Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.

UHASSELT
KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Faculty of Sciences
School for Information Technology

Master of Statistics and Data Science

Master's thesis

Modeling the association between Gestational Weight Gain trajectories and maternal/neonatal outcomes

Josline Adhiambo Otieno

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics and Data Science, specialization Biostatistics

SUPERVISOR :
dr. Zoe PIETERS

SUPERVISOR :
Nelson Owuor ONYANGO
Dr. Eric OHUMA
Dr. Lucas MALLA

2020
2021

Maastricht University

Faculty of Sciences
School for Information Technology
Master of Statistics and Data Science

Master's thesis

Modeling the association between Gestational Weight Gain trajectories and maternal/neonatal outcomes

Josline Adhiambo Otieno

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics and Data Science, specialization Biostatistics

SUPERVISOR :
dr. Zoe PIETERS

SUPERVISOR :
Nelson Owuor ONYANGO
Dr. Eric OHUMA
Dr. Lucas MALLA

Acknowledgements

First and above all, I would like to express my sincere gratitude to the Almighty God for His immense love and protection throughout my academic journey at Hasselt University.

Secondly, I would like to underscore the support and guidance of my project supervisors Prof. Dr. Eric Ohuma, Dr. Zoe Piëters, Dr. Nelson Onyango, and Dr. Lucas Malla who made their time for me to help at every instance when I needed help the most.

I am greatly indebted to Hasselt University fraternity for making my stay at Hasselt life changing, and VLIR-UOS for awarding me a prestigious scholarship to pursue my Masters program.

Last but not least, my special thank you goes to the love of my life, Dr. Dioggban Jakperik, my lovely siblings, Colphax, Veron, Charles and Joan, my adorable son, Ian Brilly and my Dad, Mr. George Otieno who have been extremely supportive throughout this journey.

To my special friends, Ndivhuwo Millicent Mudzudzanyi, Karimatou Djenabou, Violet Ankunda, Lembris Njotto, Adama Kazienga, among others, thanks so much for the support and encouragement throughout my study.

Abstract

Background: Gestational weight gain (GWG) refers to the total amount of weight gained between the conception and delivery of an infant. Growth and development of a foetus in a pregnant woman require sufficient GWG. Inappropriate GWG can expose both mother and her baby to adverse health outcomes such as gestational diabetes, hypertension, small- and large-for gestational age babies among others. This study focused on application of multilevel approach to characterize the GWG patterns, and to quantify the association of these patterns and the neonatal outcomes using generalized linear models.

Methods: Using a sample of GWG data set from the INTERBIO-21^{st} project consisting of 2820 mothers, who were enrolled in the study before their 14^{th} gestational age, a random effects model was fitted with weight measurements at each prenatal visit regressed as a function of gestational age. Fractional Polynomials provided the best fitting power of gestational age in the model. Additionally, generalized linear models using 'log' link function were used to model the relation between the random effects parameters and the risk of neonatal/ maternal outcomes including small- and large-for gestational age, birth length, head circumference, gestational diabetes among others.

Results: The study demonstrated that there was less and much variability respectively within and between subjects, in terms of their gestational weights. The findings also showed that increase in the deviations from the global average weight at the minimum value of GA is associated with risk of large for gestational age neonates, newborns with above 90^{th} centile for both birth head circumference and length, pregnancy induced hypertension, gestational diabetes, and increased risk of C-section. On the other hand, increased rate of weight gain from the average resulted in chances of large for gestational age neonates, above 90^{th} centile for birth head circumference newborns, pregnancy induced hypertension, preterm premature rupture of the membranes (PPROM), and failure to progress.

Conclusions: Women who had a low starting weight on average had a faster rate of weight gain, and vice versa. Risk of adverse maternal/ neonatal outcomes attributes to starting pregnancy weight and the rate of GWG. Maternal characteristics such as country of residence, age, blood pressure, having had previous preterm births, neonatal deaths and C-section were also associated with neonatal/ maternal outcomes.

Key Words: Gestational Weight Gain, Pregnancy, Multi-level model, Fractional polynomials, Generalized linear model, Relative risk.

Contents

List of Figures

List of Tables

1 Introduction

The gestation period during which an embryo develops inside the uterus is a critical period. This period is measured as time between conception and delivery, lasting for approximately 39 weeks (Hoffman *et al.*, 2008). Most women go through discomforts due to the big changes to their body needed to support the foetus development from conception to birth. Some of the issues include heartburn, bloating, depression, fatigue, lower pelvic cramping, sore breasts, weight gain, nausea, vomiting among others (O'Brien and Naber, 1992).

Gestational weight gain (GWG) is a fundamental transformation that occurs in a woman's body during pregnancy since it helps the foetus to grow and develop. It refers to the total amount of weight gained between the conception and delivery of an infant. Weight gain or loss during gestational period might have an impact on the immediate or future health of either the mother or her baby. Previous studies showed that women who have excess or inadequate GWG are associated with different health outcomes. The outcomes refer to measurable results for a mother and/ or an infant based on the standards of safety as set by clinical studies (Scarf *et al.*, 2018). For instance, a meta analysis conducted by Voerman *et al.* (2019), concluded that the risk for adverse maternal and infant outcomes varied by GWG and across the range of prepregnancy weights. The GWGs of $14 - 16\ kgs$ (for underweight), $10 - 18\ kgs$ (normal weight), $2 - 16\ kgs$ (overweight), $2 - 6\ kgs$ (obesity grade 1), $0 - 4\ kgs$ (obesity grade 2), and $0 - 6\ kgs$ (obesity grade 3), were linked to low to moderate adverse outcomes such as preeclampsia, gestational hypertension, gestational diabetes, cesarean delivery, preterm birth, and small or large size for gestational age at birth. Voerman *et al.* (2019) also mentioned that adverse outcomes occurred in 37.2% of women, among which 34.7% were underweight, and 61.1% were found obese grade 3.

Within the context of Oken *et al.* (2008), a U-shaped association between maternal GWG and child weight outcomes was observed after adjusting for sociodemographic characteristics such as maternal age, smoking, marital status, household income, paternal education, child race/ethnicity, gestation length, sex, age, but not maternal body mass index (BMI), recording a higher risk of obesity of 10% and 35% in, respectively the lowest (< 10 pounds) and highest (≥ 45 pounds) category of maternal weight gain. According to Drehmer *et al.* (2013), less GWG can contribute to preterm birth and small for gestational age (SGA) neonates which may lead to failure to initiate breastfeeding, high risk for diseases and neonatal morbidity, and infant development delays. Factors that may influence the differences in the weight gain among the pregnant mothers include age, socioeconomic status, ethnicity, level of education, intention for the pregnancy, prenatal advice, and psycho-social characteristics such as attitude towards weight gain, social support, depression, stress, anxiety, and self-efficacy (Hickey, 2000).

Starting pregnancy at a healthy weight and maintaining appropriate GWG is essential for the health of the mother and her feotus. The prevalence of GWG risk factors is relatively high in low-middle income countries (LMICs) (Wang *et al.*, 2020). This could be due to inaccessibility and unaffordability of adequate food (FAO, 2015), having little information about healthy GWG and how to maintain it through good nutritional status and interventions. To lower the adverse health outcomes for both mother and child of less or excessive weight gain, National Academy of Medicine (NAM) formerly known as US Institute of Medicine (IOM), issued updated guidelines for GWG accord-

ing to pre-pregnancy BMI categories of the World Health Organization (WHO) (Rasmussen *et al.*, 2009). This was as shown in Table 1 (NRC, 2010).

Table 1: Recommended amount of GWG in the 2009 IOM guidelines

Prepregnancy BMI category	Total Weight gain
underweight ($< 18.5 kg/m^2$)	$12.5 - 18\ kgs$
normal weight ($18.5 - 24.9 kg/m^2$)	$11.5 - 16\ kgs$
overweight ($25 - 29.9 kg/m^2$)	$7 - 11.5\ kgs$
obese ($\geq 30 kg/m^2$)	$5 - 9\ kgs$

Despite these guidelines, most women still gain weight outside the recommended limits (Li *et al.*, 2015; Johnson *et al.*, 2015). A study which consisted of women who participated in the 2010 and 2011 Pregnancy Risk Assessment Monitoring System from 28 states, investigated the prevalence of GWG according to the guidelines. It was revealed that 20.9%, 32.0%, and 47.2% of expectant mothers gain inadequate, adequate and excessive gestational weight respectively (Deputy *et al.*, 2015). In a systematic review and meta analysis on the prevalence of GWG, it was showed that 27.8% and 39.4% of women globally had gained weight respectively above and below the 2009 IOM guidelines (Martínez-Hortelano *et al.*, 2020). Therefore, weight management in pregnancy is a concern.

1.1 Problem Statement

Women with high or low GWG are likely to experience various adverse maternal outcomes as well as increased risk of neonatal outcomes (Kirkegaard *et al.*, 2014). However, efforts to improve weight management in pregnancy are hindered by; (1) absence of consensual GWG policy worldwide (Scott *et al.*, 2014), (2) lack of evidence that weight interventions based on targets during pregnancy are achievable and improve clinical endpoints, (3) lack of evidence of long term effects controlling GWG for either mother or child (Thangaratinam *et al.*, 2012; Muktabhant *et al.*, 2015; Goldstein *et al.*, 2018). Currently, there are no GWG recommendations used in routine clinical management especially in low middle income countries (LMICs) within Africa. Therefore, more evidence is needed to determine how to help women start pregnancy at a healthy weight for them, how to achieve an appropriate GWG and timely postpartum weight loss. The present study focuses (1) on statistical modelling of the GWG trajectories, using a multilevel approach, and (2) quantifying the association of GWG trajectories and neonatal outcomes.

1.2 Justification of the study

Gestational weight gain is an important indicator which can be used to establish interventions with an aim to improve maternal and neonatal outcomes. The 2009 IOM guidelines are considered more relevant for women in high-income settings, however they can be used as benchmarks to monitor GWG in LMICs (Coffey, 2015). Experience shows that LMICs have unhealthy dietary patterns and high prevalence of sedentary lifestyle, thus there is still increased burden of excessive GWG (Orach and Garimoi, 2009). Inadequate or excess weight gain during pregnancy is linked to worse outcomes of pregnancy. Implying that there is need to develop programmes that can assist pregnant women to gain weight within the recommended values to ensure a healthy growth of the feotus.

The findings of the study can enhance the knowledge on the GWG patterns and association with neonatal outcomes which is useful in the formulation of comprehensive policies for weight management during pregnancy especially in LMICs. Eventually, this will enable mothers and their babies to improve on their health by keeping a moderate weight gain. The information is also relevant to the health-care providers to expectant mothers in an effort to offer counselling and guidance on recommended weight gain during the routine prenatal care.

1.3 Organization of the report

The sections of this report are structured as follows: Section 2 presents the objectives and research questions of the study. Section 3 outlines the study design, gives the data description, exploratory data analysis and a brief discussion on the models used. Section 4 contains the results of the analysis. Section 5 discusses the findings of the analysis. Finally, conclusions and recommendation on the findings are made in Section 6.

2 Objectives and Research Questions of the Study

2.1 Objectives

The main objective of this study is to model gestational weight gain trajectories for predicting maternal/ neonatal outcomes.

The specific objectives are to;

1. Characterize gestational weight gain (GWG) patterns.
2. Quantify the association between the amount and rate of weight gain on maternal/ neonatal outcomes.

Figure 1 illustrated the sketch of the data analysis process carried out to achieve the aforementioned objectives.

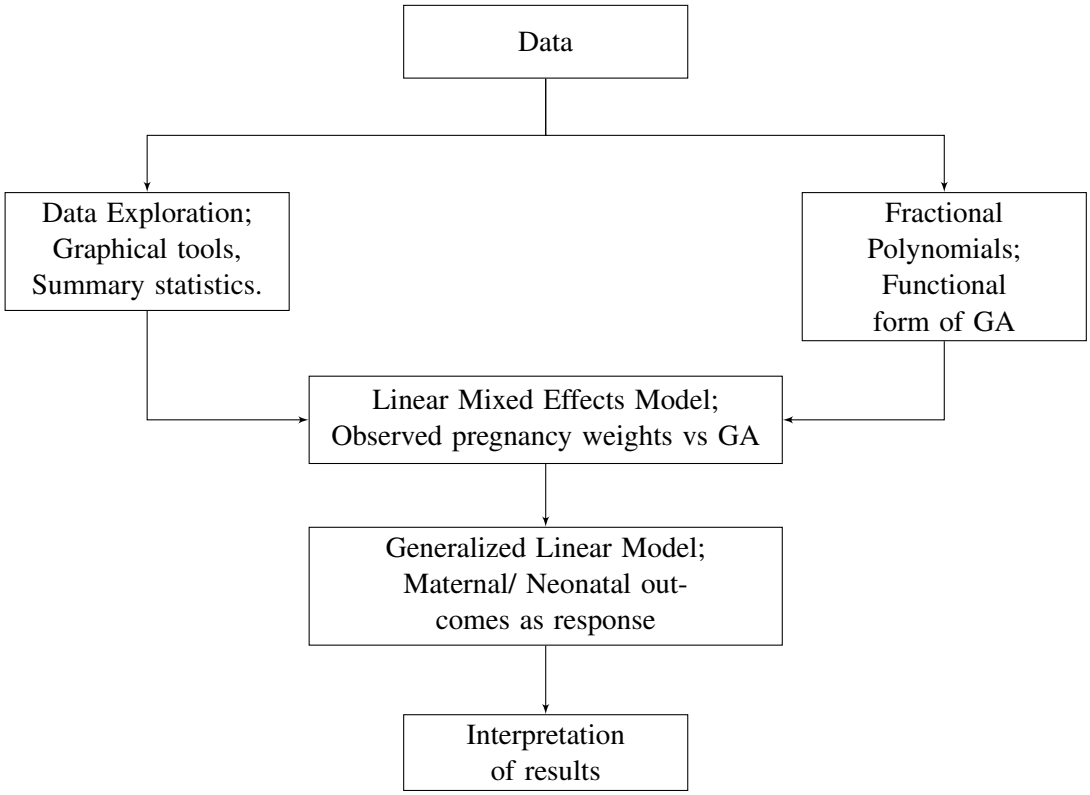


Figure 1: Schematic diagram of the analysis to achieve the objectives.

2.2 Research Questions

The research questions of the study are:

1. What is the pattern of weight gain during the pregnancy period?
2. What is the association between GWG trajectories and maternal/ neonatal outcomes?

3 Methodology

3.1 Study design

This study is based on INTERBIO-21^{st} project, a phase II of the International Fetal and Newborn Growth Consortium for the 21^{st} Century (INTERGROWTH-21^{st}) study. Phase I was an international, multi-center, population-based, research initiative conducted between 2008 to 2015 under coordination of the University of Oxford (Kennedy *et al.*, 2018). The study focused on understanding the changes in the human growth and neuro-development from early pregnancy until 2 years of age with an aim to construct international growth standards to complement WHO child growth standards.

Phase II, a newborn case-control study, conducted between 2012 to 2018, focused on investigating the effects of nutrition, environmental exposures, and clinical conditions on the fetal growth and neuro-development in healthy and complicated pregnancies from seven different populations globally; Kilifi (Kenya), Nairobi (Kenya), Pelotas (Brazil), Karachi (Pakistan), Mae Sot (Thailand), Oxford (UK), and Soweto (South Africa) (Kennedy *et al.*, 2018). Cases composed of newborn preterm (< 38 gestation weeks) and SGA babies with newborn controls constituting of term babies and appropriate for gestational age (AGA), respectively. A large cohort of women, aged 18 years and older, with at most 35 kg/m^2 of BMI, and had naturally conceived pregnancies, was enrolled before their 14^{th} gestational week (Kennedy *et al.*, 2018). The participants were exposed to different environmental risk factors including healthy, sub-optimal conditions and other adverse maternal/ neonatal risk factors (Papageorghiou *et al.*, 2014). The study was divided into; a fetal study in which expectant mothers were monitored from the early pregnancy on-wards to capture every information on the fetal growth patterns, and a neonatal study which monitored women since their delivery time. These studies used INTERGROWTH-21^{st} project tools to record information on anthropometric measurements and pregnancy outcomes on these women both from resource constrained and unconstrained settings (Kennedy *et al.*, 2018). The studies were supported by a grant from the Gates Foundation.

3.2 Description of the data-set

GWG data was collected from $2,820$ mothers, who were enrolled in the study before their 14^{th} gestational week. Data on their newborns was collected until they reached 2 years of age. $13,396$ records are available on GWG, anthropometric measurements, and maternal/ neonatal outcomes for the sample. Table 2 displays the description of the variables in the data set. The data set was highly unbalanced, with study subjects visiting the clinic at different times during pregnancy (between 9 to 35 weeks) for different number of visits, ranging from 2 to 5, whereas gestational week at delivery ranged between 22 to 43 weeks. The woman's weight recorded at enrollment ranged between 33 to 109.2 Kgs.

Table 2: Description of Variables in the data-set

Variables				
Demographic Variables	Father's Age	Alcohol intake	Failure to progress (FTP)	Birth weight (BW)
Patient ID	#previous pregnancy	**Maternal outcomes**	CPD	BW centile
Maternal date of birth	Previous preterm	Hypertension	PPROM	BW zscores
Marital status	#neonatal deaths	PIH	Placental abruptio	Below C10 BW
Years of formal education	Previous Caesarean section	Gestational diabetes	Gestational Age category	Above C90 BW
Level of education	Maternal request	Preeclampsia	**Neonatal outcomes**	BHC
Occupation	Other Maternal reason	Hellp Preeclampsia	Newborn Sex	BHC centiles
Weight	Other Fetal reason	Preterm Labour	Fetal Anaemia	BHC zscores
Height	Date of visit	Mode of delivery	Fetal death	Below C10 HC
BMI	Gestational Age	Vaginal bleeding	Fetal distress	Above C90 HC
BMI category	Gestational age at delivery	Placenta praevia	Newborn ISCU	Birth length (BL)
Systolic blood pressure	Pregnancy weight	Breech presentation	#days in ISCU	BL centiles
Diastolic blood pressure	Weight record	Uterine rupture	Seizures	Below C10 BL
Country	Smoking status	Reduced fetal movement	Preterm	Above C90 BL

Hellp - haemolysis, elevated liver enzymes, low platelet count. **BMI** - body Mass Index. **#** - total number of. **PIH** - Pregnancy Induced Hypertension. **CPD** - Cephalopelvic disproportion. **PPROM** - preterm premature rupture of the membranes. **ISCU** - intensive special care unit. **BHC** - birth head circumference.

