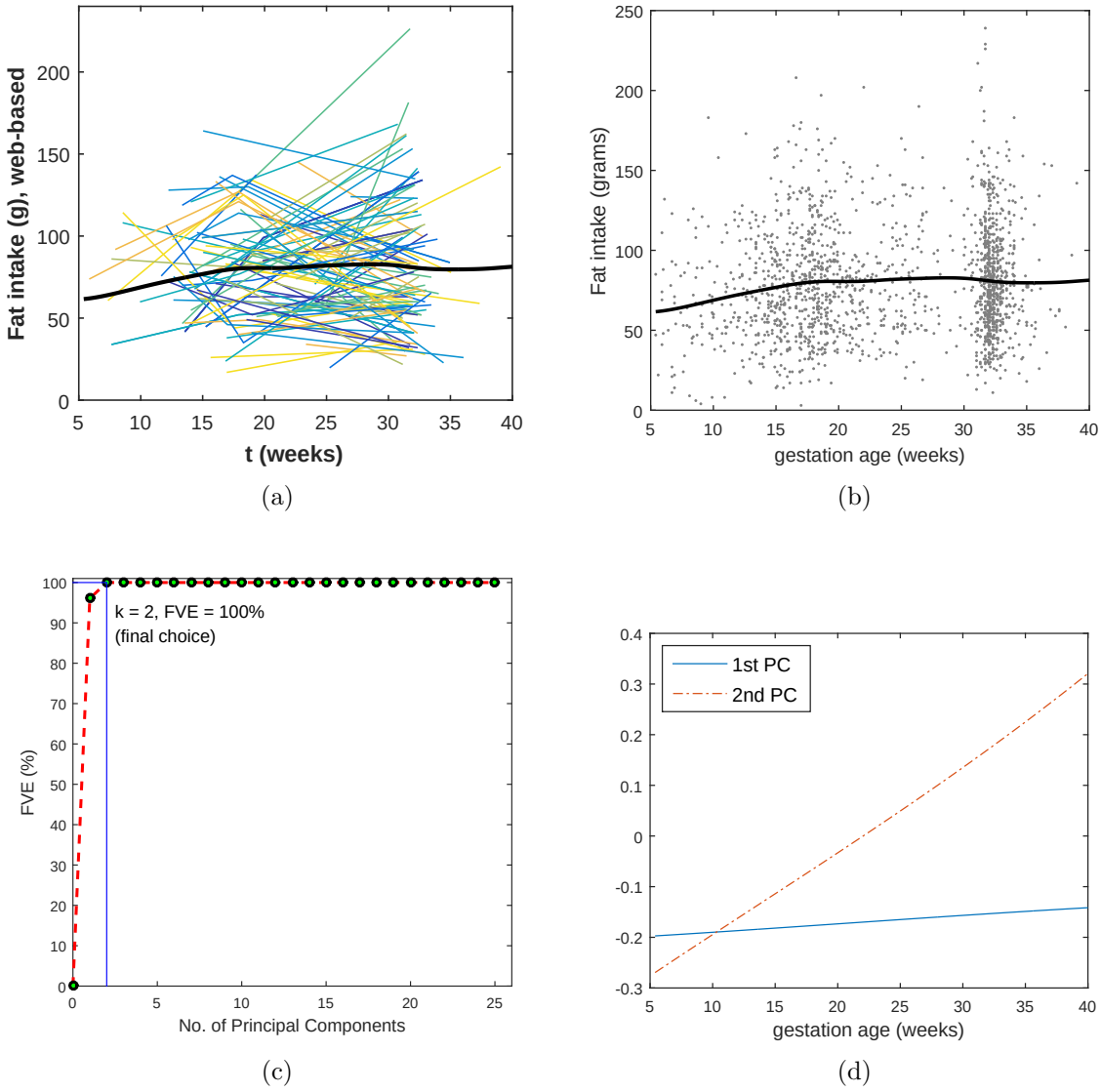


Figure 14: (a) Observed individual trajectories of the fat intake of 100 randomly selected subjects completing the web-based 24 hour food recall, overlaid with the smooth estimate of the mean function of the all the fat intake data measured from the web-based 24-hour food recall; (b) All the records of fat intakes measured from the web-based 24-hour food recall, overlaid with the smooth estimate of the mean function; (c) Scree plot of the fat intake data measured from the web-based 24-hour food recall; (d) First two PC functions, indicating the dominant mode of variation.

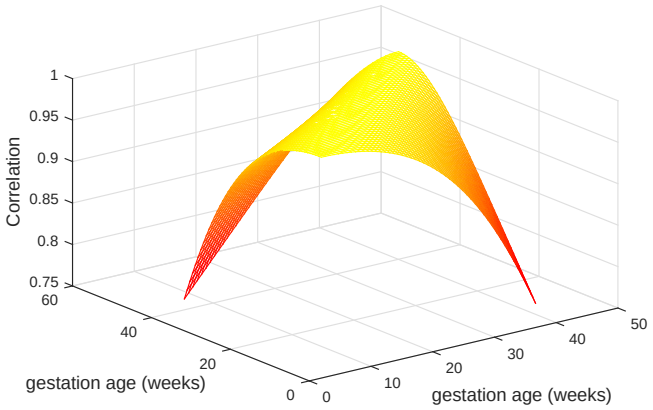


Figure 15: Smooth estimate of the correlation of the fat intake measured from the web-based 24-hour food recall. The dependence between the fat intake at early time and later times dies off relatively rapidly

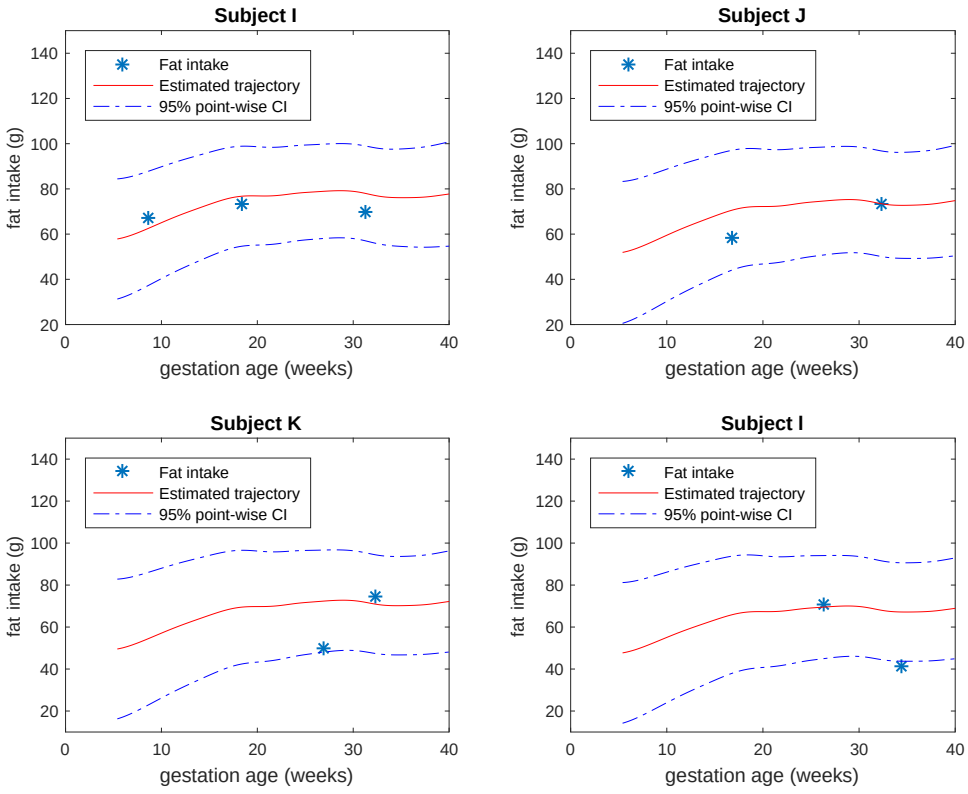


Figure 16: Predicted trajectories and confidence bands of the fat intakes of 4 random subjects who completed the web-based 24-hour food recall. Some points have considerable distances to the predicted trajectories, or even outside the confidence bands.

the total GWG, and provide informative references for interventions such as counseling on dietary intakes and physical activities. In this part, we are interested in how GWG is affected by the possible factors, such as pre-pregnancy BMI, nutrient intakes and physical activities.

