Examining Gestational Weight Gain and Perinatal Outcomes in Twin versus Singleton Pregnancies

Michelle Celia Dimitris,

Epidemiology, Biostatistics, and Occupational Health,

McGill University, Montreal

September 2019

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy (PhD) Epidemiology.

Table of Contents

Abstract

Gestational weight gain is a modifiable characteristic of pregnancy that is rarely studied in twins. Among singleton pregnancies, gestational weight gain is correlated with perinatal outcomes, such as gestational diabetes, preterm birth, small/large for gestational age and Caesarean delivery, as well as long-term maternal and child weight. Although twin pregnancies have increased in recent years and are at higher risk for many of these outcomes, maternal weight gain is infrequently studied in this population. Moreover, the time-varying nature of gestational weight gain and its inherent connection to gestational duration confer unique methodological issues to its study. Objectives of my dissertation were threefold: (1) Evaluate the accuracy and precision of methods for estimating maternal weight gain between prenatal visits; (2) Investigate the association between time-varying gestational weight gain and time-to-delivery; (3) Determine the extent to which increased risk of gestational diabetes in twin pregnancies is mediated by gestational weight gain. I use serial weight measurements and gestational diabetes screening information abstracted from medical charts, as well as maternal/pregnancy characteristics obtained from a hospital-maintained database, for a large cohort and case-cohort of twin and singleton pregnancies, respectively, delivered at Magee-Women's Hospital in Pittsburgh Pennsylvania. For Objective 1, I compare both individual-level and pooled methods with varying degrees of flexibility and parametricity; I conclude that both individual-level and pooled methods with high flexibility and low parametricity performed best. For Objective 2, I examine the relationship between time-varying gestational weight gain z-score and time-to-delivery using a flexible extension of the Cox proportional hazards model; I conclude that relatively low gestational weight gain is associated with higher risk for early preterm spontaneous delivery, while relatively high gestational weight gain is associated with higher risk for late preterm no labour delivery. For Objective 3, I combine twin and singleton data, and use causal mediation analyses to disentangle pathways between plurality (i.e. twin vs singleton) and gestational diabetes that are mediated/not mediated by increased gestational weight

gain; I conclude that the relationship between plurality and gestational diabetes is primarily mediated by pathways not related to gestational weight gain. I anticipate that the results of my dissertation will shed light on the relationship between gestational weight gain and perinatal outcomes in twin pregnancies, particularly how it resembles/differs from that in singletons. Broadly, I expect that my findings will have implications for gestational weight guidelines, especially among twin pregnancies for which guidelines are currently provisional.

Résumé

La prise de poids est une caractéristique modifiable de la grossesse rarement étudiée chez les grossesses gémellaires. Parmi les grossesses uniques, la prise de poids gestationnelle est corrélée aux complications périnatales, telles que le diabète gestationnel, les naissances prétermes, les néonates petits / grands pour l'âge gestationnel et l'accouchement par césarienne, ainsi que le poids à long terme de la mère et de l'enfant. Bien que les grossesses gémellaires aient augmenté au cours des dernières années et courent un risque plus élevé pour beaucoup de ces issues, la prise de poids maternelle est rarement étudiée dans cette population. De plus, la nature variable du gain de poids pendant la grossesse et son lien inhérent avec la durée gestationnelle confèrent à son étude des problèmes méthodologiques uniques. Les objectifs de ma thèse étaient les suivants: (1) Évaluer l'exactitude et la précision des méthodes d'estimation du gain de poids maternel entre les visites prénatales; (2) Étudier le lien qui existe entre le gain de poids pendant la grossesse et l'âge gestationnel à l'accouchement; (3) Déterminer dans quelle mesure le risque accru de diabète gestationnel lors de grossesses gémellaires est fonction du gain de poids gestationnel. J'utilise des mesures de poids en série et des informations de dépistage du diabète gestationnel extraites de dossiers médicaux, ainsi que des caractéristiques maternelles / de grossesse obtenues à partir d'une base de données gérée par un hôpital, pour une grande cohorte et une cohorte de grossesses gémellaires et simples, respectivement, accouchées à Magee- Hôpital des femmes à Pittsburgh en Pennsylvanie. Pour l'objectif 1, je compare à la fois des méthodes individuelles et des méthodes combinées avec divers degrés de flexibilité et de paramétrie; je conclus que les méthodes combinées au niveau individuel, avec une grande flexibilité et un faible paramétrie, ont donné les meilleurs résultats. Pour l'objectif 2, j'examine la relation entre le score z du gain de poids gestationnel, qui varie avec le temps, et le délai avant l'accouchement à l'aide d'une extension flexible du modèle des risques proportionnels de Cox; je conclus que le gain de poids gestationnel relativement faible est associé à un risque plus élevé d'accouchement prématuré, tandis

qu'un gain de poids gestationnel relativement élevé est associé à un risque plus élevé d'accouchement prématuré tardif sans travail. Pour l'objectif 3, je combine des données simples et utilise des analyses de médiation causale pour démêler les voies entre la pluralité (gémellaires vs uniques) et le diabète gestationnel qui sont médiées / ne sont pas médiées par un gain de poids accru pendant la grossesse; je conclus que la relation entre la pluralité et le diabète gestationnel repose principalement sur des voies non liées à la prise de poids pendant la grossesse. Je prévois que les résultats de ma thèse éclairciront la relation entre le gain de poids gestationnel et les conséquences périnatales chez les grossesses gémellaires, en particulier en quoi elles ressemblent / diffèrent de celles des grossesses uniques. En gros, je m'attends à ce que mes conclusions aient des conséquences pour les recommandations relatives au gain de poids gestationnel, en particulier pour les grossesses gémellaires pour lesquelles les recommandations sont actuellement provisoires.

Acknowledgements

I wish to thank my committee members, Dr. Jennifer Hutcheon and Dr. Katherine Himes, as well as Dr. Lisa Bodnar, Dr. Michal Abrahamowicz, and Dr. Marie-Eve Beauchamp, for their ideas, feedback, and support. I thank Sara Parisi for her technical support of all manuscripts and Dr. Yishu Wang for her technical support of my second manuscript, and in particular her development of the flexible Cox proportional hazards model. I extend a special thanks to my doctoral supervisors, Dr. Jay Kaufman and Dr. Robert Platt, for their feedback, mentorship, and support, as well as a heartfelt thank you to my family and friends for their constant love and encouragement. Funding sources that supported my doctoral training were: Graduate Award (Institute for Health and Social Policy, 2016), Alexander McFee and Graduate Excellence Fellowships (McGill University Faculty of Medicine, 2016-2017), Ferring Fellowship (McGill University Faculty of Medicine, 2017-2018), Fonds de Recherche Santé Doctoral Award (Government of Quebec, 2018-2020), and stipend support from National Institutes of Health R01-NR014245 grant. Support for conference travel include Travel Award – Institute Community Support (CIHR 2019) and Graduate Research Enhancement and Travel Award (McGill University Department of Epidemiology, Biostatistics, and Occupational Health, 2019).