3.3 Exploratory Data Analysis

Data exploration was carried out to get a general overview of the data set. In this study, both graphical representation and summary statistics were used to explore the data. For instance, individual profiles plot was fitted to assess the variability of weight between and within subjects over the gestational age; mean evolution plot and a plot of squared residuals of the smoothed mean curve were used to investigate the evolution of the mean structure and explore the variability of pregnancy weight over gestational age, respectively.

3.4 Statistical Methodology

3.4.1 Previous approaches to analyzing GWG

Various approaches have been used in the past studies to analyze the association between GWG and maternal/neonatal outcomes, either using single measurement of GWG or repeated measurements of weight gain. However, the analysis based on single measurement ignores the correlation between total GWG and gestational duration, and hence possibly influence the relationship between the GWG and outcomes that are linked to gestational age such as birth weight, neonatal death among others (Hutcheon *et al.*, 2018).

In Hinkle *et al.* (2016), a logistic regression was used to estimate the relative risks (RR) of neonatal mortality predicted by total GWG. This was adjusted for gestational age, $z-$score, overall and within study sites. Directed acyclic graphs and simulation approach was used to assess any confounding attributed to gestational age and any potential biases from using total GWG, respectively. However, it was mentioned that z-score model is subject to reduced precision when $z-$chart is applied to external studies outside the underlying population.

Log-binomial regression models had been used to model GWG and pre-term-birth, small-for-gestational-age, large-for-gestational-age births while accounting for smoking, race, and education. It revealed a U-shape association between GWG z-score and pre-term birth but overestimated the relationship between rate of GWG, GWG adequacy ratio, and GWG and pre-term birth (Bodnar *et al.*, 2015). Bodnar *et al.* (2015) noted that the findings imply that analysis involving gestational

age-dependent outcomes misspecify the relationship if the three traditional measures are used. In spite of elimination of the potential for gestational age-related bias using z-score charts, they still have a disadvantage. The available z-scores charts are applicable to specific populations under study Johansson *et al.* (2016) and cannot be used for generalization.

Mitchell *et al.* (2016) investigated correlation between time to delivery and time varying covariates using Cox proportional hazards model. Since this approach requires daily weight gain measurements for the whole pregnancy period, Mitchell *et al.* (2016) based their analysis on a simulated data set of non-linear patterns of GWG, then compared the results with a binomial model having preterm birth as a dichotomized outcome. Although this analysis used the repeated measurements and improves precision, useful information is lost since a binary outcome is redefined to a continuous variable (Mitchell *et al.*, 2016). Karachaliou *et al.* (2015) used multiple and log Poisson regression models with continuous and binary outcomes respectively to examine the link between trimester-specific GWG with fetal growth, obesity risk, and cardio-metabolic health outcomes from childbirth to 4 years of age. The approach is straightforward but does not account for information on previous trimester weight gain trajectories (Hutcheon *et al.*, 2018).

Kleinman *et al.* (2007) compared area under the gestational weight gain curve, total GWG, rate of GWG, and Institute of Medicine categories to explore the prediction of birth weight and maternal weight retention at 6, 12, 24 and 36 months postpartum based on GWG. It was stated that area under the weight gain curve can be interpreted easily, reflects timing of weight gain better, and does not rely on assumptions about the shape of the weight gain curve. However, Hutcheon *et al.* (2018) argued that since the method depends on gestational duration, it may be less useful for analysis that involves adverse outcomes at earlier gestational ages.

A semiparametric, group-based, latent class, trajectory model was used to estimate the overall GWG and classify first- and second-/third-trimester trajectories to assess tracking, while a robust Poisson regression model was used to estimate the relative risk of SGA and LGA outcomes by the probability of trajectory membership (Pugh *et al.*, 2017). The method does not depend on any assumption, however the number of classes chosen by the model may not have a meaningful relationship with clinically important outcomes, since it is data driven (Hutcheon *et al.*, 2018).

A recent study on cohort of normal weight, overweight, and obese women was conducted to explore the feasibility of using non-linear mixed model called Super Imposition by Translation And Rotation (SITAR) to GWG trajectories (Riddell *et al.*, 2017). The model is suitable since it allows for summary of complex weight trajectories into three parameters; timing of growth, the acceleration of gain, and absolute amount of gain. The disadvantage is that the model has convergence issues when fitting all the three parameters, although the reduced models fit well (Riddell *et al.*, 2017).

Of all these methods discussed, a choice of the most applicable method depends on the context of the research question. However, analysis using repeated measurement of GWG is most suitable since the single measurement ignores the pattern in which weight was gained (Hutcheon *et al.*, 2018). In this study a multilevel modelling approach was considered because it accounts for correlation between a subject's repeated weight measurements, and allows each subject's trajectory to vary about the population average (Hutcheon *et al.*, 2018; Tuerlinckx *et al.*, 2006; Verbeke and Molenberghs, 2000).

3.4.2 Fractional Polynomials

Together with data exploration, it was necessary to further explore the functional form of gestational age in the linear mixed model. One way to represent the non-linear function of gestational age, is by use of the conventional polynomial models in which the exponents of GA are positive integers. However, this type of polynomials suffer from different limitations; low order polynomials offer only a few model shapes, while high order polynomials often do not fit the data well especially at the extremes of the observed range of the covariate. Moreover, these polynomials do not have asymptotes hence cannot represent the curvature in the data well. An extended family of curves known as Fractional Polynomials (FP) which was proposed by Royston and Altman (1994) and Royston and Sauerbrei (2008), can be used as an alternative since they provide flexible parameterization for continuous variables.

An FP model is a parametric approach which assumes that at each GA, the pregnancy weight measurement has a normal distribution with mean and standard deviation varying smoothly with GA (Ohuma *et al.*, 2019). The exponents of FPs are selected from a restricted set of both integer and non-integer values, $\{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$. A general form of a fractional polynomial of order m for covariate X can be expressed as:

$$FPm(X) = \beta_0 + \beta_1 X^{p_1} + \beta_2 X^{p_2} + \cdots + \beta_m X^{p_m} \quad \text{where, } p_1 \leq p_2 \leq \ldots \leq p_m,\ X^0 = ln(X)$$

The degree of an FP model above is defined by the number of powers of the covariate, X. FPs are suitable since they offer similar model fits as the conventional polynomials yet with few terms, have asymptotes, and offer a wide range of family of curves by allowing for non-integer powers, logarithms as well as repeated powers (Royston and Altman, 1994).

FP regressions with degree greater than 2 are less often used in practice, thus FP1 and FP2 were considered sufficient in this study. A total of eight and 32 possible models for respectively FP1 and FP2 were fitted for pregnancy weights as a function of gestational age based on maximum likelihood. Afterwards, the best fitting model for FP1 and FP2 was chosen using the deviance ($-2ln\lambda$). The deviance for straight line model was used as reference for the deviance of the FP1 and FP2 models. A larger deviance is preferred as it corresponds to a better fit of the model. A closed test procedure that ensures approximately correct Type I error rate for each test was applied to select the best fitting model of the FP1 and FP2 (Royston, 2017). The procedure is carried out in three steps that include overall association test, non-linearity test, and test between a simpler and more complex FP model (Benner, 2014; Royston, 2017). The steps are as follows.

(i). First, test the best fitting FP2 against the null model using likelihood ratio test with χ^2_4 at $\alpha = 5\%$ level of significance. If significant, move to the next step, otherwise stop with the chosen model being null model.

(ii). Next, repeat the same test using χ^2_3 at $\alpha = 5\%$ level of significance of the best fitting FP2 against linear model. If significant, continue to the final step, otherwise stop and conclude straight line as the final model.

(iii). Finally, test the best fitting FP2 against the best fitting FP1 based on χ^2_2 at $\alpha = 5\%$ level of significance. If significant, the final model is FP2, otherwise FP1.

3.4.3 Linear Mixed Model

Multiple measurements on weight are obtained per expectant mother during the routine prenatal visits. This implies that measurements obtained from one pregnant individual are correlated, but

measurements between individuals are assumed to be independent. A linear mixed model (LMM) that takes into account the correlation between repeated measurements within subjects and between-subject variability due to clustering, is called a mixed effects model (Verbeke and Molenberghs, 2000).

A linear mixed model is an extension of a simple linear model that incorporates both fixed and random effects (Verbeke and Molenberghs, 2000). The fixed effects model average evolution, while random effects model the extent to which the evolution vary across levels of each subject. LMM assumes that a continuous response variable is linearly related to a set of explanatory variables. In this study, a model having both random slope and intercept was proposed to model the pregnancy weights as a linear function of gestational age, whereby the slope depends on the gestational age of a woman. The model was formulated as follows.

$$Y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})GA_{ij} + \epsilon_{ij}, \qquad \boldsymbol{b_i} = (b_{0i},\ b_{1i})' \sim N(0, D), \quad \boldsymbol{\epsilon_i} \sim N(0, \Sigma_i)$$

where Y_{ij} is the measurement of weight in $kilograms$ for subject i recorded at visit j (at time GA_{ij}, gestational age in weeks), $i = 1, \ldots, 2820,\ j = 1, \ldots, n_i$, $\boldsymbol{Y_i} = (Y_{i1}, Y_{i2}, \ldots, Y_{in_i})'$ is an n_i-dimensional vector of all the repeated Y_{ij}'s, β_0 and β_1 represent the vectors of global-average regression coefficients, b_{0i} and b_{1i} are vectors of woman-specific regression coefficients which denote the deviations of weight of the i^{th} subject from the global-average weight, ϵ_{ij} is an error term, D and Σ_i are $(n_i \times n_i)$ co-variance matrices which depend on i only through their dimensions n_i, with the elements of $\boldsymbol{b_i}$ and $\boldsymbol{\epsilon_i}$ assumed independent.

3.4.4 MLE versus REML Estimation procedure

The model parameters to be estimated include the vector of fixed effects, $\boldsymbol{\beta}$, subject-specific random effects, $\boldsymbol{b_i}$, and the co-variance parameters, D and Σ_i. Since $\boldsymbol{b_i}$ and $\boldsymbol{\epsilon_i}$ are assumed to be independent and distributed multivariate normal, maximum likelihood (ML) estimation method can be used to estimate the parameters in the model. However, the maximum likelihood estimates are biased having smaller variances. This is attributed to ML estimation technique considers the fixed effects as unknown when estimating variance parameters, yet it does not account for the degrees of freedom lost by estimating the fixed effects. In this study, linear mixed model was fitted under restricted maximum likelihood (REML) method since it accounts for the degrees of freedom lost when estimating the vector of fixed effects. This produces unbiased variance estimates. The procedure ensures that the fixed effects are not contained in the log-likelihood function when estimating the variance parameters but estimates them in the second step based on a posterior distribution having obtained the unbiased variance component estimates. The limitation of using REML is that it only allows for comparison of nested models with similar mean structures.

3.4.5 Test for the need of Random slope

To assess the need of having both random effects in the model, a model with and without random slope were fitted based on restricted maximum likelihood estimation method. Since under the hierarchical interpretation of the model, H_0 is on the boundary of the parameter space, that is, hierarchical models cannot allow for negative variance components (D matrix should be positive semi definite), the classical likelihood inference based on a single χ^2 distribution cannot be used (Verbeke and Molenberghs, 2000). An asymptotic null-distribution of the likelihood ratio test statistic, $-2ln\lambda_N$ under a mixture of χ^2_1 and χ^2_2 with equal weights of 0.5 was used to test for the effect of

the random slope based on the following hypothesis.

$$H_0 : D = \begin{pmatrix} d_{11} & 0 \\ 0 & 0 \end{pmatrix}$$

3.4.6 Generalized Linear Model

Generalized linear models (GLMs) are extension of the classical linear regression framework that allow the response variable to have a non-normal error distribution mainly from the exponential family, given the values of the explanatory variables in the model. GLMs relate the linear predictor to the predicted value of the response via a link function, and variance of each observation as a function of its predicted value via variance function. In this study, a family of a GLM known as log-binomial model was considered for analysis.

To assess the association between the maternal/ neonatal outcomes with other covariates, relative risk (RR) was preferred over odds ratio (OR). RR refers to the ratio of the probability of an event in an exposed group to the probability of the event in an unexposed group, $\left(\frac{p_1}{p_0}\right)$. On the other hand, OR expresses the ratio of odds of an outcome in an exposed group to the odds of an outcome in the non-exposed group, $\left(\frac{p_1/(1-p_1)}{p_0/(1-p_0)}\right)$. Both are measures of association for binary data, however, OR is commonly reported in literature and is usually misinterpreted as RR (Viera, 2008). This can be explained by the ease of obtaining ORs from fitted logistic regression model based on "canonical link" function in most standard statistical softwares. OR gives a good approximation of RR if the event is rare (less than 10% of the unexposed group), especially in case-control studies, unfortunately it is not appropriate when events are common as it exaggerates RR, mainly in cohort and cross-sectional studies (Viera, 2008; Schmidt and Kohlmann, 2008). RR is advantageous since it has an intuitive interpretation compared to OR, and it can also be applied to groups with different prevalence of an event.

In this study, the model of focus was log-binomial model using "log" link function since it naturally generates relative risks estimates, when studying the association of a set of predictors on a single binary outcome. RR takes any non-negative real number, though log-relative risk is mostly used to alleviate that restriction. A limitation that may arise when fitting this kind of model is failure to converge or yielding inadmissible fitted values. Unlike logistic regression model, which uses "logit" link function to ensure that the fitted probabilities are within $[0, 1]$, "log" link function allows probabilities beyond this boundary.

Let Y be a binary random outcome with sample space $\{0, 1\}$, $\boldsymbol{X}^T = (1, x_1, x_2, \ldots, x_p)$ be a set of covariates. It follows that, the log-binomial model assumes the probability of Y equals to 1 given k covariates. This can be formulated as:

$$log[P(Y = 1|\boldsymbol{X})] = \beta^T \boldsymbol{X} \qquad \text{for } \beta^T \boldsymbol{X} \leq 0$$

where $\beta^T = (\beta_0, \beta_1, \beta_2, \ldots, \beta_p)$ is a $(p = k + 1)$ vector of parameters to be estimated. Exponentiating the coefficient vector, β^T results to adjusted risk ratios. For instance, RR of $\boldsymbol{X}_1$ compared to $\boldsymbol{X}_0$ is denoted by:

$$\frac{P(Y = 1|\boldsymbol{X}_1)}{P(Y = 1|\boldsymbol{X}_0)} = \frac{e^{\beta^T \boldsymbol{X}_1}}{e^{\beta^T \boldsymbol{X}_0}} = e^{\beta^T (\boldsymbol{X}_1 - \boldsymbol{X}_0)}$$

The maximum likelihood estimate of β^T can be obtained from maximizing the logarithm of likelihood function of the log binomial regression model above.

$$\max_{\beta^T} l(\beta^T) = \sum_{i=1}^{n} y_i \beta^T \boldsymbol{X}_i + (1 - y_i) log(1 - e^{\beta^T \boldsymbol{X}_i}) \quad \text{subject to } \beta^T \boldsymbol{X}_i \leq 0 \; \forall \; 1 \leq i \leq n.$$

where y_i is the observed outcome, and $\boldsymbol{X}_i^T = (1, x_{i1}, x_{i2}, \ldots, x_{ip})$ is a $(p = k + 1)$ vector of the observed covariates. The constraint ensures that $P(Y = 1|\boldsymbol{X}) = e^{\beta^T \boldsymbol{X}} \leq 1$.

Fitting log-binomial regression using glm function in R statistical software generates the MLE of β^T based on an algorithm known as iteratively reweighted least squares (IRLS), which involves solving a weighted least squares problem in each iteration. To ensure that the estimates are bounded within the restricted parameter space, a criteria called step-halving is implemented within the IRLS procedure. This follows that in case a positive estimate is realized during the fitting process, the update is halved repeatedly until the value is recomputed. If it is still positive, the updated value is halved again until the fitted value is negative. However, according to different authors, this procedure sometimes fail to converge even so the value is within the parameter space (Williamson *et al.*, 2013; Schwendinger *et al.*, 2021). The formulated log-binomial model for this study was as follows.

$$log\Big[P(Y_{ij} = 1 \mid b_{0i}, b_{1i}, \boldsymbol{Z_i})\Big] = \beta_0 + \beta_1 \hat{b}_{0i} + \beta_2 \hat{b}_{1i} + \boldsymbol{Z_i \beta_i}$$

where $Y_{ij} \sim Bin(n, p)$ is a binary variable for the maternal/ neonatal outcome of subject i recorded at delivery time j, b_{0i} and b_{1i} are random effects of the linear mixed model, and $\boldsymbol{Z_i}$; other covariates (maternal characteristics) adjusted in the model.

In case non-convergence occurred while fitting the model, a Poisson regression model with a robust variance estimator was used as an approximation (Carter *et al.*, 2005). The robust sandwich estimation was considered to obtain valid confidence intervals since Poisson model exaggerates standard errors. One limitation worth noting is that Poisson regression does not guarantee that the resulting estimates are bound within the parameter space for the log-binomial model.

3.5 Software

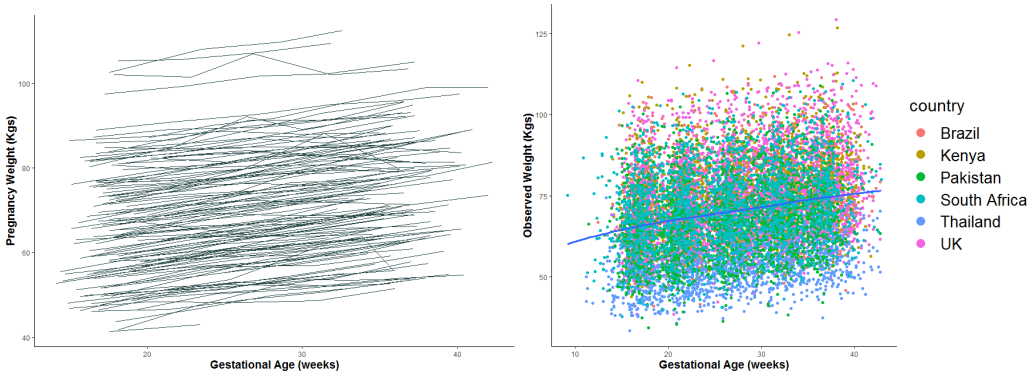
Analyses were performed in R statistical software version $4.0.3$ $(2020 - 10 - 10)$ (R Core Team, 2020), and SAS software version 9.4 (SAS Institute Inc., 2013). The $lmer$ function in the $lme4$ package was used to fit the linear mixed model, whereas the glm function in the $stats$ package was applied to fit the generalized linear models based on log link function. SAS software All the analyses were conducted under 5% significance level.