In the collected data of APrON study, the weight gain can be directly calculated from the weight records, by defining the total GWG as

$$G = \max\{W_H, \mathrm{W}_C\} - \mathrm{W}_0. \tag{3.3.2}$$

That is, the greater one between the self-reported highest weight during pregnancy (W_H) and the measured weight in the 3rd trimester (W_C) are selected and treated as the weight right before birth. This value minus the pre-pregnancy weight (W_0) is the total GWG. This GWG is naturally derived from the weight records, yet has some potential drawbacks. (i) When the self-reported W_H is less than the measured W_C, the latter is used as the highest weight. However, we note from previous part that subjects are unlikely to schedule a visit after 37 weeks. The W_C measured before 37 weeks may be different from her weight right before birth. (ii) It is possible that the weight of a pregnant woman drops during the last few weeks. The highest weight may be greater than the weight right before delivery. The definition in (3.3.2) may be overestimating the true GWG.

Due to the discrepancy between the measurements from the two different food intake assessing instruments, we model the data collected by the two instrument separately. Taking the web-based data as example, G defined in (3.3.2) is the response, and the average nutrient intake of macronutrients (fat, protein, and carbohydrate), average of the PA indices are predictors. Neither the multivariate linear regression or the functional regression model explained the variance of the weight gain well. Adding the

Table 8: Comparison of the weight gain regression models with directly calculated weight gain

Response	Predictor(s)	RMSE	Adjusted R^2
G	BMI_p	1.57	0.08
G	BMI_p, $\overline{T}$, $\overline{TI}$	1.57	0.08
G	BMI_p, $\hat{T}(t)$, $\hat{TI}(t)$	1.57	0.08
G	BMI_p, $\overline{F}, \overline{P}, \overline{C}$, $\overline{TI}$	1.56	0.09

RMSE: Root Mean Square Error; BMI_p: prepregnancy BMI; $\overline{T}$: average total calories intake;$\overline{F}$: average fat intake; $\overline{P}$: average protein intake; $\overline{C}$: average Carbohydrate intake; $\overline{TI}$: average total PA index; $\hat{T}(t)$: estimated total calories intake trajectory; $\hat{TI}(t)$: estimated total PA index trajectory.

nutrient intake/PA data as additional predictors to the SLR model with prepregnancy BMI as the predictor did not improve the RMSE or R^2; see Table 8.

Natural log transform is a widely used technique in data analysis [120]. We transform the weight, and the difference is now the log of a fraction, representing the relative gain of a subject.

$$LG := \log(\max\{W_H, W_C\}) - \log(W_0) = \log\left(\frac{\max\{W_H, W_C\}}{W_0}\right).$$

However, even after the log-transformation, the SLR model with LG as the response and prepregnancy BMI as predictor still does not explain the variance of the weight gain well, with $R^2 < 0.1$.

Exploring the data we found that, for subjects with extremely large or small weight gains, the regression models above could not predict the weight gain well. As mentioned above, this method of weight gain calculation has possibility to overestimate the total GWG. An alternative approach is to estimate the weight gain from the estimated

trajectory:

$$G' = (\text{weight estimated at GA at birth}) - (\text{weight estimated at } t = 0).$$

The log transformed value can be accordingly defined:

$$LG' = \log \frac{\text{weight estimated at GA at birth}}{\text{weight estimated at } t = 0}.$$

The weight gain estimated from trajectory could account for the weight gain in the last few weeks even if the subject does not have any weight record. Moreover, FPCA method has a smoothing effect. For subjects with large/small weight gain, the estimated trajectory tends to have a lowered/increased weight gain, and thus the estimated weight gain is closer to the mean weight gain. This effect can be seen from Figure 17. The weight gain estimated from the trajectory, G', is highly correlated with the weight gain directly calculated, G, but tends to be smaller than G when G is large, or larger than G when G is small. This approach of estimating GWG may reduce the bias of under-reporting the prepregnancy BMI or over-reporting the W_H, as noted in [10, 79]. As it is shown in Table 8 that the nutrient intake/PA information do not improve R^2 of the model, we consider using only prepregnancy BMI as predictor, and employ all the weight data instead of dividing them by the type of instruments. The new model works very well in terms of the R^2. For a simple linear regression model, where we use pre-pregnancy BMI as the only predictor, the R^2 is 0.48, which means near half of the variance of the G' can be accounted by the variance of pre-pregnancy BMI. The performances of the candidate models are listed in Table 9. As seen from the table, the model where LG' is the response has the highest adjusted R^2 and smallest RMSE.

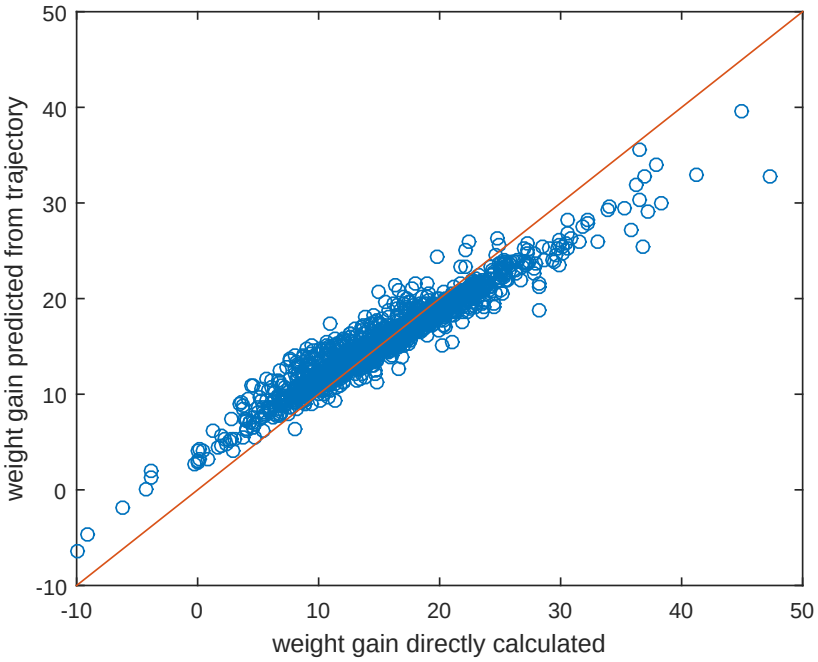


Figure 17: Weight gain predicted from the trajectory vs. weight gain directly calculated