Contribution to Original Knowledge

Manuscript 1: Evaluates several methods for interpolating gestational weight gain between measurements in twin and singleton pregnancies and assesses whether accuracy/precision vary by characteristics gestational age and total number of visits.

Manuscript 2: Examines the relationship between time-varying gestational weight gain z-score and gestational age at birth using a survival analysis framework. Analyses are conducted using a novel extension of the Cox proportional hazards model, which incorporates both time-dependent effects and non-linear relationships.

Manuscript 3: Investigates the extent to which increased risk of gestational diabetes in twin pregnancies is mediated by increased gestational weight gain. Analyses incorporate twin and singleton pregnancies from the same study population to elucidate causal relationships.

Contribution of Authors

Manuscript 1:

Dimitris MC, Hutcheon JA, Platt RW, Himes KP, Bodnar LM, Kaufman JS.

MD and JH proposed the current study; JH and LB lead the parent studies; MD designed study, conducted analysis, and wrote manuscript; JH, RP, KP, LB, JK provided feedback; All authors read and approved the final study.

Manuscript 2:

Dimitris MC, Hutcheon JA, Platt RW, Abrahamowicz M, Beauchamp ME, Himes KP, Bodnar LM, Kaufman JS.

MD and JK proposed the current study; JH and LB lead the parent studies; MD designed study, conducted analysis and wrote manuscript; MA and MB provided analytical support for the extended survival analysis model. JH, RP, KP, LB, JK provided feedback; All authors read and approved the final study.

Manuscript 3:

Dimitris MC, Kaufman JS, Bodnar LM, Platt RW, Himes KP, Hutcheon JA.

JH proposed the current study; JH and LB lead the parent studies; MD designed study, conducted analysis, and wrote manuscript; JK, LB, RP, KP, JH provided feedback; All authors read and approved the final study.

1. Introduction

Gestational weight gain is a risk factor for maternal and infant health yet is infrequently studied in twin pregnancies. Among singletons, both inadequate and excessive gestational weight gain are correlated with maternal and infant health outcomes[1,2]; however, many studies either improperly or imprecisely study gestational weight gain as a cumulative total or average rate[3]. Methodologically rigorous studies are needed for refinement of gestational weight gain guidelines in pregnancy. A primary goal of my dissertation was to investigate the relationship between gestational weight gain and perinatal outcomes in twin pregnancies; a secondary goal was to compare the relationship with that observed among singleton pregnancies.

1.1. Objectives

1. Evaluate methods for interpolating gestational weight gain between measurements in twin and singleton pregnancies.
2. Describe the relationship between time-varying gestational weight gain and gestational age at birth in twin pregnancies and compare to that among singletons.
3. Determine the extent to which increased risk of gestational diabetes in twin pregnancies is mediated by gestational weight gain.

To achieve these objectives, I analyzed previously-collected data for a large cohort of twin and case-cohort of singleton pregnancies delivered between 1998 and 2013 at Magee Women's Hospital in Pittsburgh, Pennsylvania. I used advanced statistical methods, including extended survival analysis that incorporates both time-dependent and non-linear effects, and causal mediation analyses to elucidate relationships and mechanisms involving gestational weight gain. I anticipate that the results of my dissertation will inform gestational weight gain guidelines for twin pregnancies, as well as support methodologically-rigorous research in this domain.

2. Literature Review

Twin Pregnancies are Increasingly Common

Twin pregnancies have substantially increased in the United States, Canada, and similar settings in recent decades[4,5]. In the United States in 2016, 33.4 per 1000 live births, or 131,723 total live births, were twins[4]. Twin births in this country have nearly doubled since 1971 due to increases in both maternal age and use of fertility treatment[6]. Studies have also described increased risk of dizygotic twin pregnancy associated with increased pre-pregnancy body mass index (BMI)[7], even prior to the advent of assisted reproductive technology[8]. Although the rate of twin births has somewhat stabilized since 2006[9], maternal age at pregnancy and pre-pregnancy BMI continue to rise[4,10]. Proportions of births among mothers aged 30 – 49 have steadily increased, while those among mothers aged 15 – 29 continue to decrease[4]. Additionally, a United States-based study found that prevalence of pre-pregnancy obesity increased from 17.6% in 2003 to 20.5% in 2009[10]. Advanced maternal age is associated with higher rates of twin pregnancy[11] caused by increased likelihood of multiple follicle development per cycle[12] and use of fertility treatment[6]; it follows that twin pregnancies will continue to comprise a meaningful proportion of births. The Society for Maternal Fetal Medicine has called for greater inclusion of twin pregnancies in studies investigating causes or predictors of perinatal outcomes[13].

Twin Pregnancies are at Higher Risk of Adverse Perinatal Outcomes

Compared to singletons, twin pregnancies experience higher risk of maternal and infant morbidity[14,15]. Twin pregnancies are at increased risk for gestational hypertension, gestational diabetes, Caesarean delivery, postpartum hemorrhage, and mortality among mothers, as well as cardiac defects, preterm birth, low birth weight, cerebral palsy, and mortality among infants[15,16]. In the United States in 2016, approximately 60% twins and 8% of singletons were born before 37 weeks' gestational age (GA), which exceeds a seven-fold increased risk of preterm birth[4]. It is estimated that twin pregnancies

are responsible for 15 – 20% of preterm births[17]. Additionally, studies of glucose intolerance in pregnancy estimate a two-fold increased risk of gestational diabetes mellitus (GDM) in mothers of twin versus singleton pregnancies[18,19], although differences in GDM by plurality are not always observed[20,21].

Gestational Age at Birth is an Important Outcome in Twin and Singleton Pregnancies

Conventionally, GA at birth is categorized as preterm (<37 weeks' GA) and term (≥37 weeks' GA) birth. Studies further subdivide GA at birth as extreme prematurity (<28 weeks' GA), severe prematurity (28-31 weeks' GA), moderate prematurity (32-33 weeks' GA), and late preterm/near term (34-36 weeks' GA), early term (37-38 weeks' GA), late term (39-41 weeks' GA) and post term (≥42 weeks' GA)[17,22], which is consistent with a continuum of risk; namely, risk of infant morbidity/mortality decreases as GA approaches term and increases after term[23].

Prior research suggests that optimal GA at birth is slightly lower in twins than in singletons. A recent systematic review concluded that minimal risk of infant morbidity/mortality occurred at 36 weeks' GA for monochorionic and 37 weeks' GA for dichorionic twin pregnancies (chorionicity refers to number of placentae)[24]. Conversely, infant mortality is minimized at 38 weeks' GA among singleton pregnancies[25,26]. It is therefore possible that the conventional preterm birth definition has different clinical implications in twin and singleton pregnancies.

GA at birth itself is a clinically relevant outcome in both twin and singleton pregnancies and demonstrates similar patterns of risk; namely, infant morbidity/mortality decreases as GA approaches optimum and increases after optimum in both singleton and twin pregnancies[23–26]. In contrast to preterm birth, recent studies in singletons have investigated GA at birth as a continuous outcome[27]. Specifically, time-to-event analyses enable researchers to differentiate between subdivisions of preterm birth without unnecessary stratification or categorization; this may be more statistically efficient, especially when investigating small effects[27]. Importantly, this approach enables comparison of effects

between twins and singletons without assuming that the dichotomous outcome of preterm birth is an equally important endpoint in both pregnancies.