4 Results

4.1 Exploratory Data Analysis

Individuals profiles plot was used to visualize the general pattern of evolution of weight of each expectant mother within the gestation period. Since the study involved a large data set, the plot was fitted for 120 randomly selected individuals as illustrated in Figure 2a. The plot suggested almost a linear relation between gestational age (GA) and the observed weights of each subject, with less and more variability observed, respectively within and between individuals. This implied that a model with both random intercept and random slope would be suitable for analysis. It was also noted that measurements were taken at different time points and some subjects had incomplete profiles due to different enrollment and delivery times.

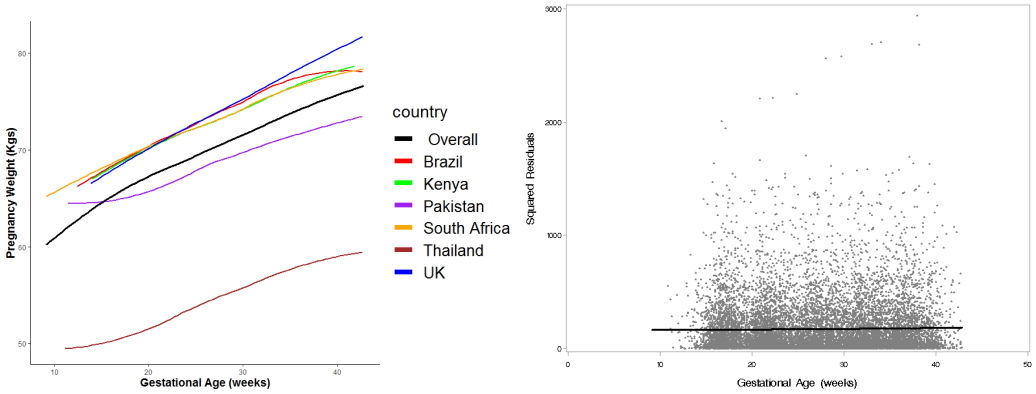
Due to the unbalanced nature of the data set, LOESS smoothing technique was used to give an overview of the mean structure for the linear mixed model. Figure 2b displays the plot constructed for mean evolution of weight versus the gestational age, also stratified by country in Figure 3a. The smoothed mean curve showed almost a linear average trend in weight over GA, suggesting that a model with GA in linear form might be adequate. In addition, Figure 3a illustrated that Thailand had lower values of recorded weights during the gestation period compared to other countries.



(a) Individual profiles plot **(b)** Mean evolution plot

Figure 2: (a) Individual profiles plot of weight versus gestational age for 120 randomly selected women. (b) Mean evolution plot using LOESS smoothing technique to investigate the evolution of the mean structure of weight versus gestational age.

It was useful to explore the evolution of the variance structure of the data set in order to choose an appropriate covariance structure to describe the variability in the data when fitting a random effects model. Figure 3b displays a plot of squared ordinary least squares residuals constructed to visualize the average evolution of the variance of weight as a function of GA. It was observed that the variance function was relatively constant, hence a model with constant variance was suitable.



(a) Mean Curve **(b)** Squared Residuals

Figure 3: Mean and variance structures for weight as a function of gestational age. (a) Mean evolution plots using LOESS smoothing technique to investigate the evolution of the mean structure of weight versus gestational age, overall and stratified by country. (b) Squared residuals of the smoothed mean curve against the gestational age using LOESS smoothing technique.

The total number of women recorded at the first prenatal visit for each country and overall, together with their corresponding mean, median, minimum, and maximum values of observed weight and GA was displayed in Table 3. The global average of pregnancy weight and GA recorded at the first visit ranged between $33.3 - 110.5\ kgs$ and $9.14 - 34.57$ weeks, respectively. There was an unequal distribution of the number of women recruited in the study, having the highest number from the UK and the lowest from Brazil. The summary statistics (Table 3) showed that the average starting pregnancy weight was approximately equal for all the countries except for Thailand. Similarly, the average starting gestational age for women in all the countries was almost the same, but have different range values.

Table 3: Summary statistics of gestational age and weight at the first prenatal visit by Country

Country/ Variable	# Women	Mean (sd)	Median	Minimum	Maximum	Mean (sd)	Median	Minimum	Maximum
Weight (Kgs)						**Gest. Age (weeks)**			
1. Brazil	349	68.98 (11.58)	67.70	41.0	105.3	17.44 (1.56)	17.43	12.43	26.43
2. Kenya	547	68.76 (11.66)	67.80	43.8	110.0	17.30 (1.42)	17.29	14	33.57
3. Pakistan	492	64.71 (11.54)	63.85	34.3	100.2	17.82 (2.2)	17.29	11.43	28.14
4. South Africa	455	68.93 (11.78)	68.40	37.8	102.4	17.22 (2.45)	17.14	9.14	30.43
5. Thailand	354	50.28 (7.48)	49.65	33.3	76.0	17.38 (2.67)	17.14	11.14	34.57
6. UK	623	68.60 (12.19)	66.40	38.4	110.5	17.34 (1.59)	17	13.86	27.86
Overall	**2820**	**65.75 (12.84)**	**64.9**	**33.3**	**110.5**	**17.41 (2.0)**	**17.14**	**9.14**	**34.57**

sd - standard deviation. **#** - Number of.

Table 4 shows the summary statistics of maternal characteristics for the continuous variables measured during the recruitment time. Women recruited in the study were on average between $18 - 50$ years old, with an average BMI of $24.88\ kg/m^2$, $64.17\ kgs$ weight , $111.33\ mmHg$ systolic blood pressure, $70.24\ mmHg$ diastolic blood pressure, $160.37\ cm$ height, and 13.20 years of formal education. Values of other summary measures were approximately the same except for Thailand having recorded low values for both weight and number of years of formal education.

Table 4: Summary statistics for continuous maternal anthropometric measurements

Var./ Country	Mean (sd)	Min	Max	Median	Mean	Min	Max	Median	Mean	Min	Max	Median
Age (years)					**BMI**				**Weight (Kgs)**			
Brazil	29.1 (5.30)	18.54	45.92	29.51	25.50	15.8	35.5	25.1	67.45	38.9	104.8	66.1
Kenya	30.89 (4.08)	19.65	42.78	30.84	25.59	17.1	35.9	25.4	67.06	41.6	106.6	66.2
Pakistan	30.50 (4.55)	18.56	43.12	30.32	25.28	14.8	34.9	25.0	63.15	34.7	92.9	62.6
South Africa	31.47 (5.79)	19.25	44.70	31.39	26.61	15.7	34.9	26.6	67.46	37.2	101.5	66.5
Thailand	26.74 (6.01)	18.03	44.94	25.67	21.12	12.4	38.0	20.8	48.86	33.0	75.0	48.2
UK	31.59 (4.74)	18.56	44.0	32.05	24.44	15.5	38.0	23.7	66.89	36.5	109.2	64.7
Overall	**30.45 (5.18)**	**18.03**	**45.92**	**30.58**	**24.88**	**12.4**	**38**	**24.5**	**64.17**	**33**	**109.2**	**63.3**
Systolic BP					**Diastolic BP**				**Height (cm)**			
Brazil	114.76	88	155	114	72.08	52	97	72	162.49	147.3	196.9	162.2
Kenya	109.67	68	149	110	69.35	44	109	69	161.79	143.5	176.5	162.0
Pakistan	106.59	80	158	107	69.46	44	92	70	158.02	137.1	184.0	158.0
South Africa	114.62	84	168	114	72.31	39	137	71	159.19	141.8	193.6	159.3
Thailand	104.14	90	150	100	65.93	50	100	65	152.14	152.6	114.1	182.2
UK	116.29	82	162	115	71.54	39	103	71	165.34	141.9	186.5	165.3
Overall	**111.33**	**68**	**168**	**110**	**70.24**	**39**	**137**	**70**	**160.37**	**114.1**	**196.9**	**160.3**
Number years	**of formal**	**education**										
Brazil	12.28	1	22	11								
Kenya	15.78	8	20	16								
Pakistan	14.74	0	27	16								
South Africa	11.77	0	18	12								
Thailand	4.99	0	30	5								
UK	15.96	5	28	16								
Overall	**13.20**	**0**	**30**	**14**								

The frequencies of the categorical variables describing maternal characteristics stratified by country were shown in Table 5. It was noted that most study participants were well educated and the highest number (42.45%) had obtained a university degree. Most participants were non-students, especially no student was recruited from Thailand. Additionally, a greater percentage of women were either married or cohabiting (86.31%), neither smoked (96.17%) nor took alcohol (97.84%). South Africa recorded the highest number of women who had previous pregnancies compared to other countries. High number of women had missing values for the previous preterm birth, neonatal deaths, and cesarean section. It was noted that these values were missing by design, meaning that they did not have any previous pregnancies, except for few values ($< 5\%$) that were completely not recorded.

Table 5: Summary statistics for categorical maternal anthropometric measurements

Variable	Frequency (%)						
	Brazil	**Kenya**	**Pakistan**	**South Africa**	**Thailand**	**UK**	**Overall**
Level of Education							
No school attended	0 (0)	0 (0)	3 (0.61)	1 (0.22)	76 (21.47)	0 (0)	80 (2.84)
Primary	48 (13.75)	2 (0.37)	17 (3.46)	15 (3.3)	151 (42.66)	0 (0)	233 (8.26)
Secondary	144 (41.26)	5 (0.91)	84 (17.07)	330 (72.53)	113 (31.92)	215 (34.51)	419 (14.86)
Professional/Technical training	23 (6.59)	179 (32.72)	21 (4.27)	77 (16.92)	5 (1.41)	114 (18.30)	891 (31.6)
University	134 (38.40)	361 (66.0)	367 (74.59)	32 (7.03)	9 (2.54)	294 (47.19)	1197 (42.45)
Occupational Status							
Clerical support, service or sales	104 (29.80)	59 (10.79)	0 (0)	64 (14.07)	3 (0.85)	197 (31.62)	427 (15.14)
Housework	83 (23.78)	18 (3.29)	321 (65.24)	47 (10.33)	232 (65.54)	94 (15.09)	795 (28.19)
Managerial/Professional/Technical	75 (21.49)	424 (77.51)	125 (25.41)	14 (3.08)	13 (3.67)	293 (47.03)	944 (33.48)
Skilled manual work	19 (5.44)	11 (2.01)	34 (6.91)	38 (8.35)	31 (8.76)	6 (0.96)	139 (4.93)
Student	19 (5.44)	25 (4.57)	7 (1.42)	18 (3.96)	0 (0)	16 (2.57)	85 (3.01)
Unskilled manual work	4 (1.15)	0 (0)	2 (0.41)	48 (10.55)	67 (18.93)	16 (2.57)	137 (4.86)
Other	45 (12.89)	10 (1.83)	3 (0.61)	226 (49.67)	8 (2.26)	1 (0.16)	293 (10.39)
Marital Status							
Married/Cohabiting	318 (91.12)	497 (90.86)	490 (99.59)	182 (40)	348 (98.31)	599 (96.15)	2434 (86.31)
Separated/Divorced	1 (0.29)	0 (0)	1 (0.20)	2 (0.44)	0 (0)	2 (0.32)	6 (0.21)
Single	29 (8.31)	50 (9.14)	1 (0.20)	271 (59.56)	4 (1.13)	22 (3.53)	377 (13.37)
Widowed	1 (0.29)	0 (0)	0 (0)	0 (0)	2 (0.56)	0 (0)	3 (0.11)
BMI Category							
Normal weight	163 (46.70)	239 (43.69)	217 (44.11)	159 (34.95)	255 (72.03)	365 (58.59)	1398 (49.57)
Underweight	5 (1.43)	11 (2.01)	24 (4.88)	9 (1.98)	62 (17.51)	16 (2.57)	127 (4.5)
Overweight	128 (36.68)	209 (38.21)	177 (35.98)	174 (38.24)	33 (9.32)	169 (27.13)	890 (31.56)
Obese	53 (15.19)	87 (15.90)	74 (15.04)	113 (24.84)	4 (1.13)	73 (11.72)	404 (14.33)
-	0 (0)	1 (0.18)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.04)
Alcohol intake							
No	339 (97.13)	545 (99.63)	492 (100)	421 (92.53)	353 (99.72)	609 (97.75)	2759 (97.84)
Yes	10 (2.87)	2 (0.37)	0 (0)	34 (7.47)	1 (0.28)	14 (2.25)	61 (2.16)
Smoking Status							
No	328 (93.98)	547 (100)	491 (99.80)	431 (94.73)	332 (93.79)	583 (93.58)	2712 (96.17)
Yes	21 (6.02)	0 (0)	1 (0.20)	24 (5.27)	22 (6.21)	40 (6.42)	108 (3.83)
Previous Pregnancy							
No	204 (58.45)	213 (38.94)	121 (24.59)	29 (6.37)	114 (32.20)	190 (30.50)	871 (30.89)
Yes	145 (41.55)	334 (61.06)	371 (75.41)	426 (93.63)	240 (67.80)	433 (69.50)	1949 (69.11)
Previous preterm birth							
No	108 (30.95)	267 (48.81)	240 (48.78)	273 (60)	195 (55.08)	300 (48.15)	1383 (49.04)
Yes	14 (4.01)	14 (2.56)	76 (15.45)	83 (18.24)	24 (6.78)	64 (10.27)	275 (9.75)
NA	227 (65.04)	266 (48.63)	176 (35.77)	99 (21.76)	135 (38.14)	259 (41.57)	1162 (41.21)
Previous neonatal deaths							
No	118 (33.81)	274 (50.09)	264 (53.66)	328 (72.09)	207 (58.47)	361 (57.95)	1552 (55.04)
Yes	2 (0.53)	7 (1.28)	52 (10.57)	28 (6.15)	12 (3.39)	3 (0.48)	104 (3.69)
NA	229 (65.62)	266 (48.63)	176 (35.77)	99 (21.76)	135 (38.14)	259 (41.57)	1164 (41.28)
Previous CS							
No	249 (71.35)	202 (36.93)	224 (45.53)	177 (38.90)	30 (8.47)	196 (31.46)	1078 (38.23)
Yes	34 (9.74)	79 (14.44)	152 (30.89)	111 (24.40)	3 (0.85)	40 (6.42)	419 (14.86)
NA	66 (18.91)	266 (48.63)	116 (23.58)	167 (36.70)	321 (90.68)	387 (62.12)	1323 (46.91)

Table 6 presents frequencies and percentages of maternal outcomes stratified by country. It can be observed that few participants experienced adverse outcomes including 3.72% hypertensive, 5.78% having pregnancy induced hypertension, 6.60% gestational diabetic, 1.13% with preeclampsia, 0.25% having haemolysis, elevated liver enzymes, low platelet count (HELLP) preeclampsia, 3.55% underwent preterm labour, 41.74% had ceaserian section, 0.74% had vaginal bleeding, 0.43% placenta praevia, 2.52% breech precentation, 0.07% uterine rapture, 1.28% reduced fetal movement, 7.66% failure to progress, 1.67% cephalopelvic disproportion, 3.72% with preterm premature rupture of the membranes (PPROM), 0.50% placenta abruptio, and 0.53% extremely preterm birth. Greater number of missing values recorded were by design, that is, that some of the maternal outcomes did not apply to women who had vaginal spontaneous delivery.