Table 9: Comparison of the weight gain regression models with log transformed/untransformed, estimated weight gain from trajectory

Estimated Model	RMSE	Adjusted R^2
$G = 20.62 - 0.23\text{BMI}_p$	5.80	0.03
$LG = 0.42 - 0.01\text{BMI}_p$	0.07	0.25
$G' = 23.80 - 0.39\text{BMI}_p$	3.88	0.18
$LG' = 0.45 - 0.01\text{BMI}_p$	0.05	0.47

Table 10: Comparison of the weight gain regression models with and without nutrient intake/PA information

Response	Predictor(s)	RMSE	Adjusted R^2
LG'	BMI_p	0.03	0.59
LG'	BMI_p, $\overline{T}$, $\overline{TI}$	0.03	0.60
LG'	BMI_p, $\hat{T}(t)$, $\hat{TI}(t)$	0.03	0.60

The estimated prediction equation is

$$\widehat{LG'} = 0.45 - 0.01\text{BMI}_p. \tag{3.3.3}$$

We found that adding the food intake or physical activities data to the predictor has no significant improvement for the R^2 of the regression model. Taking the web-based data as example, the comparison of the models with/without the estimated trajectories of total calories and total PA index is shown in Table 10.

We conclude that the majority of variance has already been explained by pre-pregnancy BMI. The scatter plot of the model (3.3.3) is shown in Figure 18, which clearly shows that the weight gain predicted from trajectory is negatively associated with pre-pregnancy BMI. For an underweight woman, she is likely to gain more weight during pregnancy than a pregnant woman who is overweight before pregnancy.

3.4 Discussion

GWG is an important indicator of maternal and fetus health. Inadequate and excessive GWGs are associated with various negative pregnancy outcomes [17, 42] and have negative impact for maternal and infant health in the long run [117]. However, in developed countries such as the United States and Canada, meeting the recommended

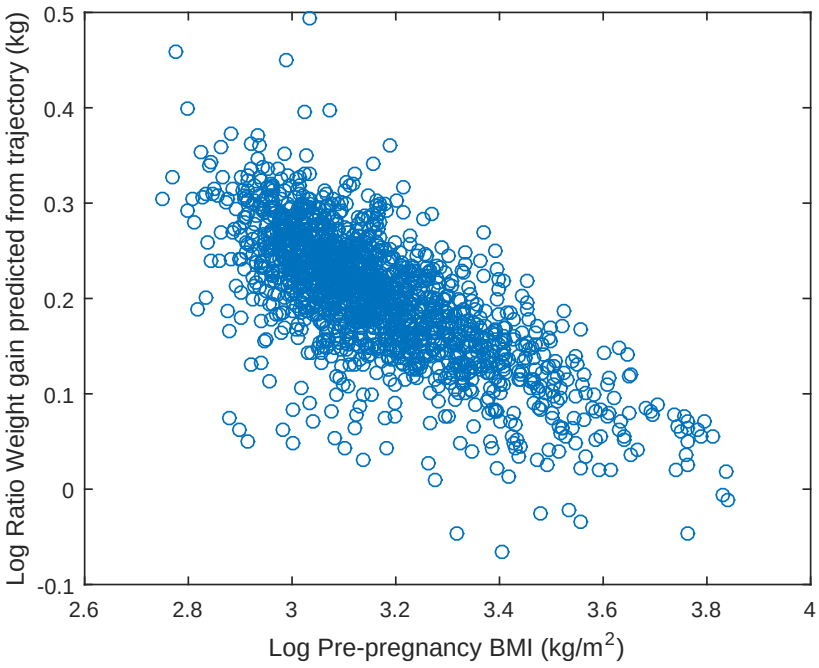


Figure 18: Scatter plot of the log transformed weight gain vs pre-pregnancy BMI

GWG guideline is difficult for the majority of pregnant women [51, 98]. This situation makes it important to monitor GWG during pregnancy to promote healthy weight gain. Previous studies focused on the numerical aspect of GWG, such as total GWG and weight gain rate. Yet in clinical practice, the weight trajectory may be more informative for health care providers. Literatures show that interventions such as customized counseling and education combined with individual weight monitoring were effective to help pregnant women establish healthy weight gain. These interventions call for a personalized weight growth trajectory as reference, especially a predicted weight trajectory with partial existing weight records. If a weight trajectory could be predicted with several existing weight records, the trajectory will help monitor the maternal health, and the proper interventions can be used. APrON is a large cohort study of maternal nutrition and GWG. In the study, the gestational weights and food intakes as well as PA of a large cohort of pregnant women were measured. Our study

has two aims, (i) estimation of the trajectories of weight and nutrient intakes; (ii) modeling of the relationship between weight and prepregnancy BMI together with food intake/PA to predict GWG.

We used FPCA for longitudinal sparse data to estimate the weight trajectory of each pregnant women during pregnancy. The modeling performance is satisfactory for weight, with RMSE=2.1 (kg). Compared to the NLME approach with RMSE=2.6 (kg), this is a significant improvement. It is demonstrated in Figure 10 that the FPCA approach adapts to the observations better, especially when the trajectory is not in the same pattern with the overall mean. The PFCA approach shows much better flexibility and robustness in the modeling.

For the nutrient intakes, the agreement between observations and estimated trajectories are not as satisfactory as weight; see Figure.13 and 16. The performance difference may due to, but not restricted to the following reasons:

1. Variability of nutrients intakes are much larger than body weight, not only for different subjects, but also for a single subject from different time points. A subject might have a large intake of fat while taking little carbohydrates in a certain day, but may have different intake or even doing the opposite on another day. The 24-hour food recall only captures food intake in a single day and could result in quite different intake patterns from one day to another. On the contrary, body weight is a cumulative variable which does not change too much in a certain period of time. The trajectory pattern follow the same overall trend with less variability.

2. Large measurement errors of food intake make the trajectory prediction difficult. Measurement of body weight has much smaller measurement errors on the scale of kilograms. On the other hand, the nutrients intakes computed by 24h food

recall questionnaires are subject to considerable measurement errors [16]. They come from various sources, including inaccurate estimation of food serving sizes, omitted food items, etc

3. More information are available for the weight trajectory modeling than nutrient intakes. We require each subject in the modeling to have a valid self-reported pre-pregnancy weight. So for weight measurements, we have at least one extra record of weight than that of food intake/PA. Most subjects also have the number of W_H. We have more data pointnmms for the weight trajectory estimation than for the nutrient intakes estimation.

For the second aim, the association between the covariates and the directly calculated GWG obtained by (3.3.2). In terms of R^2, the association is not strong. However, the estimated GWG from the trajectory is strongly correlated with pre-pregnancy BMI.This result is consistent with the GWG guidelines proposed by IOM and Health Canada as per [123, 69]. That is, on average, women with lower pre-pregnancy BMI should gain more weight than those with higher pre-pregnancy BMI.

Although it is reasonable and also promising to include the food intake/PA as explanatory variables for the estimation of GWG, it does not improve the model fit significantly. The R^2 does not change (0.51-0.52 vs 0.51) much. Both numeric and functional form of the food intake/PA were tried, and different combinations of the food intake/PA function forms did not make a meaningful contribution to R^2, which are not reported here.