Glucose Intolerance is an Important Outcome in Twin and Singleton Pregnancies

Gestational diabetes is defined as glucose intolerance, namely low insulin production and/or high blood glucose, first recognized in pregnancy[28]. Associations between high blood glucose/GDM and poor perinatal outcomes are well-established[29,30]. A recent study found that GDM is associated with gestational hypertension/preeclampsia, Caesarean delivery, preterm birth, large for GA, neonatal hypoglycemia, and jaundice in singleton pregnancies, while GDM is associated with Caesarean delivery, preterm birth, large for GA, and jaundice, but not gestational hypertension/preeclampsia or neonatal hypoglycemia in twin pregnancies[31]. In contrast, a systematic review concluded that GDM in twin pregnancies is associated with neonatal intensive care unit admissions only, and not GA at birth, small for GA, large for GA, low Apgar score, respiratory distress, or neonatal hypoglycemia[32]. Among dichorionic twin pregnancies delivered at Magee Women's Hospital (MWH) in Pittsburgh, Pennsylvania, GDM is associated with preeclampsia but not Cesarean delivery, large for GA, small for GA, or neonatal intensive care unit admissions[33]. After pregnancy, researchers have found similar time from delivery to abnormal glucose tolerance in twin and singleton pregnancies, although these effects were not adjusted for maternal age[34]. It is possible that GDM as conventionally defined is not as clinically relevant due to increased metabolic requirement among twin pregnancies[35]. One study found an association between glycemic control and small-for-gestational-age among twin pregnancies with GDM, which supports the hypothesis that high blood glucose is indicative of greater metabolic requirement among twin pregnancies[36]. However, it is notable that GDM is correlated with several adverse maternal outcomes among both pluralities. Furthermore, conventional diagnostic criteria for GDM are in flux, even among singleton pregnancies. A recent study of glucose intolerance in singleton pregnancies found that sub-GDM hyperglycemia is associated with Caesarean delivery,

shoulder dystocia, and large for GA; the definition of GDM was broadened as a result[29]. Glucose intolerance may present a continuum of risk for poor perinatal outcomes thematically similar to that of GA at birth.

Gestational Weight Gain is a Modifiable Pregnancy Characteristic

Gestational weight gain (GWG), or maternal weight gain during pregnancy, is a modifiable risk factor for maternal and infant health. A Cochrane review of randomized controlled trials found that nutrition- and exercise-based interventions decrease the risk of excessive GWG by 20% and increase the risk of insufficient GWG by 14%[37]. Although evidence that supported a concurrent risk reduction of preterm birth is limited, it is notable that both excessive and insufficient GWG are associated with preterm birth[38]; thus, interventions that decrease risk of excessive and increase risk of insufficient GWG may decrease and increase risk of preterm birth, respectively, potentially resulting in null effect estimates of interventions on preterm birth. A Cochrane review of nutrition- and exercise-based interventions specifically aimed at preventing GDM found an average change in GWG of -0.89 kg as well as a 15% reduced risk of GDM[39]. More recently, a meta-analysis of 45 randomized-controlled trials (RCTs) found 44% and 38% reduced risk of GDM associated with diet and physical activity interventions, respectively[40], while a systematic review of 13 RCTs among mothers with overweight/obesity pre-pregnancy BMI found that physical activity interventions decreased both average GWG by -1.14 kg and risk of GDM by 29%[41]. It is notable that the Cochrane review of nutrition- and exercise-based interventions for prevention of GDM found a 20% decreased risk of preterm birth[39]. Studies aimed at preventing GDM typically enrol participants at high risk for this condition, including those with overweight or obese pre-pregnancy BMI. Since low GWG may be less harmful in this population[42,43], this effect might be attributed to decreased risk of high GWG; this is consistent with observational studies of GWG and preterm birth among pregnancies with overweight

and obese BMI[44]. Effects of nutrition and exercise-based interventions on GWG among twin pregnancies are unknown, as twins were generally excluded from these trials[37,39].

Studying Gestational Weight Gain is Methodologically Complex

It is important to carefully consider the exposure definition when investigating the effect of GWG on perinatal outcomes. For example, many studies characterize this exposure as total GWG in kilograms, or average rate of GWG in kilograms per week over the entire pregnancy. However, longer gestational duration typically results in an increased total GWG, even in absence of causal effect[3]. Likewise, shorter gestational duration necessitates an increased contribution of first trimester weight gain to the cumulative rate of GWG over pregnancy. Since GWG is minimal in the first trimester, cumulative rate of GWG over the entire pregnancy is similarly intertwined with gestational duration. These measures are inappropriate when the outcome is caused or defined by lower GA at birth[3]. For example, the association between GWG and preterm birth is overestimated when investigating either total GWG and cumulative rate of GWG[1].

Recent studies have responded to this issue by characterizing GWG as rate in second/third trimester only or in relation to Institute of Medicine (IOM) GWG guidelines[45,46]. Although this improves upon prior research, it may be considered an oversimplification of a time-varying exposure and doesn't allow for the investigation of patterns of GWG. Moreover, adequacy in relation to the IOM GWG guidelines is still linked to gestational duration if derived from total GWG. Standards for relative GWG by GA, stratified by pre-pregnancy body mass index (BMI), are available for both twin and singleton pregnancies[47,48], yet current studies of GWG rarely leverage z-score charts. Researchers are increasingly studying the impact of GWG trajectory on maternal and child health[27], and the IOM has called for more of these studies to inform future GWG guidelines[49].

Causal inference is challenging in studies of GWG, since GWG may be regarded as an ill-defined exposure. Specifically, GWG may be intervened upon in several ways, including by nutrition

and exercise interventions as previously described[37,39], which may take many forms. Moreover, change in weight during pregnancy is a proxy for changes in several maternal physiological characteristics, such as fat, protein, fat-free mass, and total body water, as well as fetal size, all of which may have different relationships with adverse perinatal outcomes[50]. Additionally, it has been estimated that as much as 46% of GWG may be due to genetic causes[51]. Caution must be exercised when inferring causality, particularly in regard to identifiability conditions of well-defined interventions and consistency[52]; in other words, increases or decreases in GWG caused by different mechanisms may not result in equally increased or decreased risks for perinatal outcomes. Nonetheless, maternal weight is routinely monitored during pregnancy, relatively easy to measure in both clinical and research settings, and practice guidelines have been established for its recommended range. Observational studies of GWG and perinatal outcomes are important to refine ranges for existing guidelines and inform future intervention-based studies of GWG and maternal and infant health.