Table 6: Frequencies of Maternal outcomes

Variable	Frequency (%)						
	Brazil	**Kenya**	**Pakistan**	**South Africa**	**Thailand**	**UK**	**Overall**
Hypertension							
No	336 (96.28)	543 (99.27)	463 (94.11)	429 (94.29)	343 (96.89)	601 (96.47)	2715 (96.28)
Yes	13 (3.72)	4 (0.73)	29 (5.89)	26 (5.71)	11 (3.11)	22 (3.53)	105 (3.72)
Pregnancy Induced Hypertension							
No	327 (93.70)	539 (98.54)	459 (93.29)	427 (93.85)	332 (93.79)	573 (91.97)	2657 (94.22)
Yes	22 (6.30)	8 (1.46)	33 (6.71)	28 (6.15)	22 (6.21)	50 (8.03)	163 (5.78)
Gestational Diabetes							
No	322 (92.26)	535 (97.81)	373 (75.81)	451 (99.12)	349 (98.59)	604 (96.95)	2634 (93.40)
Yes	27 (7.74)	12 (2.19)	119 (24.19)	4 (0.88)	5 (1.41)	19 (3.05)	186 (6.60)
Preeclampsia							
No	349 (100)	539 (98.54)	483 (98.17)	451 (99.12)	353 (99.72)	612 (98.23)	2787 (98.83)
Yes	0 (0)	8 (1.46)	9 (1.83)	3 (0.66)	1 (0.28)	11 (1.77)	32 (1.13)
NA	0 (0)	0 (0)	0 (0)	1 (0.22)	0 (0)	0 (0)	1 (0.04)
Hellp Preeclampsia							
No	349 (100)	547 (100)	489 (99.39)	454 (99.78)	352 (99.44)	622 (99.84)	2813 (99.75)
Yes	0 (0)	0 (0)	3 (0.61)	1 (0.22)	2 (0.56)	1 (0.16)	7 (0.25)
Preterm Labour							
No	305 (87.39)	542 (99.09)	475 (96.54)	452 (99.34)	342 (96.61)	604 (96.95)	2720 (96.45)
Yes	44 (12.61)	5 (0.91)	17 (3.46)	3 (0.66)	12 (3.39)	19 (3.05)	100 (3.55)
Mode of delivery							
Caesarean section	272 (77.94)	208 (38.03)	296 (60.16)	272 (59.78)	15 (4.24)	114 (18.30)	1177 (41.74)
Vaginal assisted	3 (0.86)	15 (2.74)	15 (3.05)	2 (0.44)	10 (2.82)	103 (16.53)	148 (5.25)
Vaginal spontaneous	74 (21.20)	324 (59.23)	181 (36.79)	181 (39.78)	326 (92.09)	405 (65.01)	1491 (52.87)
Assisted breech	0 (0)	0 (0)	0 (0)	0 (0)	3 (0.85)	1 (0.16)	4 (0.14)
Vaginal bleeding							
No	281 (80.52)	277 (50.64)	372 (75.61)	284 (62.42)	32 (9.04)	230 (36.92)	1476 (52.34)
Yes	2 (0.57)	4 (0.73)	4 (0.81)	4 (0.88)	1 (0.28)	6 (0.96)	21 (0.74)
NA	66 (18.91)	266 (48.63)	116 (23.58)	167 (36.70)	321 (90.68)	387 (62.12)	1323 (46.91)
Placenta praevia							
No	283 (81.09)	279 (51.01)	374 (76.02)	285 (62.64)	33 (9.32)	231 (37.08)	1485 (52.66)
Yes	0 (0)	2 (0.37)	2 (0.41)	3 (0.66)	0	5 (0.80)	12 (0.43)
NA	66 (18.91)	266 (48.63)	116 (23.58)	167 (36.70)	321 (90.68)	387 (62.12)	1323 (46.91)
Breech presentation							
No	272 (77.94)	270 (49.36)	356 (72.36)	278 (61.10)	29 (8.19)	221 (35.47)	1426 (50.57)
Yes	11 (3.15)	11 (2.01)	20 (4.07)	10 (2.20)	4 (1.13)	15 (2.41)	71 (2.52)
NA	66 (18.91)	266 (48.63)	116 (23.58)	167 (36.70)	321 (90.68)	387 (62.12)	1323 (46.91)
Uterine rupture							
No	283 (81.09)	281 (51.37)	376 (76.42)	287 (63.08)	33 (9.32)	235 (37.72)	1495 (53.01)
Yes	0 (0)	0 (0)	0 (0)	1 (0.22)	0	1 (0.16)	2 (0.07)
NA	66 (18.91)	266 (48.63)	116 (23.58)	167 (36.70)	321 (90.68)	387 (62.12)	1323 (46.91)
Reduced fetal movement							
No	278 (79.66)	265 (48.45)	367 (74.59)	287 (63.08)	32 (9.04)	232 (37.24)	1461 (51.81)
Yes	5 (1.43)	16 (2.93)	9 (1.83)	1 (0.22)	1 (0.28)	4 (0.64)	36 (1.28)
NA	66 (18.91)	266 (48.63)	116 (23.58)	167 (36.70)	321 (90.68)	387 (62.12)	1323 (46.91)
Failure to progress							
No	253 (72.49)	223 (40.77)	326 (66.26)	247 (54.29)	26 (7.34)	206 (33.07)	1281 (45.43)
Yes	30 (8.60)	58 (10.60)	50 (10.16)	41 (9.01)	7 (1.98)	30 (4.82)	216 (7.66)
NA	66 (18.91)	266 (48.63)	116 (23.58)	167 (36.70)	321 (90.68)	387 (62.12)	1323 (46.91)
Cephalopelvic disproportion							
No	256 (73.35)	279 (51.01)	372 (75.61)	279 (61.32)	29 (8.19)	235 (37.72)	1450 (51.42)
Yes	27 (7.74)	2 (0.37)	4 (0.81)	9 (1.98)	4 (1.13)	1 (0.16)	47 (1.67)
NA	66 (18.91)	266 (48.63)	116 (23.58)	167 (36.70)	321 (90.68)	387 (62.12)	1323 (46.91)
PPROM							
No	255 (73.07)	252 (46.07)	372 (75.61)	281 (61.76)	27 (7.63)	205 (32.91)	1392 (49.36)
Yes	28 (8.02)	29 (5.30)	4 (0.81)	7 (1.54)	6 (1.69)	31 (4.98)	105 (3.72)
NA	66 (18.91)	266 (48.63)	116 (23.58)	167 (36.70)	321 (90.68)	387 (62.12)	1323 (46.91)
Placental abruptio							
No	282 (80.80)	278 (50.82)	372 (75.61)	287 (63.08)	32 (9.04)	232 (37.24)	1483 (52.59)
Yes	1 (0.29)	3 (0.55)	4 (0.81)	1 (0.22)	1 (0.28)	4 (0.64)	14 (0.50)
NA	66 (18.91)	266 (48.63)	116 (23.58)	167 (36.70)	321 (90.68)	387 (62.12)	1323 (46.91)
Gestational Age category							
Full term	310 (88.83)	517 (94.52)	397 (80.69)	348 (76.48)	330 (93.22)	577 (92.62)	2479 (87.91)
Moderate to late preterm	38 (10.89)	27 (4.94)	87 (17.68)	86 (18.90)	20 (5.65)	40 (6.42)	298 (10.57)
Very preterm	1 (0.29)	3 (0.55)	7 (1.42)	13 (2.86)	1 (0.28)	3 (0.48)	28 (0.99)
Extremely preterm	0 (0)	0 (0)	1 (0.20)	8 (1.76)	3 (0.85)	3 (0.48)	15 (0.53)

The frequencies with their corresponding percentages for the neonatal outcomes stratified by country, were displayed in Table 7. Unlike the maternal outcomes, the number of babies who experienced adverse outcomes was relatively high. About 12.09% were preterm birth, 0.11% fetal anaemic, 0.07% fetal deaths, 7.52% fetal distress cases, 11.49% taken to the newborn intensive special care unit, 0.14% had seizures, 9.79% small for gestational age newborns, 7.84% large for gestational age newborns, 7.59% had head circumference below C10, 12.41% head circumference above C90, 11.06% with body length below C10, and 8.58% body length above C90. Observed missing values applied to the women who had vaginal spontaneous delivery.

Table 7: Frequencies of Neonatal outcomes

Variable	Frequency (%)						
	Brazil	**Kenya**	**Pakistan**	**South Africa**	**Thailand**	**UK**	**Overall**
Fetal Anaemia							
No	349 (100)	546 (99.82)	492 (100)	453 (99.56)	354 (100)	623 (100)	2817 (99.89)
Yes	0 (0)	1 (0.18)	0 (0)	2 (0.44)	0 (0)	0 (0)	3 (0.11)
Preterm birth							
Fullterm	310 (88.83)	517 (94.52)	397 (80.69)	348 (76.48)	330 (93.22)	577 (92.62)	2479 (87.91)
Preterm	39 (11.17)	30 (5.48)	95 (19.31)	107 (23.52)	24 (6.78)	46 (7.38)	341 (12.09)
Fetal death							
No	283 (81.09)	281 (51.37)	375 (76.22)	287 (63.08)	33 (9.32)	236 (37.88)	1495 (53.01)
Yes	0 (0)	0 (0)	1 (0.20)	1 (0.22)	0 (0)	0 (0)	2 (0.07)
NA	66 (18.91)	266 (48.63)	116 (23.58)	167 (36.70)	321 (90.68)	387 (62.12)	1323 (46.91)
Fetal distress							
No	268 (76.79)	235 (42.96)	354 (71.95)	190 (41.76)	33 (9.32)	205 (32.91)	1285 (45.57)
Yes	15 (4.30)	46 (8.41)	22 (4.47)	98 (21.54)	0 (0)	31 (4.98)	212 (7.52)
NA	66 (18.91)	266 (48.63)	116 (23.58)	167 (36.70)	321 (90.68)	387 (62.12)	1323 (46.91)
Newborn intensive special care							
No	317 (90.83)	506 (92.50)	459 (93.29)	367 (80.66)	286 (80.79)	561 (90.05)	2496 (88.51)
Yes	32 (9.17)	41 (7.50)	33 (6.71)	88 (19.34)	68 (19.21)	62 (9.95)	324 (11.49)
Seizures							
No	349 (100)	547 (100)	489 (99.39)	454 (99.78)	353 (99.72)	621 (99.68)	2813 (99.75)
Yes	0 (0)	0 (0)	1 (0.20)	0 (0)	1 (0.28)	2 (0.32)	4 (0.14)
NA	0 (0)	0 (0)	2 (0.41)	1 (0.22)	0 (0)	0 (0)	3 (0.11)
Below C10 BW							
No	319 (91.40)	507 (92.69)	443 (90.04)	417 (91.65)	279 (78.81)	579 (92.94)	2544 (90.21)
Yes	30 (8.60)	40 (7.31)	49 (9.96)	38 (8.35)	75 (21.19)	44 (7.06)	276 (9.79)
Above C90 BW							
No	322 (92.26)	503 (91.96)	479 (97.36)	411 (90.33)	346 (97.74)	538 (86.36)	2599 (92.16)
Yes	27 (7.74)	44 (8.04)	13 (2.6)	44 (9.67)	8 (2.26)	85 (13.64)	221 (7.84)
Below C10 HC							
No	334 (95.70)	528 (96.53)	459 (93.29)	441 (96.92)	251 (70.90)	593 (95.18)	2606 (92.41)
Yes	15 (4.30)	19 (3.47)	33 (6.71)	14 (3.08)	103 (29.10)	30 (4.82)	214 (7.59)
Above C90 HC							
No	305 (87.39)	457 (83.55)	456 (92.68)	375 (82.42)	347 (98.02)	530 (85.07)	2470 (87.59)
Yes	44 (12.61)	90 (16.45)	36 (7.32)	80 (17.58)	7 (1.98)	93 (14.93)	350 (12.41)
Below C10 BL							
No	309 (88.54)	510 (93.24)	447 (90.85)	400 (87.91)	273 (77.12)	569 (91.33)	2508 (88.94)
Yes	40 (11.46)	37 (6.76)	45 (9.15)	55 (12.09)	81 (22.88)	54 (8.67)	312 (11.06)
Above C90 BL							
No	325 (93.12)	502 (91.77)	450 (91.46)	408 (89.67)	335 (94.63)	558 (89.57)	2578 (91.42)
Yes	24 (6.88)	45 (8.23)	42 (8.54)	47 (10.33)	19 (5.37)	65 (10.43)	242 (8.58)

4.2 Fractional Polynomials

Based on the deviance differences from Tables 16 and 17 in the Appendix, the best fitting FP1 and FP2 models together with their Akaike information criterion (AIC) and Bayesian information criterion (BIC) values were displayed in Table 8. To select the final FP model from the two, a closed test procedure was performed as outlined in Section 3.4.2. The results were presented in Table 9.

Table 8: Best fitting Fractional Polynomials

Model	Power	Mean Structure	R^2	AIC	BIC	-2 Log Lik (Residual Deviance)
FP1	0.5	$GA^{0.5}$	0.06	106794.5	106817	106788.5
FP2	(-2, 2)	$GA^{-2} + GA^2$	0.06	106794.6	106824.6	106786.6

The likelihood ratio test of the best fitting FP2 against the null model was statistically significant ($p < 0.0001$), suggesting that test for non-linearity was necessary. Thus, a second likelihood ratio test was conducted for the best FP2 versus the linear model. The test was not statistically significant ($p = 0.3329$) at $\alpha = 5\%$ level of significance, concluding that the model with GA in linear form was adequate.

Table 9: Model Selection using closed test procedure

Model	Mean Structure	R^2	AIC	-2 Log Lik	-2 ln(λ_n)	df	$p-$value
1	$GA^{-2} + GA^2$	0.06	106794.6	106786.6		4	
2	β_0 (Null model)	-	107560.8	107556.8	770.2 (1 vs 2)	1	<0.0001
3	GA (straight line)	0.06	106794.8	106788.8	2.2 (1 vs 3)	2	0.3329

To visualize how the best fitting FPs and the chosen model fitted the data set, a plot of observed pregnancy weights as a function of GA was constructed with an overlay of these models as displayed in Figure 4.

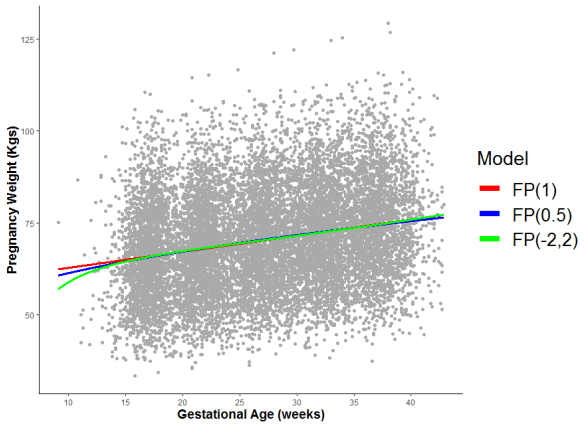


Figure 4: A plot of pregnancy weights as a function of GA with overlays of the best fitting FPs and the chosen model

Figure 4 showed that all the three FPs approximately fall on the same line on the plot, with little deviation at the lower ends.

4.3 Linear Mixed Model

From the data exploration and the fitted FP models, it was preferable to incorporate GA in the linear mixed effect model in its linear form as expressed in Section 3.4.3. Test for the need for both random intercept and random slope in the model was performed and the results were as displayed in Table 10. The results obtained implied that random slope had a statistical significant effect in the model ($p-$value< 0.0001), hence a model with both the random effects was suitable.

Table 10: Likelihood Ratio Test for Random Slope

Random Effects	AIC	-2 Log Lik	-2 ln(λ_n)	Asymptotic Null	$p-$value
Intercept + Slope	61749.59	61737.6			
Intercept	67764.04	67756.04	6018.44	$\chi^2_{1:2}$	<0.0001

The squared residuals plot of the smoothed mean curve (Figure 3b) suggested that the variability of observed weight was relatively constant over GA, therefore LMM with Compound Symmetric covariance structure was fitted. To ensure that the random effects had a meaningful interpretation, GA values were transformed by centering them around the minimum value of the recorded observations at baseline. That is, GA_{ij} was substituted by $GA_{ij} - 9.1429$ in the LMM. Table 11 indicated that GA had a statistical significant effect in the model ($p < 0.0001$).

Table 11: Type III Tests of Fixed Effect

Effect	Numerator DF	Denominator DF	$F-$value	$p-$value
GA	1	2760.3	14853	<0.0001

The parameter estimates of the fitted model were presented in Table 12. The results indicated that the overall average value of pregnancy weight is approximately 62 kgs when gestational age value is at minimum, and the value is statistically significant. In addition, there was statistically significant positive effect of GA, implying that the pregnancy weight increased over the gestational age ($p < 0.0001$).

Table 12: LMM parameter estimates

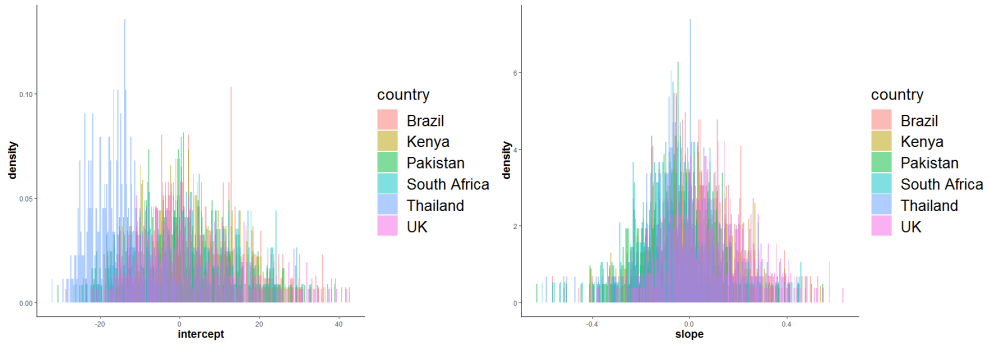
Parameter	Estimate	std. error	DF	$t-$value	$p-$value
Intercept	62.2023	0.2448	2818	254.1	<0.0001
GA	0.4406	0.0036	2760	121.9	<0.0001

The D matrix of covariance of random intercept and random slope was as shown below:

$$D = \begin{bmatrix} 167.2935 & -0.2545 \\ -0.2545 & 0.0320 \end{bmatrix}$$

The correlation between the random slope and random intercept was calculated as $corr(b_{1i}, b_{2i}) = \frac{d_{12}}{\sqrt{d_{11}}\sqrt{d_{22}}} = -0.11$. This suggests that women who had on average a low starting weight, had a faster rate of weight gain. Conversely, women with higher starting weight, had a lower rate of weight gain over the gestational age. It was noted that random intercept represented most of the variability. There was little variability within the expectant mothers $\sigma^2_{res} = 0.9037$. Figure 5 displays the distribution of the random effects stratified by country level.

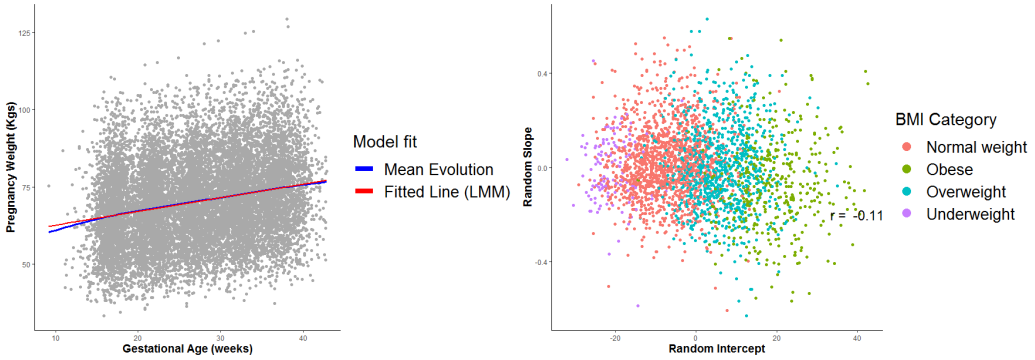
The 0 point in Figures 5a and 5b represented respectively the global average of starting pregnancy weight and the average rate of weight gain during the gestation period. The negative and positive values indicated women positioned below and above these global averages, respectively.