Several aspects of this study should be explored further. One is that the FPCA method employed here is suitable for population with individuals "measured at a dense grid of regularly spaced time points". This assumption is violated as shown in Figure 4. More points are in (i) the cluster with $(30, 35) \times (5, 27)$, corresponding to the

subjects who had records in the both the 2nd and 3rd trimester, or both the 1st and 3rd trimester; and (ii) the cluster $(37, 42) \times (5, 37)$, corresponding to the subjects who had W_H records and at least one record in the three trimesters. There are empty areas along the 45° line $T_{ij} = T_{ik}$: $(0, 13) \times (0, 13)$, $(16, 27) \times (16, 27)$, $(29, 37) \times (29, 37)$ and $(37, 42) \times (37, 42)$ as subjects who completed the questionnaire in a single trimester or visited the research center at any time after 37 weeks, are relatively rare. In the areas with few data points, the estimated mean function and covariance surface could be skewed by the few points, rather than reflecting the real population pattern. The results may be not as reliable as in the areas with more observations. It can also be seen from Figure 10 that the confidence bands are wider on these intervals. The data collection could be modified accordingly, for example, ask the participants to visit the research center in a random manner, or encourage them to visit on the boundary times of trimesters, such as 13-14 weeks, 26-28 weeks. Weight after 37 weeks can be self-reported as most women have the measurement during prenatal visit to medical clinics. New methods can also be developed to tackle such unevenly spaced time points, possibly yielding more accurate estimates.

The second aspect is the bias of the estimated total GWG from the trajectory. From Figure 18 we see that the estimated total GWG from trajectory underestimates the weight gain for women with larger weight gain, while overestimates the weight gain with smaller weight gain. That is because the FPCA method borrows information from the entire cohort, and the estimation of individual trajectories used only the first K PCs. Trajectories that have a different pattern than the sample mean are likely to have larger FPC scores on the other PCs, which are discarded during the estimation. The estimated trajectory is then closer to the mean function than it should be. Women who were underweight before pregnancy tend to have higher weight gain, which means that their weight trajectory may be steeper than the mean function, so their estimated

total GWGs are lower than actual values. Similarly for women who were overweight or obese before pregnancy, their estimated total GWGs are higher. Different estimation methods of the weight trajectories could be employed in these different prepregnancy BMI groups for more reliable estimation, or an index predictor of the prepregnancy BMI categories can be added to the estimation of trajectories.

Last but not least, in our estimation of total GWG, adding the self-reported 24 hour food intake and PA data in the regression model did not increase the percentage of total variance explained. This indicates that those self-reported data are not providing sufficiently useful information for estimating the weight gain. In Chapter 2 we also showed that for the same subject, the two different instruments can have considerable discrepancy for the measurement of nutrient intakes. There can be much noise information in the nutrient intake data, bringing difficulties for accurate modeling. The design of the data collection can be correspondingly improved. For example, repeated administrations of 24 hour food recall are more representative for the usual food intake pattern of an individual. Holmes et al. recommended four repeated 24-hour food recalls in a national study of diet and nutrition[54].

Chapter 4

Discussion

Dietary intake and weight gain during pregnancy are two important topics in maternal nutrition. In this thesis, I utilized the data collected in the APrON study and investigated the dietary profile and its association with the GWG pattern of pregnant women in Alberta, Canada.

The two different 24-hour food recall instruments employed in the APrON study were subject to report and measurement errors. The validation study recruited 58 nonpregnant volunteers to complete both instruments for the same period of 24 hours and the nutrient intakes were computed accordingly. The Bland-Altman plot suggested that the recruited subjects tended to slightly under-report the total calories consumption in the self-administered, web-based instrument. The variances of the reporting error among individuals were relatively large. Therefore, it may not be appropriate to directly pool the results from two different instruments for further analysis.

The calibration between the two instruments by SLR showed reasonable correlation for most of the key nutrients, including fat, protein, carbohydrate and vitamins C, D, and iron. But the intakes of choline and folate had weaker correlations between the two instruments. Further investigation revealed that the web-based instrument had an issue extracting the folate intake in its algorithm. For the macronutrients, the increment of Pearson's correlation coefficient by proportional energy adjustment indicated that the nutrient densities were more consistently reported using the two instruments. We recommend to apply the proportional energy adjustment when pooling the intakes

of the macronutrients from the two instruments. The LASSO based calibration significantly improved the R^2 of the calibration equation of folate, with the intakes of fiber, iron, and potassium as predictors. Choline intakes needed further investigation. For the other micronutrients, the improvement of calibration equations by LASSO regularization was present but limited. If the interviewer-administered instrument is used as the gold standard, the LASSO based calibration is suggested.

I obtained satisfactory estimates of the individual weight trajectories during pregnancy for the participants of the APrON study, using the FPCA by conditional expectation. The approach provided a useful tool to predict the weight gain trajectory from weight data of a few time points during pregnancy, and can be used for the clinical counseling. The estimated trajectories of nutrition intakes, however, had larger errors compared to the weight. The larger errors were probably due to the variation of the daily food intake, the measurement errors in the food intake assessment instruments, and fewer measurements (up to three for food intake vs. up to five for weight) obtained during the study. The GWG estimated from trajectory agreed well with the self-reported weight gain, except for those with high or low prepregnancy BMI. The increment of weight on the log scale from the trajectory showed strong negative correlation with prepregnancy BMI. Therefore, my study confirmed the feasibility of the GWG guidelines published by IOM and Health Canada, i.e., underweight women are gaining more weight during pregnancy than overweight or obese women. However, adding the dietary intake and PA information did not result in a significantly higher R^2 in the regression model, which might be due to the large noise in the nutrient intake data.

Aside from the achievements, several limitations of my study deserve further investigation, and corresponding future work would likely improve the results. In the calibration of the two instruments, the sample of non-pregnant volunteers may have

discrepancies in diet patterns with pregnant women, as pregnancy is often accompanied with diet changes [67, 20]. An algorithm issue was revealed when extracting nutrient intakes from the instruments. Researchers may desire further investigation in the two instruments to be used for pregnant women.

In the modeling of weight/food intake trajectories, the FPCA by conditional expectation approach requires the measurements of all individuals to be dense on a regularly spaced time grid, which is violated for the weight or food intake records between trimesters. Estimated trajectory may be not reliable if few measurements were made in a certain time period. These unevenly spaced records may require to modify the proposed method for the modeling of trajectories. The data collection can adjust for food intake assessments (e.g. repeated 24-hour food recalls) to capture the dietary pattern better and obtain stronger correlation with weight gain. An index predictor may be desired for different prepregnancy BMI groups, in order to cope with the discrepancy between the GWG estimated from trajectory and directly calculated for women with very large or small pregnancy BMI.

In summary, to study dietary intakes and weight gain during pregnancy, I used the data collected in the APrON study. Most key nutrients of interest had acceptable correlations between the two instruments by proportional energy adjustment or LASSO based calibration. However, directly pooling of the results from the two instruments might not be appropriate. Estimation of the weight growth trajectory was successful compared to the traditional mixed-effect models. The GWG estimated from trajectory was strongly correlated with the prepregnancy BMI. However, improvement of data collection may be desired to get better association between GWG estimates and dietary intakes.

Bibliography

[1] Abrams, B., Carmichael, S., and Selvin, S. Factors associated with the pattern of maternal weight gain during pregnancy. *Obstetrics & Gynecology 86*, 2 (1995), 170–176.