Gestational Weight Gain is Associated with Preterm Birth and Gestational Diabetes in Singleton Pregnancies

Studies that avoid the methodological pitfalls described above have described relationships between GWG and both preterm birth and GDM in singleton pregnancies. For example, studies that related total GWG to an internal standard or reference found a u-shaped relationship between GWG and preterm birth[38,53]. Specifically, both low and high total GWG z-scores were associated with increased risk of preterm birth, either within normal/overweight pre-pregnancy BMI strata[53] or across pre-pregnancy BMI categories[38].

Similarly, when either total/rate of GWG in mid-pregnancy or total GWG z-score are investigated as exposures of interest, a monotonic relationship is consistently found between GWG and GDM in singletons[38,54–57]. Specifically, studies find 40% increase odds of GDM when comparing excessive to adequate GWG prior to glucose screening[54], 43 – 74% increased odds of GDM when

comparing higher to lower quantiles of GWG rate before glucose screening test[55], and 64 – 130% increased odds of GDM when comparing higher to lower quantiles of total GWG before 24 to 28 weeks' GA in singleton pregnancies[56]. Studies are less conclusive about the role of pre-pregnancy BMI in modifying this effect; while one study finds evidence for an inverse effect in underweight pre-pregnancy BMI stratum[38], and two studies find evidence for stronger effects within overweight/obese pre-pregnancy BMI strata[55,57], others simply adjust for pre-pregnancy BMI and/or find no evidence for effect modification[54,56]. Studies additionally conclude that effects are driven by GWG in the first trimester.

Gestational Weight Gain is Insufficiently Studied in Twin Pregnancies

Both total GWG and risk of adverse perinatal outcomes are higher in twin pregnancies, yet GWG is rarely studied in this population. A recent systematic review found only 15 studies investigating GWG in twin pregnancies, and these frequently suffered from the same methodological issues found in studies of GWG among singletons[58]. Among studies of GWG and preterm birth in twin pregnancies published since this review, 9 of 12 use either total or cumulative rate of GWG and are therefore uninterpretable[45,59–65]. Among remaining studies, one found that low rate of GWG in the second trimester was associated with increased risk of spontaneous preterm birth[45], one found that low GWG before but not after 22 weeks' GA was associated with preterm birth[66], and one found no association between rate of GWG in the second/third trimester and preterm birth[46]. While GDM is more common in twins, and GWG is associated with GDM in singletons, compelling evidence linking GWG and GDM in twin pregnancies does not exist[45]. Evidence-based guidelines for GWG in singletons are available, yet IOM GWG guidelines for twins are provisional and based on limited information[49]. Methodologically rigorous studies are needed to fill a gap in knowledge about the relationship between GWG and perinatal outcomes in twins, and how it compares to that in singleton pregnancies.

3. Manuscript 1

3.1. Title

Interpolating Gestational Weight Gain Between Clinical Visits in Twin and Singleton Pregnancies

3.2. Abstract

Gestational weight gain, a modifiable predictor of maternal and infant health, is often quantified as a single summary measure, such as the cumulative total or rate of weight gain throughout pregnancy. While researchers are increasingly interested in studying gestational weight gain as a longitudinal measure, advanced statistical methods that can be used for this purpose generally require estimation between prenatal weights, which are often measured weeks apart. We evaluated the relative accuracy and precision of methods for estimating maternal weight between prenatal visits among large cohorts of twin and singleton pregnancies. We used serial weights obtained by chart abstraction for dichorionic twin (n=2066) and singleton (n=7731) pregnancies delivered from 1998-2013 at Magee-Women's Hospital in Pittsburgh, Pennsylvania. We mimicked weight gain data typically available in pregnancy cohorts by retaining 3-4 weights per pregnancy, namely pre-pregnancy weight, delivery weight, as well as weights measured at first prenatal and glucose screening visits and used these to fit interpolation models. We estimated all remaining prenatal weights using interpolation methods and calculated the difference between measured weight gain and corresponding predicted weight gain in kilograms; we calculated root mean squared error (RMSE) for each model. We evaluated 16 methods that varied on characteristics including individual-level variation, flexibility, and parametricity. Among both twin and singleton pregnancies, the best methods incorporated restricted cubic splines, random intercepts and slopes for pregnancy, and internal knots demarcating trimesters (RMSE=1.55 kg, 1.45 kg) or quantiles (RMSE=1.56 kg, 1.45 kg). Individual-level linear interpolation using most proximal measurements also performed well (RMSE 1.62 kg among twins, 1.46 kg among singletons). Overall, RMSE ranged from 1.55 to 6.09 in twins and 1.45 to 4.87 in singletons, demonstrating that choice of

model is important in accurately and precisely predicting weight gain between measurements. Although the best interpolation methods differed in complexity, they demonstrated high degrees of individual-level variation and flexibility, and less parametricity.

3.3. Introduction

Gestational weight gain (GWG), or maternal weight gain during pregnancy, is frequently studied in relation to maternal and infant outcomes. Prior research suggests that GWG is associated with adverse perinatal outcomes including gestational diabetes, Caesarean section delivery, small/large for gestational age, and preterm birth, as well as with longer-term maternal and child obesity[2,44,67]. Nonetheless, Institute of Medicine (IOM) guidelines for GWG acknowledge a paucity of evidence for health outcomes among certain groups, notably twin pregnancies[49].

While maternal weight changes over the course of pregnancy, GWG is traditionally studied as a single cumulative exposure. Specifically, total or rate of GWG are often studied as risk factors for adverse outcomes, even though these summary measures are imprecise and inappropriate in situations where outcomes are dependent on gestational duration[3]. Researchers are increasingly motivated to investigate patterns or trajectories of maternal weight change during pregnancy. Furthermore, the IOM has recommended research about the impact of longitudinal measures of GWG on maternal and infant health[49].

Maternal weight is measured at prenatal visits, yet analytical techniques that incorporate time-varying covariates may require exposure measurements that extend beyond the finite number of observed measurements. Applying certain methods, such as time-to-event analyses, to GWG research may require interpolation between observed weight measurements; we are unaware of any gold standard for such estimation over the course of pregnancy or within twins. A recent study examined the accuracy and precision of two flexible methods for estimating GWG at 28 weeks' GA and 40 weeks' GA; however, this study did not incorporate first trimester GWG, nor did investigators evaluate

method performance across the full range of gestational duration[68]. Importantly, this study did not evaluate interpolation methods in twin pregnancies, in which studies of GWG and maternal/infant health are urgently needed[58]. We therefore aimed to evaluate methods for interpolating maternal weight during pregnancy between observed measurements among both twin and singleton pregnancies, as well as investigate whether precision of methods differed by gestational age or mothers' total number of observed measurements.