(a) Random Intercept **(b)** Random Slope

Figure 5: Distribution of random slopes and intercepts from the fitted linear mixed model stratified by country.

Figure 6a illustrates graphically the model fit of the linear mixed model. Both the fitted line and the mean evolution line lie approximately on the same path, suggesting that the model fitted the data well. The scatter plot of random slope and random intercept stratified by BMI category was displayed in Figure 6b. It was observed that underweight and normal weight women had faster rate of weight gain compared to overweight or obese women.



(a) Mean Evolution Plot **(b)** Scatter plot of Random slopes and intercepts

Figure 6: (a) Mean evolution of weight versus gestational age with an overlay of fitted line from Linear mixed model, to understand the model fit. (b) A scatter plot of random slope and intercept from linear mixed model, visualizing relationship between the two, stratified by BMI category.

4.4 Generalized Linear Model

Several general linear models using 'log' link function were fitted to quantify the association between the random effects estimates from the linear mixed model with pregnancy weight regressed on gestational age, with the maternal/ neonatal outcomes. Given the low frequency of occurrence of some outcomes (Table 18 in Appendix), a sample size greater than 30 was utilized for interpretation. Simple log-binomial models with single covariate; each random effect from the LMM

(random intercept and random slope) were first fitted for each outcome (Models 1 and 2 in Table 13). Thereafter, more complex models having both covariates (Model 3 in Table 13), with interaction between these covariates, and adjusting for maternal characteristics (Tables 14 and 15) were considered. Test for interaction effect for each outcome using likelihood ratio test was not statistically significant at $\alpha = 5\%$, hence interaction was not needed in the models. Together with the estimates, graphical visualization of the relation between predicted relative risks of the outcomes and each predictor was used for interpretation of the effect measure (Figures 7 and 8).

The outcomes were categorized into 'Yes' (1), having experienced the outcome and 'No' (0), with no outcome. The 0 category was considered as the reference group. The interpretation were as follows;

Preterm birth: Figure 8a shows that women positioned above and below the global mean weight at the minimum value of GA, had neither increased nor decreased risk of preterm birth. This was confirmed by an estimate of $RR = 1; 95\% CI,\ 0.99 - 1.01$ for the random intercept. On the other hand, there was an inverse association between the rate of weight gain from the global average and the risk of having preterm birth. This association was statistically significant in all the models ($RR = 0.36; 95\%,\ 0.19 - 0.66$) except for the model that adjusted for other covariates ($RR = 0.77; 95\%,\ 0.32 - 1.89$).

Fetal distress: The findings indicate that the association between both deviations of weight from the overall average weight at the minimum GA value, and the rate of weight gain from the overall average, on the risk of fetal distress was not statistically significant. As the rate of weight gain increased by 1 kg, the risk of having fetal distress also increased by approximately 24% ($RR = 1.24; 95\% CI,\ 0.60 - 2.55$). Similarly, Figure 8c suggests that there was an inverse J shape relationship between deviations from the overall average weight at the lowest GA value and the risk of fetal distress.

New born in intensive special care unit (ISCU): Figures 8e and 8f display a negative and a positive curved association between respectively random intercept and random slope and the chances of a newborn being moved to the ISCU. The risk of having newborn in the ISCU was higher for women who had weight below the global average at the minimum value of GA, and vice versa. This was confirmed from the statistical analysis. As the deviation increased, the risk of being in ISCU decreased by 0.76% ($RR = 0.99; 95\%,\ 0.98 - 1.00$), while as the rate of weight gain increased, the risk also increased by 20% ($RR = 1.20; 95\%,\ 0.64 - 2.26$). Both the relationships were statistically insignificant. Adjusting for other maternal characteristics widened the confidence interval of the random slope estimate but the effect still remained insignificant.

Birth Weight: This was categorized into small for gestational age (SGA), defined as babies who were below the 10^{th} centile for birth weight, and large for gestational age (LGA), consisting of babies who were above the 90^{th} centile for birth weight as explained by Villar *et al.* (2014). From the findings, as the deviations from the global average weight at the minimum value of GA increased, the risk of having SGA and LGA babies reduced ($RR = 0.96; 95\%,\ 0.95 - 0.97$) and increased ($RR = 1.04; 95\%,\ 1.03 - 1.05$), respectively. On the other hand, the risk of SGA and LGA neonates respectively decreased ($RR = 0.10; 95\%,\ 0.05 - 0.20$) and increased ($RR = 8.86; 95\%,\ 4.57 - 17.19$) with the increase in the rate of weight gain (Figures 8g-8j). Incorporating other covariates in the model did not change the statistical significance of the effect measure, apart from a slight reduction on the random slope estimate for below C10 BW outcome and an increase on the same for above C90 BW outcome.

Head Circumference (HC): This consisted of two groups; newborns who were below the 10^{th} centile for birth head circumference, and the ones who were above the 90^{th} centile for birth head circumference as defined by Villar *et al.* (2014). There was a curved association of both random intercept and random slope on the risk of having below C10 HC and above C90 HC neonates (Figures 8k-8n). As the woman's weight from the overall average weight at the minimum GA increased, the risk of having below C10 HC babies decreased ($RR = 0.93; 95\%,\ 0.91 - 0.94$) while the risk for above C90 HC neonates increased ($RR = 1.03; 95\%,\ 1.02 - 1.04$). Likewise, as the rate of weight gain from the overall average increased the risk for below C10 HC babies reduced ($RR = 0.13; 95\%,\ 0.06 - 0.31$) and risk for above C90 HC newborns increased ($RR = 4.04; 95\%,\ 2.35 - 6.96$). All the changes were statistically significant and were not affected by adding other covariates in the model.

Birth Length (BL): According to the international growth standards, BL was grouped into two classes; below C10 BL defined as babies who were below the 10^{th} centile for birth length, and above C90 BL representing babies who were above the 90^{th} centile for birth length (Villar *et al.*, 2014). The statistical analysis showed that risk of having below C10 BL neonates reduced ($RR = 0.97; 95\%,\ 0.96 - 0.97$) and risk for above C90 BL births increased ($RR = 1.02; 95\%,\ 1.01 - 1.03$) with an increase of deviations from the overall average weight at minimum GA. Similarly, increase in the rate of weight gain lead to a decrease in the risk of having below C10 BL babies ($RR = 0.21; 95\%,\ 0.10 - 0.43$), and an increase for above C90 BL newborns ($RR = 1.59; 95\%,\ 1.80 - 3.13$). The effect of rate of weight gain on the above C90 BL outcome was statistically insignificant in all the fitted models.

Pregnancy induced hypertension (PIH): Graphical representation of the relative risk of PIH as a function of each random effect estimate, suggests a positive curved relationship (Figures 7a and 7b). From the analysis, increase in both the value of deviation from the global mean weight at the minimum GA, and rate of weight gain from the average lead to an increase in the risk of being hypertensive, with $RR = 1.03; 95\%,\ 1.02 - 1.04$ and $RR = 3.32; 95\%,\ 1.42 - 7.62$, respectively. A statistical significance effect was obtained in all the fitted models except a non-significance effect of rate of weight gain after adjusting for other covariates.

Gestational Diabetes: A positive curved association of risk of having gestational diabetes and starting pregnancy weight was observed (Figure 7c). Conversely, an inverse relation of rate of weight gain and the risk of diabetes during pregnancy was depicted in Figure 7d. As the starting weight from the overall average at the minimum GA increased, the risk of diabetes also increased ($RR = 1.03; 95\%,\ 1.02 - 1.04$). Inversely, increase in the rate of weight gain lead to reduced chance of being diabetic during pregnancy ($RR = 0.05; 95\%,\ 0.02 - 0.10$). All the models revealed a statistically significant effect of the associations and on the same direction.

Preterm Labour: Figures 7g and 7h display a negative and a positive curved association between random intercept and random slope with risk of preterm labour, respectively. All the effect measures estimates were non-significant except for the random intercept only model. Increase in the deviations from the average weight at the minimum value of GA lead to a low chance of having preterm labour ($RR = 0.98; 95\%,\ 0.97 - 1.00$). The risk of preterm labour increased with the increased of rate of weight gain ($RR = 1.33; 95\%,\ 0.41 - 4.23$), however the change was not significant.

Cesarean section: There was an inverse curved association of rate of weight gain with risk of having C-section (Figure 7t). This relationship was not statistically significant ($RR = 0.99; 95\%$, $0.77 - 1.28$). However, as the deviations from the overall average weight increased, the risk of having C-section also significantly increased ($RR = 1.02; 95\%,\ 1.01 - 1.02$). The effect measures had the same interpretation for all the models but was non-significant after adjusting for other covariates.

An inverse J shaped relationship was observed in the plots of relative risks of failure to progress, cephalopelvic disproportion (CPD), and preterm premature rupture of the membranes (PPROM), as a function of deviations from the global average weight at minimum value of GA (Figures 7m,7o,7q). This suggested that women with starting weight below the overall mean had an increased risk of failure to progress, CPD, and PPROM, and vice versa. Figures 7n, 7p, 7r depicted an almost positive linear association between rate of weight gain from the average and risk of these outcomes. A statistical significant effect of the effect measure was obtained for models with random slope as the only covariate on the risk of failure to progress ($RR = 2.37; 95\%,\ 1.19 - 4.63$), CPD ($RR = 5.19; 95\%,\ 1.09 - 23.74$), and PPROM ($RR = 3.52; 95\%,\ 1.26 - 9.57$).

Table 13: GLMs Parameter estimates

Outcome (n)	Par.	Model 1			Model 2			Model 3		
		RR	**LCI**	**UCI**	**RR**	**LCI**	**UCI**	**RR**	**LCI**	**UCI**
Maternal										
Preg. induced hyp	b_1	1.0296*	1.0183	1.0406				1.0310*	1.0198	1.0420
(163)	b_2				2.8648*	1.1681	6.9171	3.3231*	1.4244	7.6161
Gestational	b_1	1.0337*	1.0234	1.0438				1.0259*	1.0156	1.0360
Diabetes (186)	b_2				0.0276*	0.0139	0.0586	0.0455*	0.0225	0.0966
Preeclampsia (32)	b_1	1.0147	0.9883	1.0408				1.0154	0.9888	1.0417
	b_2				1.4442	0.1823	10.9271	1.6186	0.2107	11.7517
Preterm Labour	b_1	0.9847*	0.9693	0.9999				0.9850	0.9696	1.0003
(100)	b_2				1.4660	0.4659	4.5362	1.3284	0.4078	4.2254
Breech presentation	b_1	0.9900	0.9713	1.0085				0.9889	0.9699	1.0077
(71)	b_2				0.7887	0.2182	2.8213	0.6701	0.1757	2.4986
Reduced fetal	b_1	0.9964	0.9699	1.0227				1.0002	0.9730	1.0271
movement (36)	b_2				4.2171	0.7022	23.9791	4.2240	0.6823	24.3985
Failure to progress	b_1	0.9876*	0.9774	0.9977				0.9896	0.9792	1.0000
(216)	b_2				2.3683*	1.1927	4.6310	2.0733*	1.0216	4.1315
Cephalopelvic dis.	b_1	0.9746*	0.9510	0.9979				0.9778	0.9537	1.0017
(47)	b_2				5.1949*	1.0880	23.7355	4.2154	0.8313	20.1383
PPROM (105)	b_1	0.9858	0.9706	1.0010				0.9887	0.9732	1.0042
	b_2				3.5173*	1.2589	9.5693	3.1096*	1.0788	8.6696
Caesarean section	b_1	1.0156*	1.0124	1.0188				1.0156*	1.0124	1.0188
(1177)	b_2				0.8583	0.6576	1.1203	0.9908	0.7684	1.2777
Neonatal										
Preterm birth (341)	b_1	0.9988	0.9910	1.0065				0.9970	0.9892	1.0047
	b_2				0.3733*	0.2055	0.6799	0.3589*	0.1949	0.6607
Fetal distress (212)	b_1	1.0006	0.9906	1.0105				1.0011	0.9909	1.0113
	b_2				1.2231	0.5959	2.4895	1.2415	0.5971	2.5492
Newborn ISCU	b_1	0.9921	0.9840	1.0003				0.9924	0.9841	1.0005
(324)	b_2				1.2738	0.6822	2.3654	1.2040	0.6365	2.2599
Below C10 BW	b_1	0.9579*	0.9484	0.9673				0.9565*	0.9468	0.9663
(276)	b_2				0.1690*	0.0874	0.3305	0.0959*	0.0452	0.2033
Above C90 BW	b_1	1.0379*	1.0287	1.0468				1.0418*	1.0334	1.0503
(221)	b_2				7.6984*	3.7176	15.6389	8.8605*	4.5660	17.1944
Below C10 HC	b_1	0.9291*	0.9179	0.9403				0.9301*	0.9176	0.9428
(214)	b_2				0.2716*	0.1265	0.5871	0.1337*	0.0578	0.3093
Above C90 HC	b_1	1.0285*	1.0215	1.0355				1.0316*	1.0250	1.0382
(350)	b_2				3.4829*	1.9490	6.1557	4.0420*	2.3458	6.9646
Below C10 BL	b_1	0.9671*	0.9584	0.9757				0.9660*	0.9574	0.9747
(312)	b_2				0.3067*	0.1639	0.5768	0.2105*	0.1030	0.4299
Above C90 BL	b_1	1.0233*	1.0143	1.0322				1.0243*	1.0160	1.0327
(242)	b_2				1.3039	0.6322	2.6697	1.5866	0.8049	3.1275

b_1 - Random intercept, b_2 - Random Slope, **RR** - Risk ratio, **LCI** - Lower Confidence Interval, **UCI** - Upper Confidence Interval, * - significant at 5%, Model 1 - with random intercept as covariate, Model 2 - with random slope, Model 3 - with both random intercept and slope.

Table 14: GLMs Estimates adjusted for age, country, systolic BP, previous preterm birth, , previous Cesarean Section, and previous neonatal death.

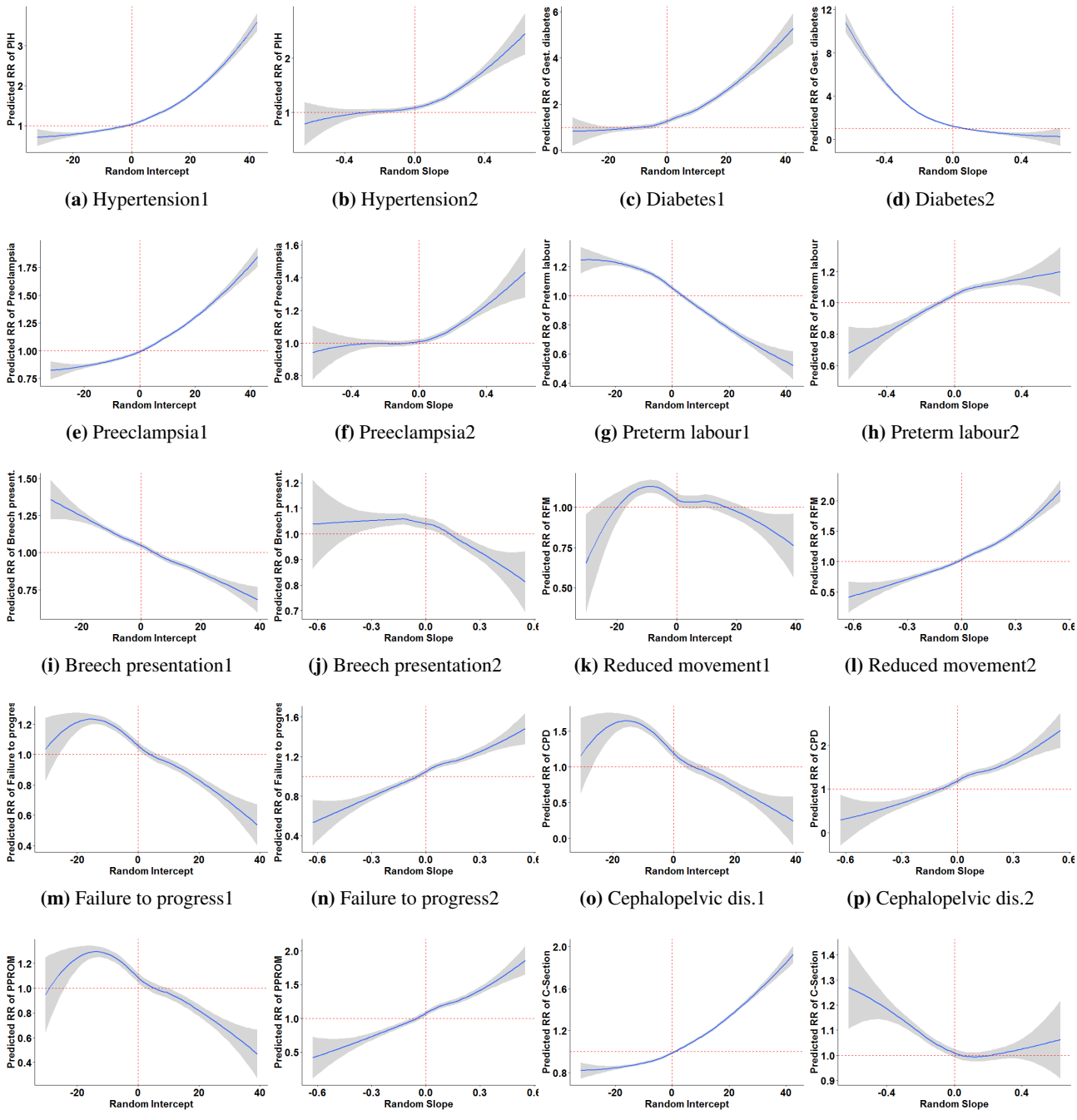
Outcome (n)	Parameter	**RR**	**LCI**	**UCI**	Outcome (n)	Parameter	**RR**	**LCI**	**UCI**
Preterm birth (341)	b_1	0.99	0.97	1.00	Newborn ISCU (324)	b_1	0.99	0.98	1.01
	b_2	0.77	0.32	1.89		b_2	1.23	0.49	3.10
	Age	1.01	0.98	1.04		Age	1.00	0.97	1.03
	Prev. preterm	1.89*	1.35	2.64		Prev. preterm	1.93*	1.32	2.84
	Systolic BP	1.01	1.00	1.02		Systolic BP	1.01*	1.00	1.03
	Prev. neo. death	1.53*	1.01	2.32		Prev. neo. death	1.63	0.98	2.70
	Prev. CS	0.60*	0.44	0.81		Prev. CS	0.56*	0.39	0.80
	Country					**Country**			
	Brazil	1.91	0.27	13.67		Brazil	0.28*	0.10	0.81
	Kenya	0.69	0.09	5.53		Kenya	0.24*	0.08	0.69
	Pakistan	3.31	0.49	22.56		Pakistan	0.29*	0.12	0.74
	SA	3.37	0.50	22.69		SA	0.74	0.32	1.73
	UK	2.07	0.30	14.16		UK	0.49	0.20	1.19
	Thailand ('ref')					Thailand ('ref')			
Below C10 BL (312)	b_1	0.97*	0.95	0.99	Above C90 HC (350)	b_1	1.02*	1.01	1.04
	b_2	0.07*	0.02	0.26		b_2	3.97*	1.86	8.47
	Age	1.01	0.96	1.05		Age	1.03	1.00	1.05
	Prev. preterm	1.11	0.67	1.85		Prev. preterm	0.84	0.55	1.28
	Systolic BP	1.00	0.98	1.01		Systolic BP	1.00	0.99	1.01
	Prev. neo. death	1.29	0.58	2.87		Prev. neo. death	0.78	0.35	1.73
	Prev. CS	0.83	0.54	1.27		Prev. CS	1.05	0.79	1.41
	Country					**Country**			
	Brazil	1.05	0.40	2.73		Brazil	0.75	0.19	2.90
	Kenya	0.41	0.15	1.14		Kenya	0.95	0.25	3.58
	Pakistan	0.36*	0.14	0.92		Pakistan	0.44	0.12	1.69
	SA	0.66	0.29	1.55		SA	0.90	0.24	3.35
	UK	0.85	0.35	2.08		UK	1.00	0.27	3.75
	Thailand ('ref')					Thailand ('ref')			
Below C10 BW (276)	b_1	0.97*	0.95	0.99	Above C90 BW (221)	b_1	1.04*	1.03	1.06
	b_2	0.02*	0.01	0.07		b_2	6.10*	2.21	16.82
	Age	0.98	0.93	1.02		Age	1.03	0.99	1.08
	Prev. preterm	1.15	0.66	2.01		Prev. preterm	0.79	0.40	1.57
	Systolic BP	1.00	0.98	1.02		Systolic BP	0.99	0.97	1.00
	Prev. neo. death	0.91	0.35	2.38		Prev. neo. death	0.88	0.25	3.11
	Prev. CS	0.52*	0.32	0.84		Prev. CS	1.17	0.77	1.79
	Country					**Country**			
	Brazil	0.91	0.30	2.78		Brazil	0.58	0.08	4.27
	Kenya	0.89	0.33	2.38		Kenya	0.65	0.09	4.61
	Pakistan	0.60	0.22	1.59		Pakistan	0.24	0.03	1.76
	SA	0.57	0.23	1.40		SA	0.57	0.08	4.02
	UK	1.03	0.40	2.62		UK	0.90	0.13	6.34
	Thailand ('ref')					Thailand ('ref')			
Above C90 BL (242)	b_1	1.03*	1.02	1.05	Below C10 HC (214)	b_1	0.98	0.95	1.01
	b_2	1.33	0.48	3.66		b_2	0.05*	0.01	0.28
	Age	1.03	0.99	1.07		Age	0.94	0.88	1.00
	Prev. preterm	0.92	0.51	1.66		Prev. preterm	1.01	0.42	2.42
	Systolic BP	0.98	0.97	1.00		Systolic BP	1.00	0.98	1.02
	Prev. neo. death	0.53	0.16	1.75		Prev. neo. death	0.60	0.12	3.04
	Prev. CS	0.81	0.54	1.20		Prev. CS	0.54	0.28	1.06
	Country					**Country**			
	Brazil	0.17*	0.05	0.57		Brazil	0.31	0.07	1.41
	Kenya	0.25*	0.09	0.74		Kenya	0.47	0.15	1.53
	Pakistan	0.24*	0.09	0.67		Pakistan	0.45	0.16	1.24
	SA	0.32*	0.12	0.87		SA	0.21*	0.07	0.68
	UK	0.36*	0.13	0.99		UK	0.53	0.16	1.76
	Thailand ('ref')					Thailand ('ref')			