[2] Agusti, A., Soler-Cataluña, J., Molina, J., Morejon, E., Garcia-Losa, M., Roset, M., and Badia, X. Does the copd assessment test (cattm) questionnaire produce similar results when self-or interviewer administered? *Quality of Life Research 24*, 10 (2015), 2345–2354.

[3] Allen, L. H. Anemia and iron deficiency: effects on pregnancy outcome. *The American journal of clinical nutrition 71*, 5 (2000), 1280s–1284s.

[4] Altman, D. G., and Bland, J. M. Measurement in medicine: the analysis of method comparison studies. *The statistician* (1983), 307–317.

[5] Andersen, L. F., Tomten, H., Haggarty, P., Løvø, A., and Hustvedt, B. Validation of energy intake estimated from a food frequency questionnaire: a doubly labelled water study. *European journal of clinical nutrition 57*, 2 (2003), 279–284.

[6] Anderson, A. S. Pregnancy as a time for dietary change? *Proceedings of the nutrition society 60*, 04 (2001), 497–504.

[7] Antelman, G., Msamanga, G. I., Spiegelman, D., Urassa, E. J., Narh, R., Hunter, D. J., and Fawzi, W. W. Nutritional factors and infectious disease contribute to anemia among pregnant women with human immunodeficiency virus in tanzania. *The Journal of nutrition 130*, 8 (2000), 1950–1957.

[8] Baecke, J. A., Burema, J., and Frijters, J. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *The American journal of clinical nutrition 36*, 5 (1982), 936–942.

[9] Bell, R. C., Barker, M., Lawrence, W., Nerenberg, K., and Olson, C. Promoting healthy maternal weights in pregnancy and postpartum. *Canadian Journal of Diabetes 39* (2015), S11.

[10] Bolton-Smith, C., Woodward, M., Tunstall-Pedoe, H., and Morrison, C. Accuracy of the estimated prevalence of obesity from self reported height and weight in an adult scottish population. *Journal of Epidemiology and Community Health 54*, 2 (2000), 143–148.

[11] Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. *Classification and regression trees*. CRC press, 1984.

[12] Briggs, M. A., Rumbold, P. L., Cockburn, E., Russell, M., and Stevenson, E. J. Agreement between two methods of dietary data collection in male adolescent academy-level soccer players. *Nutrients 7*, 7 (2015), 5948–5960.

[13] Brumback, B. A., and Rice, J. A. Smoothing spline models for the analysis of nested and crossed samples of curves. *Journal of the American Statistical Association 93*, 443 (1998), 961–976.

[14] Camargo, C. A., Rifas-Shiman, S. L., Litonjua, A. A., Rich-Edwards, J. W., Weiss, S. T., Gold, D. R., Kleinman, K., and Gillman, M. W. Maternal intake of vitamin d during pregnancy and risk of recurrent wheeze in children at 3 y of age. *The American journal of clinical nutrition 85*, 3 (2007), 788–795.

[15] Campbell, S. Fetal macrosomia: a problem in need of a policy. *Ultrasound in Obstetrics & Gynecology 43*, 1 (2014), 3–10.

[16] Carroll, R. J., Midthune, D., Subar, A. F., Shumakovich, M., Freedman, L. S., Thompson, F. E., and Kipnis, V. Taking advantage of the strengths of 2 different dietary assessment instruments to improve intake estimates for nutritional epidemiology. *American journal of epidemiology* (2012), kwr317.

[17] Cogswell, M. E., Serdula, M. K., Hungerford, D. W., and Yip, R. Gestational weight gain among average-weight and overweight womenwhat is excessive? *American Journal of Obstetrics and Gynecology 172*, 2 (1995), 705–712.

[18] Committee on Maternal Nutrition, F., and Nutrition Board, N. R. C. *Nutrition in Pregnancy and Lactation: A Report to the Children's Bureau*. National Academy of Sciences, 1967.

[19] Conway, J. M., Ingwersen, L. A., Vinyard, B. T., and Moshfegh, A. J. Effectiveness of the us department of agriculture 5-step multiple-pass method in assessing food intake in obese and nonobese women. *The American journal of clinical nutrition 77*, 5 (2003), 1171–1178.

[20] Cook, A., Holliman, R., Gilbert, R., Buffolano, W., Zufferey, J., Petersen, E., Jenum, P., Foulon, W., Semprini, A., and Dunn, D. Sources of toxoplasma infection in pregnant women: European multicentre case-control studycommentary: Congenital toxoplasmosisfurther thought for food. *Bmj 321*, 7254 (2000), 142–147.

[21] Cook, A., Pryer, J., and Shetty, P. The problem of accuracy in dietary surveys. analysis of the over 65 uk national diet and nutrition survey. *Journal of epidemiology and community health 54*, 8 (2000), 611–616.

[22] Cox, J. T., and Phelan, S. T. Nutrition during pregnancy. *Obstetrics and gynecology clinics of North America 35*, 3 (2008), 369–383.

[23] Crawford, M., Williams, G., Hassam, A., and Whitehouse, W. Essential fatty acids and fetal brain growth. *The Lancet 307*, 7957 (1976), 452–453.

[24] Csizmadi, I., Kahle, L., Ullman, R., Dawe, U., Zimmerman, T. P., Friedenreich, C. M., Bryant, H., and Subar, A. F. Adaptation and evaluation of the national cancer institute's diet history questionnaire and nutrient database for canadian populations. *Public health nutrition 10*, 01 (2007), 88–96.

[25] Danielzik, S., Czerwinski-Mast, M., Langnäse, K., Dilba, B., and Müller, M. Parental overweight, socioeconomic status and high birth weight are the major determinants of overweight and obesity in 5–7 y-old children: baseline data of the kiel obesity prevention study (kops). *International journal of obesity 28*, 11 (2004), 1494–1502.

[26] Denomme, J., Stark, K. D., and Holub, B. J. Directly quantitated dietary (n-3) fatty acid intakes of pregnant canadian women are lower than current dietary recommendations. *The Journal of nutrition 135*, 2 (2005), 206–211.

[27] Diggle, P., Heagerty, P., Liang, K.-Y., and Zeger, S. *Analysis of longitudinal data*. OUP Oxford, 2013.

[28] Eyles, D., Burne, T., and McGrath, J. Vitamin d in fetal brain development. In *Seminars in cell & developmental biology* (2011), vol. 22, Elsevier, pp. 629–636.

[29] Fan, J., and Gijbels, I. *Local polynomial modelling and its applications: monographs on statistics and applied probability 66*, vol. 66. CRC Press, 1996.

[30] Fitzmaurice, G., Davidian, M., Verbeke, G., and Molenberghs, G. *Longitudinal data analysis*. CRC Press, 2008.

[31] Florindo, A. A., Latorre, M. d. R. D. d., et al. Validation and reliability of the baecke questionnaire for the evaluation of habitual physical activity in adult men. *Revista Brasileira de Medicina do Esporte 9*, 3 (2003), 129–135.

[32] Fraser, A., Tilling, K., Macdonald-Wallis, C., Hughes, R., Sattar, N., Nelson, S. M., and Lawlor, D. A. Associations of gestational weight gain with maternal body mass index, waist circumference, and blood pressure measured 16 y after pregnancy: the avon longitudinal study of parents and children (alspac). *The American journal of clinical nutrition 93*, 6 (2011), 1285–1292.