3.4. Methods

Study Population

All pregnancies within a previously-conducted cohort study of diamniotic twin and case-cohort study of singleton pregnancies delivered between 1998-2013 and 1998-2011, respectively, at Magee Women's Hospital in Pittsburgh, Pennsylvania, United States were assessed for inclusion[47,69]. We excluded the second twin pregnancy per mother, as well as pregnancies with missing maternal anthropometry, including pre-pregnancy/delivery weight and height, absence of recorded weight measurements during pregnancy, and/or missing gestational age (GA) at birth. We excluded all twin pregnancies with monochorionic placentation, since this condition is associated with a very high-risk profile[70], which may confer less importance to GWG as a monitored characteristic of pregnancy. In the parent singleton study, the subcohort was selected by random sample within pre-pregnancy body mass index (BMI) strata (<18.5 kg/m2 or underweight, 18.5-24.9 kg/m2 or normal weight, 25-29.9 kg/m2 or overweight, 30-34.9 kg/m2or obese class I, 25-29.9 kg/m2 or obese class II, and ≥40.0 kg/m2 or obese class III). Sampling fractions were 0.64 for underweight, 0.035 for normal weight, 0.060 for overweight, 0.14 for obese class I, 0.30 for obese class II, and 0.51 for obesity class II pre-pregnancy BMI. Cases of perinatal conditions of interest (i.e. preterm birth, gestational diabetes) were oversampled in the parent singleton case-cohort[67], though we assessed only pregnancies in the subcohort for inclusion in the current study.

Data Sources

In the parent cohort and case-cohort, serial weight measurements were obtained by medical chart abstraction, while other maternal and pregnancy characteristics were obtained from the hospital-maintained Magee Obstetric, Medical, and Infant Database (MOMI) supplemented with vital statistics data collected by the state of Pennsylvania. Pre-pregnancy and delivery weights were self-reported while prenatal weights were measured; outlying pre-pregnancy weights were identified by data managers and replaced by weight at first visit if within 13 weeks' GA. Estimation of last menstrual period (LMP) in the parent studies generally followed American College of Obstetricians and Gynecologists recommendations[71]; specifically, LMP was estimated using recalled LMP and ultrasound where available.

Measured and Predicted Weights

We leveraged a cohort with many available measurements per pregnancy, retained a subset to estimate maternal weight throughout pregnancy, and compared all available measurements to those estimated in our interpolation models. Specifically, we mimicked a typical data collection schedule by retaining 4 measurements per pregnancy, namely those reported pre-pregnancy and at delivery, as well as those measured at first prenatal and glucose screening visits. We retained only 3 measurements for some pregnancies that had one prenatal visit at glucose screening (i.e. prenatal and glucose screening visits were the same). If glucose screening test date(s) were available, weight recorded at the visit on or most recently prior to the first glucose screening test date was selected; if glucose screening test date(s) and/or results were not available, weight recorded at the visit on or most recently prior to 28 weeks' gestational age was selected. Our choice of prenatal visits was guided by the data collection schedule of Project Viva, a large pregnancy cohort in the United States[72].

We used these retained measurements to interpolate maternal weight trajectory throughout pregnancy. In our study, most women had several additional prenatal visits not retained by our

mimicked data collection schedule. We compared GWG calculated using these measurements to GWG calculated using weights estimated by each interpolation model. We calculated arithmetic and absolute differences between measured and predicted GWG for all available measurements.

Fitting Interpolation Models

We compared methods that varied along three different dimensions. First, we considered models fit at the level of each individual pregnancy versus models that pooled all pregnancies and incorporated individual-level variation using random intercepts and/or slopes. Second, we compared models with increasing degrees of flexibility in how weight changed by GA, including none (i.e. last value carried forward), inclusion of linear and quadratic terms for GA, and incorporation of fractional polynomials[73]/cubic splines[74]. Third, and relatedly, we considered models with varying degrees of parametricity. For example, models incorporating linear and quadratic terms may be considered more parametric than models that linearly connect most proximal weight measurements or incorporate splines. In summary, we evaluated the relative performance of 16 methods (Table 3-1).

For the method incorporating fractional polynomials, we used a standard procedure for selecting the most parsimonious model[73]. Briefly, we fit all first-degree and second-degree fractional polynomials with every unique combination of the powers -3, -2, -1, -0.5, 0, 0.5, 1, 2, 3; per convention, we multiplied the second term by its corresponding ln-transformed covariate for all second-degree fractional polynomial combinations with two powers of the same number. First, we compared the best first-degree fractional polynomial model (i.e. that with the lowest deviance) to the linear model; a Chi-square test with one degree of freedom was performed to determine whether the arithmetic difference in deviances was significantly above zero at the p=0.05-level. Second, we compared the best second-degree fractional polynomial model to best first-degree model; a Chi-square test with two degrees of freedom was performed to determine whether the arithmetic difference in

deviances was significantly above zero at the p=0.05-level. Maternal weights were then predicted using the best fractional polynomial with random intercepts for pregnancy.

Evaluating Interpolation Methods

For each interpolation model, we calculated arithmetic differences in kilograms between measured and predicted GWG for all available maternal weights. For methods that estimate pre-pregnancy weight with imperfect accuracy/precision, we additionally evaluated the accuracy and precision of calculating GWG using predicted instead of measured pre-pregnancy weight. We evaluated methods by comparing summary statistics for differences, including measures of central tendency, such as mean and median, and measures of variation, such as standard error, interquartile range, and range. Our primary statistic of interest was the root mean squared error (RMSE), which was calculated using the following formula, where N is the number of weights: RMSE = SQRT[(Measured Weight – Predicted Weight)^2/ N] . For all summary statistics, values closest to zero indicated better model performance.

We excluded maternal weights retained to fit interpolation models when calculating summary statistics of differences in primary analyses. By definition, differences between measured and predicted weights among retained measurements must be zero for two of the individual-level methods, namely last value carried forward and linear interpolation between most proximal measurement. Conversely, differences between measured and predicted weights, even among measurements used to fit interpolation models, may be non-zero in all other individual/pooled methods. We anticipated that including retained weights for the calculation of difference between measured and predicted value would generate summary statistics artificially closer to zero for the two former methods, while effects on all other methods are unclear. We included retained measurements in difference calculations in sensitivity analyses.

We also examined whether accuracy and precision of models varied by two important characteristics, namely GA of prenatal visit and total number of weight measurements by participant. Generally, we were interested in whether the arithmetic difference (a metric of accuracy) or absolute difference (a metric of precision) varied across these characteristics; for example, whether methods predicted weights with similar accuracy/precision in the first vs. third trimester or among pregnancies with 5 vs. 20 total weight measurements. We fit generalized estimating equations (GEE)[75] using arithmetic or absolute difference in measured vs. predicted weight as the outcome, gestational age of weight measurement (days) or total number of weight measurements by pregnancy as the exposure, and correcting variance for clustering by pregnancy. Regression coefficients represent the population-average change in arithmetic or absolute difference per one-day increase in GA of weight measurement or per one-unit increase in total number of weight measurements by pregnancy. Methods whose regression coefficients were closest to zero demonstrated minimal change in accuracy/precision by the predictor of interest, and therefore indicated better model performance.