b_1 - Random intercept, b_2 - Random Slope, **RR** - Risk ratio, **LCI** - Lower Confidence Interval, **UCI** - Upper Confidence Interval, $*$ - significant at 5%.

Table 15: Estimates adjusted for age, country, systolic BP, previous preterm birth, previous Cesarean Section, and previous neonatal death.

Outcome (n)	Parameter	**RR**	**LCI**	**UCI**	Outcome (n)	Parameter	**RR**	**LCI**	**UCI**
Pregnancy induced hypertension (163)	b_1	1.04*	1.01	1.06	Gestational Diabetes (186)	b_1	1.03*	1.01	1.04
	b_2	1.81	0.35	9.27		b_2	0.24*	0.07	0.82
	Age	0.97	0.93	1.01		Age	1.03	0.99	1.07
	Prev. preterm	1.77	0.99	3.14		Prev. preterm	1.22	0.79	1.88
	Systolic BP	1.05*	1.04	1.07		Systolic BP	1.01	0.99	1.02
	Prev. neo. death	1.51	0.62	3.67		Prev. neo. death	1.30	0.81	2.08
	Prev. CS	0.62	0.38	1.01		Prev. CS	0.75	0.54	1.05
	Country					**Country**			
	Brazil	0.40	0.13	1.25		Brazil	0.53	0.15	1.89
	Kenya	0.07*	0.02	0.33		Kenya	0.25*	0.07	0.94
	Pakistan	0.37	0.12	1.13		Pakistan	1.62	0.52	5.03
	SA	0.26*	0.09	0.76		SA	0.06*	0.01	0.31
	UK	0.34	0.11	1.07		UK	0.37	0.11	1.23
	Thailand ('ref')					Thailand ('ref')			
PPROM (105)	b_1	1.00	0.97	1.03	Breech presentation (71)	b_1	0.98	0.96	1.01
	b_2	0.64	0.12	3.50		b_2	2.07	0.35	12.34
	Age	0.98	0.93	1.03		Age	0.99	0.93	1.06
	Prev. preterm	1.57	0.72	3.44		Prev. preterm	1.95	0.95	4.01
	Systolic BP	0.98	0.95	1.00		Systolic BP	1.02	1.00	1.05
	Prev. neo. death	0.67	0.10	4.41		Prev. neo. death	1.19	0.40	3.49
	Prev. CS	0.31*	0.13	0.76		Prev. CS	0.40*	0.18	0.86
	Country					**Country**			
	Brazil	0.63	0.16	2.50		Brazil	0.25	0.04	1.48
	Kenya	0.37	0.08	1.68		Kenya	0.24	0.04	1.44
	Pakistan	0.07*	0.01	0.47		Pakistan	0.51	0.11	2.27
	SA	0.18*	0.04	0.89		SA	0.27	0.06	1.17
	UK	0.81	0.22	2.98		UK	0.46	0.10	2.18
	Thailand ('ref')					Thailand ('ref')			
Cesarean section (1177)	b_1	1.00	0.99	1.00	Failure to progress (216)	b_1	0.98	0.96	1.00
	b_2	0.96	0.80	1.16		b_2	0.64	0.15	2.62
	Age	1.00	1.00	1.01		Age	1.01	0.96	1.05
	Prev. preterm	1.10*	1.03	1.18		Prev. preterm	0.65	0.32	1.34
	Systolic BP	1.00	0.98	1.00		Systolic BP	0.98	0.97	1.00
	Prev. neo. death	1.03	0.94	1.14		Prev. neo. death	1.35	0.57	3.22
	Prev. CS	1.62*	1.51	1.74		Prev. CS	0.35*	0.20	0.63
	Country					**Country**			
	Brazil	2.32*	1.23	4.36		Brazil	1.38	0.17	11.48
	Kenya	1.83	0.97	3.46		Kenya	3.16	0.42	24.07
	Pakistan	1.57	0.83	2.97		Pakistan	0.99	0.13	7.53
	SA	2.16*	1.15	4.06		SA	3.30	0.46	23.37
	UK	1.26	0.66	2.40		UK	0.70	0.08	6.29
	Thailand ('ref')					Thailand ('ref')			

b_1 - Random intercept, b_2 - Random Slope, **RR** - Risk ratio, **LCI** - Lower Confidence Interval, **UCI** - Upper Confidence Interval, $*$ - significant at 5%.



(a) Hypertension1 **(b)** Hypertension2 **(c)** Diabetes1 **(d)** Diabetes2

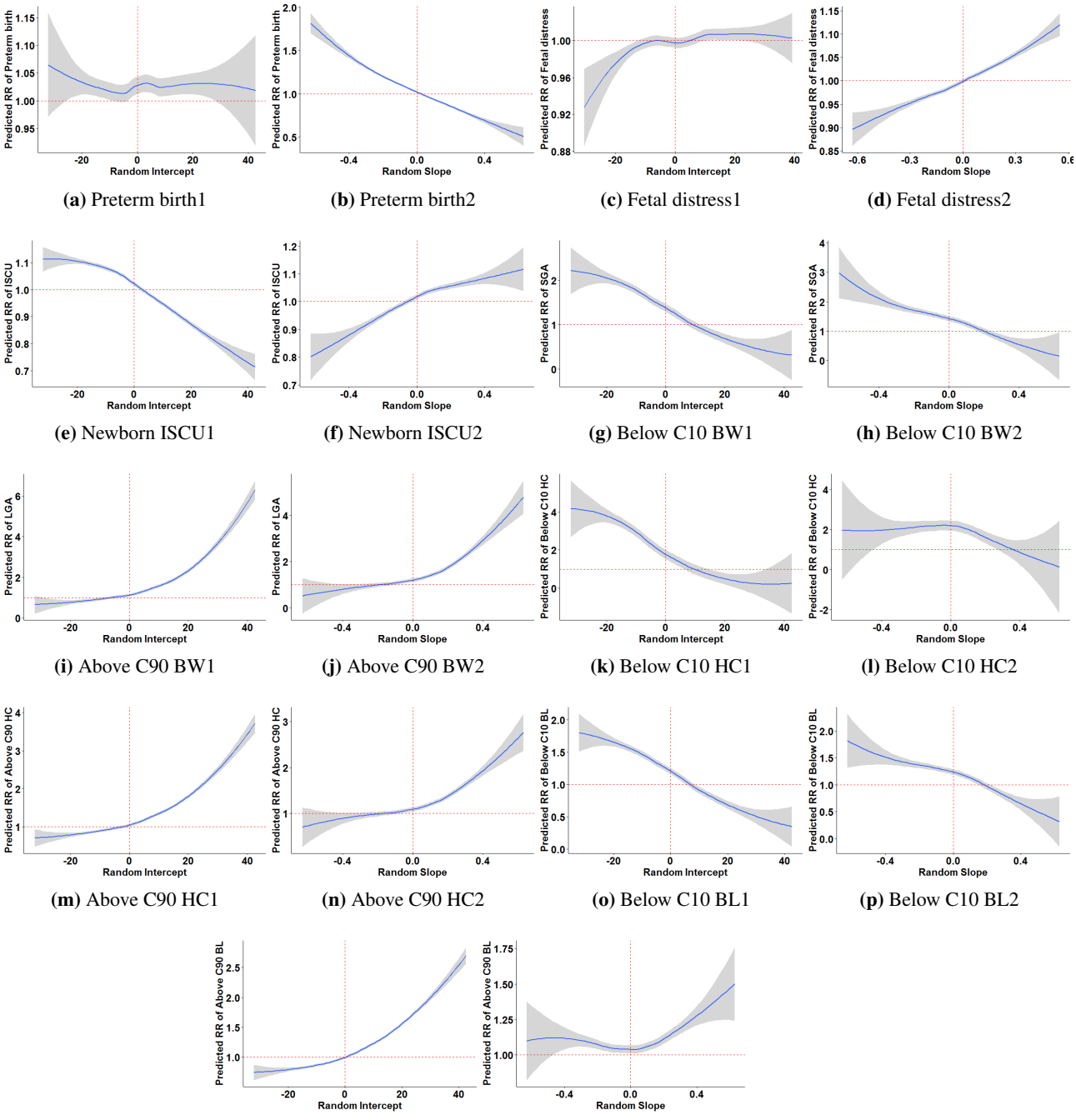
(e) Preeclampsia1 **(f)** Preeclampsia2 **(g)** Preterm labour1 **(h)** Preterm labour2

(i) Breech presentation1 **(j)** Breech presentation2 **(k)** Reduced movement1 **(l)** Reduced movement2

(m) Failure to progress1 **(n)** Failure to progress2 **(o)** Cephalopelvic dis.1 **(p)** Cephalopelvic dis.2

(q) PPROM1 **(r)** PPROM2 **(s)** Caesarean section1 **(t)** Caesarean section2

Figure 7: Graphical representation of the predicted relative risks of maternal outcomes as a function of random intercept (1) and slope (2), with the grey area representing the confidence interval. **RFM**: Reduced fetal movement, **CPD**: Cephalopelvic disproportion. The red lines indicate the reference points along the axes; $RR = 1$ represents the point of no association and 0 point on the $x-$axis denotes the global average values for starting weight (intercept) and rate of weight gain (slope).



(a) Preterm birth1 **(b)** Preterm birth2 **(c)** Fetal distress1 **(d)** Fetal distress2

(e) Newborn ISCU1 **(f)** Newborn ISCU2 **(g)** Below C10 BW1 **(h)** Below C10 BW2

(i) Above C90 BW1 **(j)** Above C90 BW2 **(k)** Below C10 HC1 **(l)** Below C10 HC2

(m) Above C90 HC1 **(n)** Above C90 HC2 **(o)** Below C10 BL1 **(p)** Below C10 BL2

(q) Above C90 BL1 **(r)** Above C90 BL2

Figure 8: Graphical representation of the predicted relative risks of neonatal outcomes as a function of random intercept (1) and slope (2), with the grey area representing the confidence interval. The red lines indicate the reference points along the axes; $RR = 1$ represents the point of no association and 0 point on the $x-$axis denotes the global average values for starting weight (intercept) and rate of weight gain (slope).

5 Discussions

In this study, a two level analysis including fitting a linear mixed effects model to assess the pattern of gestational weight gain, and generalized linear models to examine the association between the random effects generated from the former and the maternal/ neonatal outcomes, was performed.

In the multilevel analysis, weight measurements at each prenatal visit was regressed on gestational age (GA) treating each subject as a random effect in the model. As weight changes depicted almost a linear relation between GA and the observed weights for each woman, having less and much variability observed respectively within and between individuals, a model with both random intercept and slope incorporating GA in its linear form was fitted. Additionally, the best fitting power for GA was given by first degree fractional polynomial of order 1. To ensure that random effects had a meaningful zero point on the scales that did not have such a value, GA values were centered around the minimum of the first values of GA. The estimates of random intercept and slope from the fitted model varied between -32.09 to 42.59 and -0.63 to 0.63, respectively. Indicating that each woman had a different starting weight and dissimilar pattern of weight gain/ loss during the pregnancy period. A correlation of -0.11 between random slope and intercept suggested that women who had a low starting weight on average had a faster rate of weight gain, and vice versa.

In the second step, different generalized linear models were fitted to quantify the link between the obtained random effects estimates and the neonatal/ maternal outcomes. The findings indicated that increase in the deviations from the global average weight at the minimum value of GA is associated with high risk of large for gestational age neonates, newborns with above 90^{th} centile for both birth head circumference and length, pregnancy induced hypertension, gestational diabetes, and increased risk of C-section. This is consistent with the findings of Rhodes *et al.* (2003), who concluded that women with excess weight gain contribute to a higher proportion of C-section delivery. Whereas, starting pregnancy at a low weight below the average was linked with small for gestational age newborns, babies below 10^{th} centile for both birth head circumference and length. On the other hand, increased rate of weight gain from the average resulted in high chances of large for gestational age neonates, above 90^{th} centile for birth head circumference newborns, pregnancy induced hypertension, preterm premature rupture of the membranes (PPROM), and failure to progress.

GLMs with single covariates (random intercept/ slope from LMM) and including both the random effects had similar interpretation of the parameter estimates. However, adjusting for other covariates had an impact on the association between these random effects and the maternal/ neonatal outcomes. Suggesting that other important maternal factors such as country of residence, age, blood pressure, having had previous preterm births, neonatal deaths and C-section were also determinants of neonatal/ maternal outcomes.

6 Conclusions and Recommendation

In summary, the study demonstrated that starting pregnancy weight and the rate of gestational weight gain lead to risk of adverse maternal/ neonatal outcomes. To this end, it is advisable for pregnant mothers to gain weight within the recommended limits to ensure a healthy life for themselves and their babies during the pregnancy period. Additionally, the maternal characteristics mentioned in Section 5 should be taken into consideration and treated as pregnancy risk factors. Policies should be put in place to ensure that the available gestational weight gain guidelines provided by US Institute of Medicine (IOM) and WHO are strictly adhered to especially in low middle income countries.

As part of future research, this analysis could be improved by applying joint modelling composed of continuous outcome; gestational weight and binary outcome; maternal/ neonatal outcomes instead of conducting two separate statistical analysis.

References

Benner, A. (2014). Multivariable Fractional Polynomials.

Bodnar, L. M., Hutcheon, J. A., Parisi, S. M., Pugh, S. J., & Abrams, B. (2015). Comparison of gestational weight gain z-scores and traditional weight gain measures in relation to perinatal outcomes. Paediatric and perinatal epidemiology, 29(1), 11-21.

Carter, R. E., Lipsitz, S. R., & Tilley, B. C. (2005). Quasi-likelihood estimation for relative risk regression models. Biostatistics, 6(1), 39-44.

Coffey, D. (2015). Prepregnancy body mass and weight gain during pregnancy in India and sub-Saharan Africa. Proceedings of the National Academy of Sciences, 112(11), 3302-3307.