[33] Fraser, G. E., Jaceldo-Siegl, K., Henning, S. M., Fan, J., Knutsen, S. F., Haddad, E. H., Sabaté, J., Beeson, W. L., and Bennett, H. Biomarkers of dietary intake are correlated with corresponding measures from repeated dietary recalls and food-frequency questionnaires in the adventist health study-2. *The Journal of nutrition 146*, 3 (2016), 586–594.

[34] Freedman, L. S., Commins, J. M., Moler, J. E., Arab, L., Baer, D. J., Kipnis, V., Midthune, D., Moshfegh, A. J., Neuhouser, M. L., Prentice, R. L., et al. Pooled results from 5 validation studies of dietary self-report

instruments using recovery biomarkers for energy and protein intake. *American journal of epidemiology* (2014), kwu116.

[35] Fronczak, C. M., Barón, A. E., Chase, H. P., Ross, C., Brady, H. L., Hoffman, M., Eisenbarth, G. S., Rewers, M., and Norris, J. M. In utero dietary exposures and risk of islet autoimmunity in children. *Diabetes care 26*, 12 (2003), 3237–3242.

[36] Georgieff, M. K. The role of iron in neurodevelopment: fetal iron deficiency and the developing hippocampus. *Biochemical Society Transactions 36*, 6 (2008), 1267–1271.

[37] Giavarina, D. Understanding bland altman analysis. *Biochemia medica 25*, 2 (2015), 141–151.

[38] Giddens, J. B., Krug, S. K., Tsang, R. C., Guo, S., Miodovnik, M., and PRADA, J. A. Pregnant adolescent and adult women have similarly low intakes of selected nutrients. *Journal of the American Dietetic Association 100*, 11 (2000), 1334–1340.

[39] Godfrey, K., Robinson, S., Barker, D., Osmond, C., and Cox, V. Maternal nutrition in early and late pregnancy in relation to placental and fetal growth. *Bmj 312*, 7028 (1996), 410.

[40] Gordon, T., Fisher, M., and Rifkind, B. M. Some difficulties inherent in the interpretation of dietary data from free-living populations. *The American journal of clinical nutrition 39*, 1 (1984), 152–156.

[41] Group, M. V. S. R., et al. Prevention of neural tube defects: results of the medical research council vitamin study. *The lancet 338*, 8760 (1991), 131–137.

[42] Gunderson, E. P., and Abrams, B. Epidemiology of gestational weight gain and body weight changes after pregnancy. *Epidemiologic reviews 21*, 2 (1999), 261–275.

[43] Guo, W. Functional mixed effects models. *Biometrics 58*, 1 (2002), 121–128.

[44] Guo, W. Functional data analysis in longitudinal settings using smoothing splines. *Statistical methods in medical research 13*, 1 (2004), 49–62.

[45] Guo, Y., Hastie, T., and Tibshirani, R. Regularized linear discriminant analysis and its application in microarrays. *Biostatistics 8*, 1 (2007), 86–100.

[46] Hall, P., Müller, H.-G., and Wang, J.-L. Properties of principal component methods for functional and longitudinal data analysis. *The annals of statistics* (2006), 1493–1517.

[47] Hanning, R. M., Royall, D., Toews, J. E., Blashill, L., Wegener, J., and Driezen, P. Web-based food behaviour questionnaire: validation with grades six to eight students. *Canadian Journal of Dietetic Practice and Research 70*, 4 (2009), 172–178.

[48] Harms, L. R., Burne, T. H., Eyles, D. W., and McGrath, J. J. Vitamin d and the brain. *Best Practice & Research Clinical Endocrinology & Metabolism 25*, 4 (2011), 657–669.

[49] Harris, J. R. *Subcellular Biochemistry: Ascorbic Acid: Biochemistry and Biomedical Cell Biology*, vol. 25. Springer Science & Business Media, 2013.

[50] Hastie, T., Friedman, J., and Tibshirani, R. Model assessment and selection. In *The elements of statistical learning*. Springer, 2001, pp. 193–224.

[51] Helms, E., Coulson, C. C., and Galvin, S. L. Trends in weight gain during pregnancy: a population study across 16 years in north carolina. *American journal of obstetrics and gynecology 194*, 5 (2006), e32–e34.

[52] Hercberg, S. Prenatal nutrition guidelines for health professionals-gestational weight gain. majesty the queen in right of canada, represented by the minister of health canada 2010. isbn: H164-126/2010e-pdf cat. no: 978-1-100-16831-9. Tech. rep., H164-126/2010E-PDF, 2010.

[53] Hoerl, A. E., and Kennard, R. W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics 12*, 1 (1970), 55–67.

[54] Holmes, B., Dick, K., and Nelson, M. A comparison of four dietary assessment methods in materially deprived households in england. *Public health nutrition 11*, 05 (2008), 444–456.

[55] Hunt, I. F., Jacob, M., Ostegard, N., Masri, G., Clark, V. A., and Coulson, A. H. Effect of nutrition education on the nutritional status of low-income pregnant women of mexican descent. *The American journal of clinical nutrition 29*, 6 (1976), 675–684.

[56] Jebeile, H., Mijatovic, J., Louie, J. C. Y., Prvan, T., and Brand-Miller, J. C. A systematic review and meta-analysis of energy intake and weight gain in pregnancy. *American journal of obstetrics and gynecology* (2015).

[57] Jiang, T., Christian, P., Khatry, S. K., Wu, L., and West, K. P. Micronutrient deficiencies in early pregnancy are common, concurrent, and vary by season among rural nepali pregnant women. *The Journal of nutrition 135*, 5 (2005), 1106–1112.

[58] Johansson, I., Hallmans, G., Wikman, Å., Biessy, C., Riboli, E., and Kaaks, R. Validation and calibration of food-frequency questionnaire measurements in the northern sweden health and disease cohort. *Public health nutrition 5*, 03 (2002), 487–496.

[59] Kac, G., Benício, M. H., Velásquez-Meléndez, G., Valente, J. G., and Struchiner, C. J. Gestational weight gain and prepregnancy weight influence postpartum weight retention in a cohort of brazilian women. *The Journal of nutrition 134*, 3 (2004), 661–666.

[60] Kaiser, L. L., and Allen, L. Position of the american dietetic association: nutrition and lifestyle for a healthy pregnancy outcome. *Journal of the American Dietetic Association 102*, 10 (2002), 1479–1490.

[61] Kaplan, B. J., Giesbrecht, G. F., Leung, B. M., Field, C. J., Dewey, D., Bell, R. C., Manca, D. P., O'Beirne, M., Johnston, D. W., Pop, V. J., et al. The alberta pregnancy outcomes and nutrition (apron) cohort study: rationale and methods. *Maternal & child nutrition 10*, 1 (2014), 44–60.

[62] Kipnis, V., Freedman, L. S., Brown, C. C., Hartman, A., Schatzkin, A., and Wacholder, S. Interpretation of energy adjustment models for nutritional epidemiology. *American journal of epidemiology 137*, 12 (1993), 1376–1380.

[63] Kirkpatrick, S. I., Thompson, F. E., Subar, A. F., Douglass, D., Zimmerman, T. P., Kahle, L. L., George, S. M., and Potischman, N. Validity of the national cancer institute's automated self-administered 24-hour recall (asa24): results of a feeding study. *The FASEB Journal 27*, 1_MeetingAbstracts (2013), 230–4.