Sensitivity Analyses

Primary analyses included all available weight measurements, while sensitivity analyses excluded implausible measurements identified by two available methods for identifying outliers in longitudinal data. We used both conditional percentile[76] and studentized residual[77] methods for identifying outliers in longitudinal data. Briefly, the conditional percentile method calculates expected maternal weight at a given GA conditional on all previous weight measurements; outlying weight measurements are defined as falling above/below the expected value by a given number of standard deviations (we used 4 SD). Conversely, the studentized residual method employs individual-level regression and calculates the absolute value of jackknife residuals for each weight measurement; outliers are defined as being greater than or equal to a given value (we used 10 kg across GA). Both methods were fit with either linear or linear and quadratic terms, which yielded four methods for

identifying outlying maternal weights: conditional percentile (linear), conditional percentile (linear and quadratic), studentized residual (linear), and studentized residual (linear and quadratic). Although these methods for identifying outliers are conceptually similar to our proposed interpolation methods, they use all available weights rather than a subset; thus, we reserved exclusion of outliers for sensitivity analyses. All procedures described in primary analyses were repeated alternatively excluding pregnancies with any outlier maternal weights as identified by each method.

Analyses were conducted in Stata 14.2 statistical software[78]. Ethical approval was obtained from Universities of Pittsburgh and British Columbia for the parent cohorts, and additionally from McGill University Faculty of Medicine Institutional Review Board for the current study.

3.5. Results

Among pregnancies in the parent studies, we retained 2066 (70.4%) twins and 7731 (96.1%) singletons (Figures 3-1 and 3-2); when monochorionic twin pregnancies, second twin pregnancies per mother, and singleton cases are discounted, we retained 89.5% of eligible twin pregnancies.

Mothers of twin pregnancies were older, were more frequently classified as college graduates, non-Hispanic white race, married, privately insured, normal pre-pregnancy BMI, and were less likely to report smoking before/during pregnancy (Table 3-2). As expected, twin pregnancies also had higher total GWG, higher frequency of Caesarean section, and lower GA at birth. While twins generally had lower GA at first ultrasound, both twin and singleton pregnancies had similar numbers of unique weight measurements in this clinical setting. Schematics of visit pattern, including both weights selected to fit and estimated by interpolation models, are displayed in Figures 3-3 and 3-4.

For both twin and singleton pregnancies, the best methods were pooled regression using restricted cubic spline terms with random intercepts and slopes and four knots at trimesters (RMSE=1.55 kg for twins, 1.45 kg for singletons) or quantiles (RMSE=1.56 kg for twins, 1.45 kg for singletons), as well as individual-level linear interpolation using most proximal measurements

(RMSE=1.62 kg for twins, 1.46 kg for singletons). Pooled linear and quadratic regression with random intercepts and slopes also performed relatively well among both pluralities (RMSE=1.70 kg for twins, 1.57 kg for singletons). As expected, individual-level last value carried forward, which the default method for incorporating time-varying covariates in time-to-event analyses, performed worst in both twin and singleton pregnancies (RMSE=6.09 kg for twins, 4.87 kg for singletons). Generally, linear regression with or without ln-transformed weights were less accurate and precise than more flexible methods (RMSE ranged from 2.17 to 3.54 kg in twins and 1.87 to 3.79 kg in singletons). Any model that incorporated only individual-level variation in pre-pregnancy weight (i.e. random intercept) and not in trajectory (i.e. random slope) performed relatively worse (RMSE ranged from 2.81 to 3.54 kg in twins and 3.03 to 3.79 kg in singletons).

When measurements retained in interpolation models were included in summary statistics of differences (Table A3-1 for twins, Table A3-2 for singletons), individual-level linear interpolation using most proximal visits appeared slightly better than methods incorporating both splines and random intercepts. Additionally, individual-level linear and quadratic interpolation appeared somewhat better than their pooled equivalent among both pluralities. All other inferences were similar.

Measures of central tendency largely mirrored mean squared error in terms of comparative performance of methods. Notably, linear interpolation using most proximal measurements overestimated weights on average in both twins (mean=-0.32 kg SE=0.012 kg) and singletons (mean=-0.20 kg, SE=0.006 kg) compared to models incorporating splines and random intercepts and slopes with knots at trimesters (mean=-0.04 kg, SE=0.012 kg in twins and mean=-0.02 kg, SE=0.006 kg in singletons) or quantiles (mean=-0.06 kg, SE=0.012 kg in twins and mean=-0.03 kg, SE=0.006 kg in singletons). This appears to be driven by weight estimation in the first trimester of pregnancy (Figures 3-5B and 3-6B).

We observed that arithmetic differences between measured and predicted weights generally increased as gestational age increased and decreased as total number of weight measurements increased (Tables 3-5 and 3-6). Magnitudes of coefficients for gestational age in days were larger for linear interpolation using most proximal measurements (β=0.010 in twins, β=0.08 in singletons) than in either method incorporating splines and random intercepts and slopes with knots at trimesters (β=0.004, 0.005) or quantiles (β=0.005, 0.006). On average, linear interpolation using most proximal visits overestimated GWG at 20 weeks' GA by 0.73 kg in twins and 0.64 kg in singletons, and underestimated GWG at 40 weeks' GA by 0.71 kg in twins and 0.50 kg in singletons. While most methods followed this pattern, the magnitude of differences across GA was attenuated for methods incorporating both splines and random effects (overestimated by 0.16 kg in twins and 0.31 kg in singletons at 20 weeks GA; underestimated by 0.36 kg in twins and 0.44 kg in singletons at 40 weeks GA). Median, IQR, and range of differences by GA for each method and plurality are plotted in Figures 3A-P (twins) and 4A-P (singletons). Coefficients for total number of unique weight measurements were similar across method in both singleton and twin pregnancies.

Similar to arithmetic difference, change in absolute difference between measured and predicted GWG across GA was low for methods that incorporated splines (β=-0.002 in both twins and singletons; Tables 3-7 and 3-8). Pooled linear regression with random intercepts and slopes demonstrated the least change in precision by GA among twin pregnancies; however, average absolute difference ranged from 2.43 kg at 20 weeks' GA to 2.45 kg at 40 weeks' GA, indicating that overall precision for this method was low, even if average rate of change by GA was near zero. Linear interpolation using most proximal measurements was more precise at later gestational ages; specifically, absolute difference ranged from 1.35 to 0.72 kg in twins and 1.26 to 0.77 kg in singletons. Pooled methods that incorporated only random intercepts were less precise as gestational age increased.

Coefficients for total number of unique weight measurements were closest to zero for restricted cubic splines, regardless of knot placement.

3.6. Discussion

Among both twin and singleton pregnancies in our study, the best methods for interpolating gestational weight gain between measurements were pooled regression using restricted cubic spline terms with random intercepts and slopes and knots at trimesters or quantiles, as well as individual-level linear interpolation using most proximal measurements. In relation to other methods, these ranked high on the dimension of individual variation and flexibility and low on the dimension of parametricity. Specifically, the best three methods incorporated maximum individual variation, either by including both random intercepts and slopes for pregnancies, or by fitting models for each individual pregnancy. Similarly, these methods were highly flexible in relation to the data used to fit the models. Relatedly, they best methods were less parametric than traditional regression models that used linear or quadratic terms, although quadratic models that incorporated maximum individual-level variation performed only slightly worse.