Deputy, N. P., Sharma, A. J., Kim, S. Y., & Hinkle, S. N. (2015). Prevalence and characteristics associated with gestational weight gain adequacy. Obstetrics and gynecology, 125(4), 773.

Drehmer, M., Duncan, B. B., Kac, G., & Schmidt, M. I. (2013). Association of second and third trimester weight gain in pregnancy with maternal and fetal outcomes. PloS one, 8(1), e54704.

FAO. (2015). Regional overview of food insecurity: African food security prospects brighter than ever.

Goldstein, R. F., Abell, S. K., Ranasinha, S., Misso, M. L., Boyle, J. A., Harrison, C. L., ... & Teede, H. J. (2018). Gestational weight gain across continents and ethnicity: systematic review and meta-analysis of maternal and infant outcomes in more than one million women. BMC medicine, 16(1), 1-14.

Hickey, C. A. (2000). Sociocultural and behavioral influences on weight gain during pregnancy. The American journal of clinical nutrition, 71(5), 1364S-1370S.

Hinkle, S. N., Mitchell, E. M., Grantz, K. L., Ye, A., & Schisterman, E. F. (2016). Maternal weight gain during pregnancy: comparing methods to address bias due to length of gestation in epidemiological studies. Paediatric and perinatal epidemiology, 30(3), 294-304.

Hoffman, C. S., Messer, L. C., Mendola, P., Savitz, D. A., Herring, A. H., & Hartmann, K. E. (2008). Comparison of gestational age at birth based on last menstrual period and ultrasound during the first trimester. Paediatric and perinatal epidemiology, 22(6), 587-596.

Hutcheon, J. A., & Bodnar, L. M. (2018). Good practices for observational studies of maternal weight and weight gain in pregnancy. Paediatric and perinatal epidemiology, 32(2), 152.

Johansson, K., Hutcheon, J. A., Stephansson, O., & Cnattingius, S. (2016). Pregnancy weight gain by gestational age and BMI in Sweden: a population-based cohort study. The American journal of clinical nutrition, 103(5), 1278-1284.

Johnson, J. L., Farr, S. L., Dietz, P. M., Sharma, A. J., Barfield, W. D., & Robbins, C. L. (2015). Trends in gestational weight gain: the pregnancy risk assessment monitoring system, 2000–2009. American journal of obstetrics and gynecology, 212(6), 806-e1.

Karachaliou, M., Georgiou, V., Roumeliotaki, T., Chalkiadaki, G., Daraki, V., Koinaki, S., ... & Chatzi, L. (2015). Association of trimester-specific gestational weight gain with fetal growth, offspring obesity, and cardiometabolic traits in early childhood. American journal of obstetrics and gynecology, 212(4), 502-e1.

Kennedy, S. H., Victora, C. G., Craik, R., Ash, S., Barros, F. C., Barsosio, H. C., ... & Villar, J. (2018). Deep clinical and biological phenotyping of the preterm birth and small for gestational age syndromes: The INTERBIO-21 st Newborn Case-Control Study protocol. Gates open research, 2.

Kirkegaard, H., Stovring, H., Rasmussen, K. M., Abrams, B., Sørensen, T. I., & Nohr, E. A. (2014). How do pregnancy-related weight changes and breastfeeding relate to maternal weight and BMI-adjusted waist circumference 7 y after delivery? Results from a path analysis. The American journal of clinical nutrition, 99(2), 312-319.

Kleinman, K. P., Oken, E., Radesky, J. S., Rich-Edwards, J. W., Peterson, K. E., & Gillman, M. W. (2007). How should gestational weight gain be assessed? A comparison of existing methods and a novel method, area under the weight gain curve. International journal of epidemiology, 36(6), 1275-1282.

Li, C., Liu, Y., & Zhang, W. (2015). Joint and independent associations of gestational weight gain and pre-pregnancy body mass index with outcomes of pregnancy in Chinese women: a retrospective cohort study. PLoS One, 10(8), e0136850.

Martínez-Hortelano, J. A., Cavero-Redondo, I., Álvarez-Bueno, C., Garrido-Miguel, M., Soriano-Cano, A., & Martínez-Vizcaíno, V. (2020). Monitoring gestational weight gain and prepregnancy BMI using the 2009 IOM guidelines in the global population: a systematic review and meta-analysis. BMC pregnancy and childbirth, 20(1), 1-12.

Mitchell, E. M., Hinkle, S. N., & Schisterman, E. F. (2016). It s about time: a survival approach to gestational weight gain and preterm delivery. Epidemiology (Cambridge, Mass.), 27(2), 182.

Muktabhant, B., Lawrie, T. A., Lumbiganon, P., & Laopaiboon, M. (2015). Diet or exercise, or both, for preventing excessive weight gain in pregnancy. Cochrane database of systematic reviews, (6).

National Research Council. (2010). Weight gain during pregnancy: reexamining the guidelines.

O'Brien, B., & Naber, S. (1992). Nausea and vomiting during pregnancy: effects on the quality of women's lives. Birth, 19(3), 138-143.

Ohuma, E. O., Altman, D. G., & International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st Project). (2019). Statistical methodology for constructing gestational age-related charts using cross-sectional and longitudinal data: The INTERGROWTH-21st project as a case study. Statistics in medicine, 38(19), 3507-3526.

Oken, E., Rifas-Shiman, S. L., Field, A. E., Frazier, A. L., & Gillman, M. W. (2008). Maternal gestational weight gain and offspring weight in adolescence. Obstetrics and gynecology, 112(5), 999.

Orach, D., Garimoi, C. (2009). Health equity: challenges in low income countries. African health sciences, 9(s2), S49-S51.

Papageorghiou, A. T., Kennedy, S. H., Salomon, L. J., Ohuma, E. O., Cheikh Ismail, L., Barros, F. C., ... & International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). (2014). International standards for early fetal size and pregnancy dating based on ultrasound measurement of crown–rump length in the first trimester of pregnancy. Ultrasound in Obstetrics Gynecology, 44(6), 641-648.

Pugh, S. J., Albert, P. S., Kim, S., Grobman, W., Hinkle, S. N., Newman, R. B., ... & Grantz, K. L. (2017). Patterns of gestational weight gain and birthweight outcomes in the Eunice Kennedy Shriver National Institute of Child Health and Human Development Fetal Growth Studies–Singletons: a prospective study. American journal of obstetrics and gynecology, 217(3), 346-e1.

SAS Institute Inc. 2013. SAS® 9.4 Statements: Reference. Cary, NC: SAS Institute Inc.

Rasmussen, K. M., & Yaktine, A. L. (2009). Data Tables. In Weight gain during pregnancy: reexamining the guidelines. National Academies Press (US).

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rhodes, J. C., Schoendorf, K. C., & Parker, J. D. (2003). Contribution of excess weight gain during pregnancy and macrosomia to the cesarean delivery rate, 1990–2000. Pediatrics, 111(Supplement 1), 1181-1185.

Riddell, C. A., Platt, R. W., Bodnar, L. M., & Hutcheon, J. A. (2017). Classifying gestational weight gain trajectories using the SITAR growth model. Paediatric and perinatal epidemiology, 31(2), 116-125.

Royston, P. (2017). Model selection for univariable fractional polynomials. The Stata Journal, 17(3), 619-629.

Royston, P., & Altman, D. G. (1994). Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. Journal of the Royal Statistical Society: Series C (Applied Statistics), 43(3), 429-453.

Royston, P., & Sauerbrei, W. (2008). Multivariable model-building: a pragmatic approach to regression analysis based on fractional polynomials for modelling continuous variables (Vol. 777). John Wiley Sons.

Scarf, V. L., Rossiter, C., Vedam, S., Dahlen, H. G., Ellwood, D., Forster, D., ... & Homer, C. S. (2018). Maternal and perinatal outcomes by planned place of birth among women with low-risk pregnancies in high-income countries: a systematic review and meta-analysis. Midwifery, 62, 240-255.

Schmidt, C. O., & Kohlmann, T. (2008). When to use the odds ratio or the relative risk?. International journal of public health, 53(3), 165.

Schwendinger, F., Grün, B., & Hornik, K. (2021). A comparison of optimization solvers for log binomial regression including conic programming. Computational Statistics, 1-34.

Scott, C., Andersen, C. T., Valdez, N., Mardones, F., Nohr, E. A., Poston, L., ... & Abrams, B. (2014). No global consensus: a cross-sectional survey of maternal weight policies. BMC pregnancy and childbirth, 14(1), 1-10.

Thangaratinam, S., Rogozińska, E., Jolly, K., Glinkowski, S., Roseboom, T., Tomlinson, J. W., ... & Khan, K. S. (2012). Effects of interventions in pregnancy on maternal weight and obstetric outcomes: meta-analysis of randomised evidence. Bmj, 344.

Tuerlinckx, F., Rijmen, F., Verbeke, G., & De Boeck, P. (2006). Statistical inference in generalized linear mixed models: A review. British Journal of Mathematical and Statistical Psychology, 59(2), 225-255.

Verbeke, G. and Molenberghs, G. (2000). Linear Mixed Models for Longitudinal Data. Springer-Verslag, New York.

Viera, A. J. (2008). Odds ratios and risk ratios: what's the difference and why does it matter?. Southern medical journal, 101(7), 730-734.

Villar, J., Ismail, L. C., Victora, C. G., Ohuma, E. O., Bertino, E., Altman, D. G., ... & Kennedy, S. H. (2014). International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. The Lancet, 384(9946), 857-868.

Voerman, E., Santos, S., Inskip, H., Amiano, P., Barros, H., Charles, M. A., ... & Gaillard, R. (2019). Association of gestational weight gain with adverse maternal and infant outcomes. Jama, 321(17), 1702-1715.

Wang, D., Wang, M., Darling, A. M., Perumal, N., Liu, E., Danaei, G., & Fawzi, W. W. (2020). Gestational weight gain in low-income and middle-income countries: a modelling analysis using nationally representative data. BMJ global health, 5(11), e003423.

Williamson, T., Eliasziw, M., & Fick, G. H. (2013). Log-binomial models: exploring failed convergence. Emerging themes in epidemiology, 10(1), 1-10.

Tables

Table 16: Fractional Polynomials of order 1 (FP1). Deviance difference compares the fit with that of a straight line (power=1). The maximum deviance difference $\sim \chi^2_1$ and a positive value indicates a model which fits better than a straight line

Model	Power	Mean Structure	R^2	AIC	BIC	Deviance difference
1	-2	GA^{-2}	0.05	106889	106911.5	-94.2
2	-1	GA^{-1}	0.05	106832.7	106855.2	-37.9
3	-0.5	$GA^{-0.5}$	0.05	106813.4	106835.9	-18.6
4	0	$\ln GA$	0.06	106800.6	106823.1	-5.8
5	0.5	$GA^{0.5}$	0.06	**106794.5**	**106817**	0.3
6	1	GA	0.06	106794.8	106817.3	0
7	2	GA^2	0.05	106812.6	106835.1	-17.8
8	3	GA^3	0.05	106847.7	106870.2	-52.9

Table 17: Fractional Polynomials of order 2 (FP2). The maximum deviance difference $\sim \chi^2_3$ and a positive value indicates a model which fits better than a straight line

Model	Powers	Mean Structure	R^2	AIC	BIC	DD	Model	Powers	Mean Structure	R^2	AIC	BIC	DD
1	(-2, -2)	$GA^{-2} + GA^{-2}\ln GA$	0.05	106812.3	106842.3	-15.5	19	(-0.5, 1)	$GA^{-0.5} + GA$	0.06	106795.6	106825.6	1.2
2	(-2, -1)	$GA^{-2} + GA^{-1}$	0.06	106803.9	106833.9	-7.1	20	(-0.5, 2)	$GA^{-0.5} + GA^2$	0.06	106795.1	106825.1	1.7
3	(-2, -0.5)	$GA^{-2} + GA^{-0.5}$	0.06	106800.7	106830.7	-3.9	21	(-0.5, 3)	$GA^{-0.5} + GA^3$	0.06	106795.1	106825.1	1.7
4	(-2, 0)	$GA^{-2} + \ln GA$	0.06	106798.2	106828.2	-1.4	22	(0, 0)	$\ln GA + \ln GA \ln GA$	0.06	106796.3	106826.3	0.5
5	(-2, 0.5)	$GA^{-2} + GA^{0.5}$	0.06	106796.3	106826.3	0.5	23	(0, 0.5)	$\ln GA + GA^{0.5}$	0.06	106796	106826	0.8
6	(-2, 1)	$GA^{-2} + GA$	0.06	106795.1	106825.1	1.7	24	(0, 1)	$\ln GA + GA$	0.06	106795.7	106825.7	1.1
7	(-2, 2)	$GA^{-2} + GA^2$	0.06	**106794.6**	**106824.6**	2.2	25	(0, 2)	$\ln GA + GA^2$	0.06	106795.4	106825.4	1.4
8	(-2, 3)	$GA^{-2} + GA^3$	0.06	106796.1	106826.1	0.7	26	(0, 3)	$\ln GA + GA^3$	0.06	106795.3	106825.3	1.5
9	(-1, -1)	$GA^{-1} + GA^{-1}\ln GA$	0.06	106800	106830	-3.2	27	(0.5, 0.5)	$GA^{0.5} + GA^{0.5}\ln GA$	0.06	106795.9	106825.9	0.9
10	(-1, -0.5)	$GA^{-1} + GA^{-0.5}$	0.06	106798.4	106828.4	-1.6	28	(0.5, 1)	$GA^{0.5} + GA$	0.06	106795.8	106825.8	1
11	(-1, 0)	$GA^{-1} + \ln GA$	0.06	106797.1	106827.2	-0.3	29	(0.5, 2)	$GA^{0.5} + GA^2$	0.06	106795.7	106825.7	1.1
12	(-1, 0.5)	$GA^{-1} + GA^{0.5}$	0.06	106796.1	106826.2	0.7	30	(0.5, 3)	$GA^{0.5} + GA^3$	0.06	106795.6	106825.6	1.2
13	(-1, 1)	$GA^{-1} + GA$	0.06	106795.4	106825.4	1.4	31	(1, 1)	$GA + GA\ln GA$	0.06	106795.9	106825.9	0.9
14	(-1, 2)	$GA^{-1} + GA^2$	0.06	106794.8	106824.8	2	32	(1, 2)	$GA + GA^2$	0.06	106796	106826	0.8
15	(-1, 3)	$GA^{-1} + GA^3$	0.06	106795.1	106825.2	1.7	33	(1, 3)	$GA + GA^3$	0.06	106796.1	106826.1	0.7
16	(-0.5, -0.5)	$GA^{-0.5} + GA^{-0.5}\ln GA$	0.06	106797.5	106827.5	-0.7	34	(2, 2)	$GA^2 + GA^2\ln GA$	0.06	106796.7	106826.7	0.1
17	(-0.5, 0)	$GA^{-0.5} + \ln GA$	0.06	106796.7	106826.7	0.1	35	(2, 3)	$GA^2 + GA^3$	0.06	106797.4	106827.4	-0.6
18	(-0.5, 0.5)	$GA^{-0.5} + GA^{0.5}$	0.06	106796.1	106826.1	0.7	36	(3, 3)	$GA^3 + GA^3\ln GA$	0.06	106799.1	106829.1	-2.3

DD - Deviance difference.

Table 18: GLMs Parameter estimates

Outcome (n)	Par.	Model 1			Model 2			Model 3		
		RR	**LCI**	**UCI**	**RR**	**LCI**	**UCI**	**RR**	**LCI**	**UCI**
Maternal										
Hellp Preeclampsia	b_1	0.9286*	0.8590	0.9924				0.9266*	0.8583	0.9893
(7)	b_2				0.1522	0.0019	12.9372	0.0657	0.0006	8.9801
Vaginal bleeding	b_1	0.9625*	0.9260	0.9982				0.9602*	0.9238	0.9963
(21)	b_2				0.6815	0.0608	7.3833	0.3862	0.0289	4.8407
Placenta praevia	b_1	0.9925	0.9458	1.0383				0.9885	0.9415	1.0352
(12)	b_2				0.3176	0.0129	7.6849	0.2620	0.0090	7.0177
Uterine rupture (2)	b_1	0.9912	0.8736	1.1048				0.9984	0.8777	1.1167
	b_2				26.7918	0.0123	37407.1077	26.4913	0.0100	38072.4806
Placental abruptio	b_1	0.9996	0.9567	1.0419				0.9916	0.9486	1.0349
(14)	b_2				0.1049	0.0057	2.0457	0.0891	0.0041	1.9366
Neonatal										
Fetal Anaemia (3)	b_1	0.9773	0.8820	1.0661				0.9518	0.8682	1.0365
	b_2				0.0002*	0.0000	0.0984	0.0001*	0.0000	0.0433
Fetal death (2)	b_1	0.9953	0.8784	1.1089				0.9573	0.8669	1.0700
	b_2				0.0002*	0.0000	0.2640	0.0001*	0.0000	0.1385
Seizures (4)	b_1	1.0113	0.9346	1.0864				1.0080	0.9317	1.0840
	b_2				0.1628	0.0005	54.9181	0.1827	0.0005	58.4440

b_1 - Random intercept, b_2 - Random Slope, **RR** - Risk ratio, **LCI** - Lower Confidence Interval, **UCI** - Upper Confidence Interval, * - significant at 5%, Model 1 - with random intercept as covariate, Model 2 - with random slope, Model 3 - with both random intercept and slope.