[64] Kleinman, J., et al. Maternal weight gain during pregnancy: determinants and consequences. *National Center for Health Statistics* (1990).

[65] Lee, S., Young, B. E., Cooper, E. M., Pressman, E., Queenan, R. A., Olson, C. M., Guillet, R., and OBrien, K. O. Nutrient inadequacy is prevalent in pregnant adolescents, and prenatal supplement use may not fully compensate for dietary deficiencies. *ICAN: Infant, Child, & Adolescent Nutrition* (2014), 1941406414525993.

[66] Lemons, J. A., Bauer, C. R., Oh, W., Korones, S. B., Papile, L.-A., Stoll, B. J., Verter, J., Temprosa, M., Wright, L. L., Ehrenkranz, R. A., et al. Very low birth weight outcomes of the national institute of child health and human development neonatal research network, january 1995 through december 1996. *Pediatrics 107*, 1 (2001), e1–e1.

[67] Little, R. E. Moderate alcohol use during pregnancy and decreased infant birth weight. *American Journal of Public Health 67*, 12 (1977), 1154–1156.

[68] Lowell, H., and Miller, D. C. Weight gain during pregnancy: adherence to health canada's guidelines. *Health Reports 21*, 2 (2010), 31.

[69] Lowell, H., and Miller, D. C. Weight gain during pregnancy: adherence to health canada's guidelines. *Health Reports 21*, 2 (2010), 31.

[70] Lozoff, B., and Georgieff, M. K. Iron deficiency and brain development. In *Seminars in pediatric neurology* (2006), vol. 13, Elsevier, pp. 158–165.

[71] Macdiarmid, J., and Blundell, J. Assessing dietary intake: who, what and why of under-reporting. *Nutrition research reviews 11*, 02 (1998), 231–253.

[72] Mardia, K., Kent, J., and Bibby, J. Multivariate statistics, 1979.

[73] McCargar, L., Hanning, R., Jessup, L., Lambraki, I., MacDonald, C., and Minaker, L. A web-based approach to assessment of food intake and behaviour of school children and adolescents. *Canadian Journal of Dietetic Practice and Research 64*, 2 (2003), S110.

[74] McCormick, M. C. The contribution of low birth weight to infant mortality and childhood morbidity. *New England journal of medicine 312*, 2 (1985), 82–90.

[75] McIntire, D. D., Bloom, S. L., Casey, B. M., and Leveno, K. J. Birth weight in relation to morbidity and mortality among newborn infants. *New England journal of medicine 340*, 16 (1999), 1234–1238.

[76] Miese-Looy, G., Rollings-Scattergood, J., and Yeung, A. Long-term health consequences of poor nutrition during pregnancy. *Studies by Undergraduate Researchers at Guelph 1*, 2 (2008), 73–81.

[77] Mouratidou, T., Ford, F., and Fraser, R. B. Validation of a food-frequency questionnaire for use in pregnancy. *PUBLIC HEALTH NUTRITION-CAB INTERNATIONAL- 9*, 4 (2006), 515.

[78] Müller, O., and Krawinkel, M. Malnutrition and health in developing countries. *Canadian Medical Association Journal 173*, 3 (2005), 279–286.

[79] Nawaz, H., Chan, W., Abdulrahman, M., Larson, D., and Katz, D. L. Self-reported weight and height: implications for obesity research. *American journal of preventive medicine 20*, 4 (2001), 294–298.

[80] Newhook, L. A., Sloka, S., Grant, M., Randell, E., Kovacs, C. S., and Twells, L. K. Vitamin d insufficiency common in newborns, children and pregnant women living in newfoundland and labrador, canada. *Maternal & child nutrition 5*, 2 (2009), 186–191.

[81] Nyambose, J., Koski, K. G., and Tucker, K. L. High intra/interindividual variance ratios for energy and nutrient intakes of pregnant women in rural malawi show that many days are required to estimate usual intake. *The Journal of nutrition 132*, 6 (2002), 1313–1318.

[82] of Medicine (US). Subcommittee on Nutritional Status, I., during Pregnancy, W. G., of Medicine (US). Subcommittee on Dietary Intake, I., and during Pregnancy, N. S. *Nutrition during pregnancy: part I, weight gain: part II, nutrient supplements*. Natl Academy Pr, 1990.

[83] Ogutu, J. O., Schulz-Streeck, T., and Piepho, H.-P. Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. In *BMC proceedings* (2012), vol. 6, BioMed Central, p. 1.

[84] Olson, C. M. Achieving a healthy weight gain during pregnancy. *Annu. Rev. Nutr. 28* (2008), 411–423.

[85] Olson, C. M., and Strawderman, M. S. Modifiable behavioral factors in a biopsychosocial model predict inadequate and excessive gestational weight gain. *Journal of the American Dietetic Association 103*, 1 (2003), 48–54.

[86] Olson, C. M., Strawderman, M. S., and Reed, R. G. Efficacy of an intervention to prevent excessive gestational weight gain. *American journal of obstetrics and gynecology 191*, 2 (2004), 530–536.

[87] Ono, R., Hirata, S., Yamada, M., Nishiyama, T., Kurosaka, M., and Tamura, Y. Reliability and validity of the baecke physical activity questionnaire in adult women with hip disorders. *BMC musculoskeletal disorders 8*, 1 (2007), 1.

[88] Phelan, S. Pregnancy: a teachable moment for weight control and obesity prevention. *American journal of obstetrics and gynecology 202*, 2 (2010), 135–e1.

[89] Philippaerts, R. M., and Lefevre, J. Reliability and validity of three physical activity questionnaires in flemish males. *American Journal of Epidemiology 147*, 10 (1998), 982–990.

[90] Polley, B. A., Wing, R., and Sims, C. Randomized controlled trial to prevent excessive weight gain in pregnant women. *International Journal of Obesity & Related Metabolic Disorders 26*, 11 (2002).

[91] Prentice, R. L., Mossavar-Rahmani, Y., Huang, Y., Van Horn, L., Beresford, S. A., Caan, B., Tinker, L., Schoeller, D., Bingham, S., Eaton, C. B., et al. Evaluation and comparison of food records, recalls, and frequencies for energy and protein assessment by using recovery biomarkers. *American journal of epidemiology* (2011), kwr140.

[92] Rauh, K., Gabriel, E., Kerschbaum, E., Schuster, T., von Kries, R., Amann-Gassner, U., and Hauner, H. Safety and efficacy of a lifestyle intervention for pregnant women to prevent excessive maternal weight gain: a cluster-randomized controlled trial. *BMC pregnancy and childbirth 13*, 1 (2013), 1.

[93] Rice, J. A., and Silverman, B. W. Estimating the mean and covariance structure nonparametrically when the data are curves. *Journal of the Royal Statistical Society. Series B (Methodological)* (1991), 233–243.

[94] Rifas-Shiman, S. L., Rich-Edwards, J. W., Willett, W. C., Kleinman, K. P., Oken, E., and Gillman, M. W. Changes in dietary intake from the

first to the second trimester of pregnancy. *Paediatric and perinatal epidemiology 20*, 1 (2006), 35–42.

[95] Rogers, I., and Emmett, P. Diet during pregnancy in a population of pregnant women in south west england. alspac study team. avon longitudinal study of pregnancy and childhood. *European journal of clinical nutrition 52*, 4 (1998), 246–250.