Our findings are in contrast with traditional methods for quantifying gestational weight gain. Previous research has discussed bias introduced by the inappropriate use of total GWG as an exposure measurement when investigating outcomes that occur either during pregnancy or are intertwined with GA at birth[3]. Our study additionally notes that linear rate of change, a cumulative measure frequently employed in studies of GWG, is less accurate/precise than at least 5 alternative methods in both singleton and twin pregnancies. Our study also highlights the importance of incorporating variation in both pre-pregnancy weight and weight gain trajectory in all interpolation models. Although trajectories may appear similar by visual inspection, models that omit random slopes performed several magnitudes worse than equivalents that include these terms. Our estimates of mean arithmetic difference and

RMSE were smaller in magnitude than in a prior study, which investigated interpolation methods at 28 and 40 weeks' GA among singleton pregnancies[68].

We suggest that more complex exposure conceptualizations may limit measurement error and/or bias in studies of GWG. For example, prior research has found that investigating total cumulative rate of GWG instead of total GWG z-score may overestimate the relationship between low GWG and preterm birth[1]. Methods evaluated in our study have potential applications to research investigating relationships between time-varying GWG and perinatal outcomes, particularly within a time-to-event framework. Specifically, the relationship between cumulative total or rate of GWG and preterm birth may alternatively be investigated as that between time-varying GWG and time-to-delivery[27]. Additionally, GWG at a specific GA may be of interest when a perinatal outcome occurs or is measured during pregnancy; examples of such outcomes previously investigated in relation to GWG include infant/fetal growth, gestational diabetes, or preeclampsia[67,69,79,80]. Methods beyond total/cumulative rate of GWG may be helpful in instances where pattern or GWG at a specific GA is of interest, and/or measured GWG data are not uniformly collected at the same GAs for all participants.

While linear interpolation using most proximal measurements and random effects models incorporating restricted cubic splines performed similarly, these methods have different benefits and challenges. Individual linear interpolation is conceptually straightforward; it represents "connecting the dots" between available weight measurements for each pregnancy. It is also relatively simple to implement in statistical software, albeit computationally time-consuming. Conversely, restricted cubic splines and random effects are statistical methods that may be difficult to succinctly communicate in the context of clinical journals; incorporating these when interpolating GWG may unnecessarily inhibit uptake of research findings. However, these methods performed marginally better than individual-level linear interpolation between proximal weights. Additionally, pooled methods benefit from borrowing information from other pregnancies, which may be useful in situations with sparse data. Similar to

individual-level interpolation, random effects models with splines are computationally taxing; researchers may also experience convergence issues. Moreover, when examining the effect of GWG patterns/trajectories on perinatal outcomes, it may be difficult to justify pooling all pregnancies for the purpose of estimating exposure. While incorporating individual-level variation through both random intercepts and slopes should ameliorate this concern, researchers may be reluctant to estimate exposure data by pooling participants both with and without the outcome of interest. For this reason, we recommend interpolating GWG using random effects models with splines in primary and individual-level linear interpolation using most proximal measurements in sensitivity analyses.

Root mean squared error for the best interpolation methods ranged from approximately 1.45 to 1.75 kg. Each researcher must assess whether this error is practically acceptable in the context of their GWG study. Measurement error in exposure typically biases effect estimates towards the null when non-differential by outcome. Thus, acceptability of these interpolation methods may depend on magnitude of expected effect. Additionally, practicality of expanded data collection schedules and/or availability of alternative exposure conceptualizations may also be important to consider. For example, if RMSE estimated for these methods is deemed acceptable, this may justify more cost-efficient data collection routines in resource-poor settings. Moreover, while GWG researchers have emphasized importance of collecting serial weight measurements, it is notable that self-reported pre-pregnancy weight may experience average underestimation as high in magnitude as 3 kg[81]. Thus, the accuracy and precision of our recommended methods, even when only 3-4 weight measurements per participant are used to fit models, may instead justify more emphasis on increasing accuracy and precision of pre-pregnancy weight in the conduct of GWG research.

Our study had many strengths. We leveraged previously-collected clinical data for two large pregnancy cohorts, which captured either a census or random sample of pregnancies delivered at a large obstetric hospital in Pittsburgh, Pennsylvania. We tested a comprehensive collection of

interpolation methods, incorporating varying degrees of individual-level variation, flexibility, and parametricity across gestational duration in both twin and singleton pregnancies. GWG is infrequently studied among twins, although these pregnancies are at higher risk for perinatal outcomes and experience greater GWG on average.

We also note key weaknesses of our study. Self-reported pre-pregnancy weight has been associated with measurement bias up to 3 kg in previous research[81]. However, implausible pre-pregnancy weights were replaced with weight at first visit if it occurred before 13 weeks' GA[67]. Although we attempted to replicate the data collection schedule of Project Viva, not all pregnancies had recorded glucose screening tests. In instances where glucose screening test was not recorded and/or timing was not available, we estimated this visit that which was on or closest to 28 weeks' GA. Distribution of GA among our selected visits may differ slightly from that of Project Viva, which reports first prenatal visit at median 10 weeks' GA and glucose screening visit at median 28 weeks' GA among singletons[72]. Gestational age may be estimated with different accuracy/precision in twin and singleton pregnancies, which would limit comparability of our findings; however, gestational age estimated by either first or second trimester ultrasound was found to perform similarly in both twins and singletons conceived by in vitro fertilization[82].

Importantly, we cannot guarantee that RMSE estimated for linear interpolation using most proximal measurement is generalizable to pregnancies with four or fewer unique weight measurements. All unique weight measurements among pregnancies with only two prenatal visits both on or before 28 weeks' GA, or only one prenatal visit, would have been selected for inclusion in interpolation models. Linear interpolation using most proximal measurements and last value carried forward methods would have predicted these weights with zero error, since this method estimates only weights between selected prenatal visits. Conversely, all other individual-level/pooled methods do not necessarily estimate selected prenatal weight measurements with exact accuracy and precision. Thus predictive

value of this method may be evaluated even among pregnancies with two or fewer prenatal visits, even though all of these measurements would have been included in interpolation models. It is notable that magnitudes for coefficients in regression models that investigated predictors of differences were small across methods; nonetheless, we cannot rule out the possibility that linear interpolation using most proximal measurements is insufficient among pregnancies with two or fewer weight measurements during pregnancy.

In conclusion, we recommend interpolation methods with high flexibility, high individual-level variation, and low parametricity for the estimation of GWG between weight measurements. In particular, the results of our study suggest that methods incorporating restricted cubic splines, random intercepts and slopes for pregnancy, and knots at either trimesters or quantiles are best for interpolating GWG. Individual-level linear interpolation using most proximal measurements is an acceptable alternative, and individual level linear and quadratic interpolation, or pooled equivalent with random intercepts and slopes, performs only marginally worse than more flexible/less parametric alternatives. Our study has implications for investigating GWG trajectory over the course of pregnancy, and specifically among twins. Additionally, our methods and approach may have applications elsewhere in biomedical research in instances where metrics are measured at discrete points in time and change somewhat but not entirely predictably over time.