R-code

```
#----------packages
library("tidyverse")
library(dplyr)
library(lme4)
library(lattice)
library(ggplot2)
library(gridExtra)
library(gam)
library(mfp)
library(latticeExtra)
library(msm)
library(sandwich)
setwd("C:\\Users\\HP\\Desktop\\2nd yearUHasselt 2nd semester\\
Thesis materials\\Thesis Data")
interbio<-read.csv("INTERBIO data.csv", header = TRUE) #load data
dim(interbio)#13396    94

#-----------------------------------------------------------------
#                       EDA
#-----------------------------------------------------------
interb1<-interbio
interb1$mse_maternal_dob<-as.Date(interb1$mse_maternal_dob,
format = "%d/%m/%Y")
interb1$pfu_visit_date<-as.Date(interb1$pfu_visit_date,
format = "%d/%m/%Y")
```

```
length(unique(interb1$ptid)) #total number of women = 2820

#-----------------------------------------------------------------
#                     Overall Profile and Mean Plots
#-----------------------------------------------------------
#-----------------------------------------------------profiles
set.seed(99999)
#random sample of data
interb1_120<-interb1[interb1$ptid %in%sample(unique(interb1$ptid), 120,
replace = FALSE),]
ggplot(interb1_120, aes(x= gestweeks, y= pw)) +
  geom_line(aes(group = ptid), colour="darkslategray")+
  labs(x="Gestational Age (weeks)", y="Pregnancy Weight (Kgs)")+
  theme_classic()+
  theme(
    axis.title.x = element_text(size = 14, face = "bold"),
    axis.title.y = element_text( size = 14, face = "bold")
  )
#overall mean evolution
ggplot(interb1, aes(gestweeks, pw))+
  geom_point(color="darkgray")+
  #geom_point(aes(colour=country, group = country))+
  geom_point(aes(colour=country))+
  geom_smooth(se=FALSE, method = "loess",size=1.2)+
  labs(x="Gestational Age (weeks)", y="Observed Weight (Kgs)")  +
  theme_classic()+
  theme(
    axis.title.x = element_text(size = 14, face = "bold"),
    axis.title.y = element_text( size = 14, face = "bold"),
    legend.text = element_text(size=20),
    legend.title = element_text(size=22),
    legend.key.size = unit(1, 'cm')
  )+
  guides(color = guide_legend(override.aes = list(size = 5)))

#mean evolution by country
TheColors <- c("black","red","green","purple","orange","brown","blue")
ggplot(interb1, aes(gestweeks, pw, colour=country)) +
  #geom_point(color="gray") +
  geom_smooth(method = "loess", se = FALSE,size=1) +
  geom_smooth(aes(group = 1, col = " Overall"), method = "loess",
  se = FALSE,size=1.2)+ #, span = 0.07
  scale_color_manual("country", values = TheColors)+
  labs(x="Gestational Age (weeks)", y="Pregnancy Weight (Kgs)")  +
  theme_classic()+
  theme(
    axis.title.x = element_text(size = 14, face = "bold"),
    axis.title.y = element_text( size = 14, face = "bold"),
```

```
    legend.text = element_text(size=20),
    legend.title = element_text(size=22),
    legend.key.size = unit(1, 'cm')
  )+
  guides(color = guide_legend(override.aes = list(size = 3)))

#.summaries for maternal characteristics by country
interb2upxrt<- interb1 %>% group_by(country,ptid) %>%
  summarise(n=n(),
            minwg=min(wg),
            maxwg=max(wg),
            age=(as.numeric(last(as.Date(pfu_visit_date))-
            first(as.Date(mse_maternal_dob))))/365,
            sysBP=mean(mse_15_bp_systolic)
            #...+ other variables in the results
  ) %>%
  as.data.frame()
interb2upxrt %>% group_by(country) %>%
  summarise(n=n(),
            min=min(minwg),
            max=max(maxwg),
            mean_age=mean(age, na.rm = TRUE),
            sd_age=sd(age, na.rm = TRUE),
            min_age=min(age, na.rm = TRUE),
            max_age=max(age, na.rm = TRUE),
            md_age=median(age, na.rm = TRUE),
            mean_sysBP=mean(sysBP, na.rm = TRUE)
            #...+ other variables in the results
  )%>%
  as.data.frame()
#Global maternal xrts
interb2upxrtGlobal<- interb1 %>% group_by(ptid) %>%
  summarise(n=n(),
            minwg=min(wg),
            maxwg=max(wg),
            age=(as.numeric(last(as.Date(pfu_visit_date))-
            first(as.Date(mse_maternal_dob))))/365,
            sysBP=mean(mse_15_bp_systolic)
            #...+ other variables in the results
  ) %>%
  as.data.frame()
interb2upxrtGlobal %>% #group_by(country) %>%
  summarise(n=n(),
            min=min(minwg),
            max=max(maxwg),
            mean_age=mean(age, na.rm = TRUE)
            #...+ other variables in the results
  )%>%
```

```
  as.data.frame()
#-------------------------------------------------------------------------#
#                      frequencies for maternal characteristics           #
#-------------------------------------------------------------------------#
interb1<-interb1[interb1$country=="UK",]
length(unique(interb1$ptid))#623
#level of education
interb_edu<- interb1 %>%
group_by(ptid, mse_03_highest_level_education) %>%
  summarise(n=n()) %>%
  as.data.frame()
interb_edu %>% group_by(mse_03_highest_level_education) %>%
  summarise(tot_wom=n())%>%
  as.data.frame()
#...same code for other maternal characteristics

#Fractional polynomials
fp_model <- mfp(pw~fp(gestweeks, df=4, select = 0.5),
                family=gaussian, data=interb1, verbose = TRUE)
#or
interb1$p1<- (interb1$gestweeks)^(-2)
fp1_1<-lm(pw ~ p1, data = interb1); summary(fp1_1)
#...same procedure for other FP1s
interb1$p2_2a<- (interb1$gestweeks)^(-2)
interb1$p2_2b <- (interb1$gestweeks)^(-1)
fp2_2<-lm(pw ~ p2_2a + p2_2b, data = interb1); summary(fp2_2)
#...same procedure for other FP2s

#FP plots
interb111<-subset(interb1, select = c("pw","gestweeks"))
## Modify dataset
interb111<- within(interb111, {
  GESTWEEKS <- gestweeks
  PW <- pw
  ## Category indicator variable
  indicator <- cut(gestweeks, breaks = c(-Inf,10,20,30,40,50,Inf),
  labels = c("i1","i2","i3","i4","i5","i6"))
  ## Variables for spline
  s1        <- pmax(gestweeks -  0, 0)
  s2        <- pmax(gestweeks - 10, 0)
  s3        <- pmax(gestweeks - 20, 0)
  s4        <- pmax(gestweeks - 30, 0)
  s5        <- pmax(gestweeks - 40, 0)
  s6        <- pmax(gestweeks - 50, 0)
})
## Plot raw data as a scatter plot
ggplot(data = interb111,
       mapping = aes(x = GESTWEEKS, y = PW)) +
```

```
  geom_point()+
  scale_x_continuous(breaks = seq(from = 0, to = 50, by = 5)) +
  theme_bw() +
  theme(legend.key = element_blank())
## Define a helper function for ggplot2
stat_fun <- function(fun, args, mapping) {
  stat_function(fun = fun, args = args, mapping = mapping, n = 13396,
  size = 1.3, alpha = 1)
}
## Define prediction function
PredFun <- function(GESTWEEKS, model) {
  ## Category indicator variable
  indicator <- cut(GESTWEEKS, breaks = c(-Inf,10,20,30,40,50,Inf),
  labels = c("i1","i2","i3","i4","i5","i6"))
  ## Variables for spline
  s1        <- pmax(GESTWEEKS -  0, 0)
  s2        <- pmax(GESTWEEKS - 10, 0)
  s3        <- pmax(GESTWEEKS - 20, 0)
  s4        <- pmax(GESTWEEKS - 30, 0)
  s5        <- pmax(GESTWEEKS - 40, 0)
  s6        <- pmax(GESTWEEKS - 50, 0)
  ## Predict
  predict(object = model, newdata = data.frame(GESTWEEKS = GESTWEEKS,
                                                indicator = indicator,
                                                s1 = s1, s2 = s2, s3 = s3,
                                                s4 = s4, s5 = s5, s6 = s6))
}
#--
## Fractional polynomial model1
lmFractional1 <- lm(formula = PW ~ GESTWEEKS,
                    data    = interb111)
## Fractional polynomial model2
lmFractional2 <- lm(formula = PW ~ I(GESTWEEKS^(0.5)), data = interb111)
## Fractional polynomial model3
lmFractional3 <- lm(formula = PW ~ GESTWEEKS + I(GESTWEEKS^(-2)) +
I(GESTWEEKS^(2)),
                    data    = interb111)
## Create plot base
plotBasePoly <- ggplot(data = interb111,
                       mapping = aes(x = GESTWEEKS, y = PW)) +
  #layer(geom = "point") +
  geom_point(color="darkgray")+
  scale_x_continuous(breaks = seq(from = 0, to = 50, by = 5)) +
  scale_color_manual(name = "Model", breaks =  c("FP(1)","FP(0.5)",
  "FP(-2,2)"),
                     values = c("FP(1)"= "red","FP(0.5)"= "blue",
                     "FP(-2,2)"= "green") ) +
  theme_classic() +
```

```
  theme(legend.key = element_blank())+
  theme(
    axis.title.x = element_text(size = 14, face = "bold"),
    axis.title.y = element_text( size = 14, face = "bold"),
    legend.text = element_text(size=20),
    legend.title = element_text(size=22),
    legend.key.size = unit(1, 'cm')
  )+
  guides(color = guide_legend(override.aes = list(size = 3)))+
  labs(x="Gestational Age (weeks)", y="Pregnancy Weight (Kgs)")
## Create layers
layerPoly1    <- stat_fun(fun = PredFun,
args = list(model = lmFractional1),  mapping = aes(color = "FP(1)"))
layerPoly2  <- stat_fun(fun = PredFun,
args = list(model = lmFractional2),  mapping = aes(color = "FP(0.5)"))
layerPoly3 <- stat_fun(fun = PredFun,
args = list(model = lmFractional3), mapping = aes(color = "FP(-2,2)"))
## Plot together
plotBasePoly + layerPoly1 + layerPoly2 + layerPoly3

#making intercept meaningful in the LMM
#-------------------- centering around min at the first visit
lmm_dat<-subset(interb1, select = c("ptid","pw","mse_07_weight",
"gestweeks","ga_at_delivery"))
#View(lmm_dat)
data_c<-lmm_dat  %>%
  mutate(gestweeks_c=gestweeks - 9.1429)%>% #min=9.1429
  as.data.frame()

#-----linear mixed model
lmm1<-  lmer(pw ~ gestweeks_c + (gestweeks_c | ptid), data_c,
             control = lmerControl(calc.derivs = FALSE))
library("lmerTest")
summary(lmm1); anova(lmm1)
##---extracting estimates of random effects per id
random_est <-as.data.frame(ranef(lmm1)$ptid)
colnames(random_est) <- c("intercept", "slope")

#plot of slope and intercept
interb2<- interb1 %>% group_by(ptid) %>%
  summarise(bmic=first(BMIcat),
            country=first(country)) %>%
  as.data.frame()
rand_bmi<-cbind(interb2, random_est)

ggplot(na.omit(rand_bmi), aes(x= intercept, y= slope)) +
  geom_point(aes(color= factor(bmic))) +
  labs(x="Random Intercept", y="Random Slope",color="BMI Category")+
```

```
  theme_classic()+
  theme(
    axis.title.x = element_text(size = 14, face = "bold"),
    axis.title.y = element_text( size = 14, face = "bold"),
    legend.text = element_text(size=20),
    legend.title = element_text(size=22),
    legend.key.size = unit(1, 'cm')
  )+
  guides(color = guide_legend(override.aes = list(size = 5)))+
  annotate(x=40, y=-0.2,
           label=paste("r = ", round(cor(rand_bmi$intercept,
           rand_bmi$slope),2)),
           geom="text", size=6)
#histogram by country
ggplot(rand_bmi, aes(intercept, fill = country)) +
  geom_histogram(alpha = 0.5, aes(y = ..density..),
  position = 'identity', bins = 300)+
  theme_classic()+
  theme(
    axis.title.x = element_text(size = 14, face = "bold"),
    axis.title.y = element_text( size = 14, face = "bold"),
    legend.text = element_text(size=20),
    legend.title = element_text(size=22),
    legend.key.size = unit(1, 'cm')
  )+
  guides(color = guide_legend(override.aes = list(size = 1)))
#-----------------------------------------------
#plot of fitted model overlay raw data

#prediction and observed
interb1$pred2 <- predict(lmm1,re.form=NA)## population level
d<-ggplot(interb1, aes(gestweeks, pw))+
  geom_point(color="darkgray")+
  geom_smooth(se=FALSE, method = "loess",size=1.2,
  aes(color = "Mean Evolution") )+
  labs(x="Gestational Age (weeks)", y="Pregnancy Weight (Kgs)")  +
  theme_classic()+
  theme(
    axis.title.x = element_text(size = 14, face = "bold"),
    axis.title.y = element_text( size = 14, face = "bold"),
    legend.text = element_text(size=20),
    legend.title = element_text(size=22),
    legend.key.size = unit(1, 'cm')
  )+
  guides(color = guide_legend(override.aes = list(size = 3)))

d + geom_line( aes(y=pred2, color="Fitted Line (LMM)"), size=1)+
  scale_color_manual(name = "Model fit",
```

```
                    breaks = c("Mean Evolution", "Fitted Line (LMM)"),
                    values = c("Mean Evolution" = "blue",
                    "Fitted Line (LMM)" = "red"))

#GLMs
interb1_tics<- interb1 %>%
  select(2,8:16,18:52,61:65,67:76,88,91:94,78)%>%
  group_by(ptid) %>%
  summarise_all(first) %>%
  as.data.frame()
#merge with random_est
finaldata1<-cbind(random_est, interb1_tics)
#PIH
finaldata1$dev_085_preg_induce_hypertension<-
factor(finaldata1$dev_085_preg_induce_hypertension)
bin_log4<-glm(dev_085_preg_induce_hypertension ~ intercept, finaldata1,
              family = binomial(link="log"), na.action = na.exclude)
              summary(bin_log4)
RR4 <-exp(cbind(RR = coef(bin_log4), confint(bin_log4)))
round(RR4, digits=4)
#...same procedure used for other GLMs
#---below_C10_HC
library(sandwich)
bin_log24c2<-glm(below_C10_BW ~ intercept + slope , finaldata1,
                 family = poisson(link="log"))
cov.bin_log24c2 <- vcovHC(bin_log24c2, type="HC0")
std.err <- sqrt(diag(cov.bin_log24c2))
r.est <- cbind(Estimate= coef(bin_log24c2), "Robust SE" = std.err,
               "Pr(>|z|)" = 2 * pnorm(abs(coef(bin_log24c2)/std.err),
               lower.tail=FALSE),
               LL = coef(bin_log24c2) - 1.96 * std.err,
               UL = coef(bin_log24c2) + 1.96 * std.err)
rexp.est <- exp(r.est[, -3]); round(rexp.est, digits=4)
#-------------predicted probabilities
predicted_probs24c<-as.data.frame(predict(bin_log24c2, type = "response",
se.fit=TRUE))
inter_data24c<-cbind(finaldata1$ptid,predicted_probs24c,
finaldata1$intercept, finaldata1$slope)
#rename columns
colnames(inter_data24c)<-c("ptid","prob","serrors", "residualscale",
"intercept","slope")
data_bin_log24c2 <-mutate(inter_data24c, Relative = prob / lag(prob))

ggplot(data_bin_log24c2, aes(x = slope, y = Relative))+
  geom_smooth(se=TRUE, method = "loess",size=1)+
  labs(x="Random Slope", y="Predicted Relative Risk of Below C10 BW")+
  theme_classic()+
  theme(
```

```
    axis.title.x = element_text(size = 14, face = "bold"),
    axis.title.y = element_text( size = 14, face = "bold")
  )

#Adjusting for other covariates
#same covariates for all outcomes
#preterm labour
bin_log8d<-glm(dev_089_preterm_labour ~ intercept + slope+ mage +
                 dev_054_any_babies_preterm +mse_15_bp_systolic+
                 dev_056_any_neonatal_deaths
               +dev_158_prev_caesarean_section+ country,
               finaldata1,family=poisson(link="log"),
               na.action = na.exclude)
cov.bin_log8d<- vcovHC(bin_log8d, type="HC0")
std.err <- sqrt(diag(cov.bin_log8d))
r.est <- cbind(Estimate= coef(bin_log8d), "Robust SE" = std.err,
               "Pr(>|z|)" = 2 * pnorm(abs(coef(bin_log8d)/std.err),
               lower.tail=FALSE),
               LL = coef(bin_log8d) - 1.96 * std.err,
               UL = coef(bin_log8d) + 1.96 * std.err)
s <- deltamethod(list(~ exp(x1), ~ exp(x2), ~ exp(x3), ~ exp(x4),
~ exp(x5), ~ exp(x6)
                      , ~ exp(x7), ~ exp(x8), ~ exp(x9), ~ exp(x10),
                      ~ exp(x11), ~ exp(x12)
                      , ~ exp(x13)),
                 coef(bin_log8d), cov.bin_log8d)
rexp.est <- exp(r.est[, -3])
rexp.est[, "Robust SE"] <- s
round(rexp.est, digits=2)
```

SAS code

```
data transformed;
set sorted_ptid;
by ptid;
*--;
retain GA_start;
retain obs 1;
ln_weight = log(pfu_01_weight+1);
if first.ptid=1 then do;
time=0;
GA_start=gestweeks;
obs=1;
end;
 else do;
time= gestweeks-GA_start;
obs+1;
end;
```

```
 drop GA_start;
run;
/*** Variance Structure ***/
/*Calculating residuals of the smoothed curve*/
proc glm data=transformed;
model pfu_01_weight=gestweeks;
output out=residuals r=residuals;
run;
data squared_residuals_overall;
set residuals;
res_square = residuals**2;
run;

/*Obtaining smoothed squared residuals vs GA curve*/
proc loess data=squared_residuals_overall;
ods output OutputStatistics=out4;
model res_square=gestweeks;
run;
proc sort data=out4 out=smoothed_residuals_sorted;
    by gestweeks;
run;proc print data=smoothed_residuals_sorted;run;
/*Plottig smoothed residuals function*/
/* Overall */
goptions reset=all ftext=swiss device=psepsf gsfname=fig1 gsfmode=replace
rotate=landscape;
proc gplot data=smoothed_residuals_sorted;
plot DepVar*gestweeks=1 Pred*gestweeks=2 / overlay haxis=axis1 vaxis=axis2;
symbol1 c=grey v=dot h=0.5 mode=include;
symbol2 c=black i=join w=2 mode=include;
axis1 label=(h=1.5 'Gestational Age (weeks)') value=(h=1)
minor=none;
axis2 label=(h=1.5 A=90 'Squared Residuals') value=(h=1)
minor=none;
run;
quit;
```