[96] Rooney, B. L., and Schauberger, C. W. Excess pregnancy weight gain and long-term obesity: One decade later. *Obstetrics & Gynecology 100*, 2 (2002), 245–252.

[97] Royston, E., et al. The prevalence of nutritional anaemia in women in developing countries: a critical review of available information. *World health statistics quarterly 35*, 2 (1982), 52–91.

[98] Schieve, L. A., Cogswell, M. E., and Scanlon, K. S. Trends in pregnancy weight gain within and outside ranges recommended by the institute of medicine in a wic population. *Maternal and child health journal 2*, 2 (1998), 111–116.

[99] Scholl, T. O., Chen, X., San Khoo, C., and Lenders, C. The dietary glycemic index during pregnancy: influence on infant birth weight, fetal growth, and biomarkers of carbohydrate metabolism. *American journal of epidemiology 159*, 5 (2004), 467–474.

[100] Scholl, T. O., Hediger, M. L., Schall, J. I., Khoo, C.-S., and Fischer, R. L. Dietary and serum folate: their influence on the outcome of pregnancy. *The American journal of clinical nutrition 63*, 4 (1996), 520–525.

[101] Siega-Riz, A. M., Adair, L. S., and Hobel, C. J. Maternal underweight status and inadequate rate of weight gain during the third trimester of pregnancy increases the risk of preterm delivery. *The Journal of nutrition 126*, 1 (1996), 146.

[102] Siega-Riz, A. M., Bodnar, L. M., and Savitz, D. A. What are pregnant women eating? nutrient and food group differences by race. *American journal of obstetrics and gynecology 186*, 3 (2002), 480–486.

[103] Smithells, R., Sheppard, S., and Schorah, C. Vitamin dificiencies and neural tube defects. *Archives of Disease in Childhood 51*, 12 (1976), 944–950.

[104] SRD, R., Julie Armstrong, M., and SRD, R. A short review of dietary assessment methods used in national and scottish research studies.

[105] Storey, K. E., Forbes, L. E., Fraser, S. N., Spence, J. C., Plotnikoff, R. C., Raine, K. D., Hanning, R. M., and McCargar, L. J. Diet quality, nutrition and physical activity among adolescents: the web-span (web-survey of physical activity and nutrition) project. *Public health nutrition 12*, 11 (2009).

[106] Strain, J., McSorley, E. M., van Wijngaarden, E., Kobrosly, R. W., Bonham, M. P., Mulhern, M. S., McAfee, A. J., Davidson, P. W., Shamlaye, C. F., Henderson, J., et al. Choline status and neurodevelopmental outcomes at 5 years of age in the seychelles child development nutrition study. *British Journal of Nutrition 110*, 02 (2013), 330–336.

[107] Subar, A. F., Kirkpatrick, S. I., Mittl, B., Zimmerman, T. P., Thompson, F. E., Bingley, C., Willis, G., Islam, N. G., Baranowski, T.,

McNutt, S., et al. The automated self-administered 24-hour dietary recall (asa24): a resource for researchers, clinicians and educators from the national cancer institute. *Journal of the Academy of Nutrition and Dietetics 112*, 8 (2012), 1134.

[108] Suharno, D., Karyadi, D., West, C., Hautvast, J. G. A., et al. Supplementation with vitamin a and iron for nutritional anaemia in pregnant women in west java, indonesia. *The lancet 342*, 8883 (1993), 1325–1328.

[109] Suitor, C., Gardner, J., and Willett, W. A comparison of food frequency and diet recall methods in studies of nutrient intake of low-income pregnant women. *Journal of the American Dietetic Association 89*, 12 (1989), 1786–1794.

[110] Swensen, A. R., Harnack, L. J., and Ross, J. A. Nutritional assessment of pregnant women enrolled in the special supplemental program for women, infants, and children (wic). *Journal of the American Dietetic Association 101*, 8 (2001), 903–908.

[111] Tanentsapf, I., Heitmann, B. L., and Adegboye, A. R. Systematic review of clinical trials on dietary interventions to prevent excessive weight gain during pregnancy among normal weight, overweight and obese women. *BMC pregnancy and childbirth 11*, 1 (2011), 1.

[112] Thompson, F. E., Subar, A. F., Loria, C. M., Reedy, J. L., and Baranowski, T. Need for technological innovation in dietary assessment. *Journal of the American Dietetic Association 110*, 1 (2010), 48.

[113] Thorsdottir, I., Torfadottir, J. E., Birgisdottir, B. E., and Geirsson, R. T. Weight gain in women of normal weight before pregnancy: complications in pregnancy or delivery and birth outcome. *Obstetrics & Gynecology 99*, 5 (2002), 799–806.

[114] Tibshirani, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* (1996), 267–288.

[115] Udipi, S., Ghugre, P., and Antony, U. Nutrition in pregnancy and lactation. *Journal of the Indian Medical Association 98*, 9 (2000), 548–557.

[116] Villar, J., Purwar, M., Merialdi, M., Zavaleta, N., Anthony, J., De Greeff, A., Poston, L., Shennan, A., et al. World health organisation multicentre randomised trial of supplementation with vitamins c and e among pregnant women at high risk for pre-eclampsia in populations of low nutritional status from developing countries. *BJOG: an International Journal of Obstetrics & Gynaecology 116*, 6 (2009), 780–788.

[117] Wei, J.-N., Sung, F.-C., Li, C.-Y., Chang, C.-H., Lin, R.-S., Lin, C.-C., Chiang, C.-C., and Chuang, L.-M. Low birth weight and high birth weight infants are both at an increased risk to have type 2 diabetes among schoolchildren in taiwan. *Diabetes care 26*, 2 (2003), 343–348.

[118] Willett, W., and Stampfer, M. J. Total energy intake: implications for epidemiologic analyses. *American journal of epidemiology 124*, 1 (1986), 17–27.

[119] Willett, W. C., Howe, G. R., and Kushi, L. H. Adjustment for total energy intake in epidemiologic studies. *The American journal of clinical nutrition 65*, 4 (1997), 1220S–1228S.

[120] Wing, R. R., Koeske, R., Epstein, L. H., Nowalk, M. P., Gooding, W., and Becker, D. Long-term effects of modest weight loss in type ii diabetic patients. *Archives of internal medicine 147*, 10 (1987), 1749–1753.

[121] Wu, B. T., Dyer, R. A., King, D. J., Richardson, K. J., and Innis, S. M. Early second trimester maternal plasma choline and betaine are related to measures of early cognitive development in term infants. *PloS one 7*, 8 (2012), e43448.

[122] Xu, J., Luntamo, M., Kulmala, T., Ashorn, P., and Cheung, Y. B. A longitudinal study of weight gain in pregnancy in malawi: unconditional and conditional standards. *The American journal of clinical nutrition 99*, 2 (2014), 296–301.

[123] Yaktine, A. L., Rasmussen, K. M., et al. *Weight Gain During Pregnancy:: Reexamining the Guidelines*. National Academies Press, 2009.

[124] Yao, F., Müller, H.-G., and Wang, J.-L. Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association 100*, 470 (2005), 577–590.

[125] Zhang, C., Liu, S., Solomon, C. G., and Hu, F. B. Dietary fiber intake, dietary glycemic load, and the risk for gestational diabetes mellitus. *Diabetes care 29*, 10 (2006), 2223–2230.