3.7. Tables

Table 3-1. Interpolation Methods Evaluated

Method	Level of Data Aggregation
Last Value Carried Forward	Individual
Linear Interpolation, Proximal	Individual
Linear Interpolation, All	Individual
Linear Interpolation, Ln-Transformed	Individual
Quadratic Interpolation, All	Individual
Linear Mixed Effects, Random Intercept	Pooled
Linear Mixed Effects, Random Intercept and Slope	Pooled
Quadratic Mixed Effects, Random Intercept	Pooled
Quadratic Mixed Effects, Random Intercept and Slope	Pooled
Ln-Transformed Mixed Effects, Random Intercept	Pooled
Ln-Transformed Mixed Effects, Random Intercept and Slope	Pooled
Fractional Polynomial Mixed Effects, Random Intercept	Pooled
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept	Pooled
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept and Slope	Pooled
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept	Pooled
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept and Slope	Pooled

Table 3-2. Sample Characteristics by Plurality**.** Continuous variables presented as median (25th percentile; 75th percentile), categorical variables presented as number (percentage), and covariates at the level of the mother/pregnancy unless otherwise stated.

Characteristic	Twins (n=2066)	Singletons (n=7731)
Maternal Age (years)[a]	30.5 (5.9)	28.9 (6.0)
Maternal Education		
<High School	113 (5.5)	658 (8.5)
High School/GED	414 (20.0)	1954 (25.3)
Some College/Associates	444 (21.5)	2071 (26.8)
College Graduate	1075 (52.0)	3038 (39.3)
missing	20 (1.0)	10 (0.1)
Race		
Non-Hispanic White	1632 (79.0)	5732 (74.1)
Non-Hispanic Black	356 (17.2)	1651 (21.4)
Hispanic	17 (0.8)	69 (0.9)
Other	61 (3.0)	279 (3.6)
Married		
Yes	1445 (69.9)	4619 (59.7)

No	619 (30.0)	3112 (40.3)
missing	2 (0.1)	0 (0.0)
Insurance		
Private/Other	1311 (63.5)	4340 (56.1)
Medicaid/Self-Pay	755 (36.5)	3390 (43.8)
missing	0 (0.0)	1 (0.0)
Parity		
Nulliparous	962 (46.6)	3365 (43.5)
Primiparous	638 (30.9)	2601 (33.6)
Multiparous	466 (22.6)	1765 (22.8)
Pre-pregnancy BMI Category		
Underweight	65 (3.1)	1222 (15.8)
Normal Weight	981 (47.5)	1353 (17.5)
Overweight	528 (25.6)	1324 (17.1)
Obese	492 (23.8)	3832 (49.6)
Pre-existing Diabetes		
Yes	59 (2.9)	209 (2.7)
No	2001 (96.9)	7522 (97.3)
missing	6 (0.3)	0 (0.0)
Pre-existing Hypertension		
Yes	115 (5.6)	533 (6.9)
No	1945 (94.1)	7198 (93.1)
missing	6 (0.3)	0 (0.0)
Ever Smoker		
Yes	290 (14.0)	1428 (18.5)
No	1776 (86.0)	6303 (81.5)
Assisted Reproductive Technology		
Yes	617 (29.9)	Not Available
No	1439 (69.7)	Not Available
missing	10 (0.5)	Not Available
Number of Unique Weight Measurements	12.0 (10.0; 14.0)	12.0 (10.0; 14.0)
Gestational Age at Birth (weeks)	36.3 (33.9; 37.9)	39.3 (38.3; 40.3)
Ultrasound		
Yes	2065 (100.0)	7634 (98.7)
No	1 (0.0)	97 (1.3)
Gestational Age at First Ultrasound (weeks)	12.0 (8.0; 18.3)	15.6 (9.6; 19.1)
missing	100 (4.8)	595 (7.7)
Type of Delivery		
Spontaneous	1178 (57.0)	3978 (51.5)
Induced	341 (16.5)	2704 (35.0)
No Labor	547 (26.5)	1034 (13.4)
missing	0 (0.0)	15 (0.2)
Delivery Route (Twin A/Singleton)		

Vaginal	936 (45.3)	5464 (70.7)
Caesarean Section	1130 (54.7)	2260 (29.2)
missing	0 (0.0)	7 (0.1)
Delivery Route (Twin B)		
Vaginal	847 (41.0)	N/A
Caesarean Section	1219 (59.0)	N/A
Stillbirth (Twin A/Singleton)		
Yes	9 (0.4)	33 (0.4)
No	2057 (99.6)	7698 (99.6)
Stillbirth (Twin B)		
Yes	17 (0.8)	N/A
No	2049 (99.2)	N/A
Total Gestational Weight Gain (kg)	16.8 (12.0; 21.8)	13.2 (8.6; 17.2)

[a]*Mean(SE)*

Table 3-3. Summary statistics for differences between measured and predicted values of maternal weight in twin pregnancies (n=16338 weight measurements). PPW refers to pre-pregnancy weight.

Method	Mean (SE)	Median (IQR)	Range	RMSE (Measured PPW)	RMSE (Predicted PPW)
Last Value Carried Forward	4.59 (0.031)	4.08 (1.81; 6.80)	-22.68; 31.75	6.09	6.09
Linear Interpolation, Proximal	-0.32 (0.012)	-0.13 (-1.16; 0.63)	-24.10; 13.41	1.62	1.62
Linear Interpolation, All	-0.57 (0.017)	-0.31 (-1.93; 0.94)	-26.41; 15.77	2.30	3.36
Linear Interpolation, Ln-Transformed	-0.37 (0.017)	-0.15 (-1.65; 1.05)	-24.38; 15.96	2.17	2.96
Quadratic Interpolation, All	0.09 (0.014)	0.20 (-0.78; 1.09)	-25.95; 12.58	1.75	1.78
Linear Mixed Effects, Random Intercept	-0.46 (0.025)	-0.66 (-2.52; 1.39)	-18.87; 23.61	3.20	5.81
Linear Mixed Effects, Random Intercept and Slope	-0.54 (0.018)	-0.39 (-1.96; 0.96)	-23.98; 16.09	2.37	3.41
Quadratic Mixed Effects, Random Intercept	0.15 (0.022)	0.12 (-1.58; 1.83)	-18.60; 22.84	2.87	5.40
Quadratic Mixed Effects, Random Intercept and Slope	0.11 (0.013)	0.18 (-0.84; 1.14)	-20.60; 14.91	1.70	2.03
Ln-Transformed Mixed Effects, Random Intercept	-0.39 (0.028)	-0.45 (-2.58; 1.77)	-22.74; 21.19	3.54	6.82
Ln-Transformed Mixed Effects, Random Intercept and Slope	-0.30 (0.018)	-0.19 (-1.68; 1.12)	-21.69; 16.32	2.26	3.14
Fractional Polynomial Mixed Effects, Random Intercept	-0.01 (0.022)	-0.03 (-1.60; 1.61)	-18.76; 22.45	2.81	5.37
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept	0.02 (0.022)	-0.04 (-1.61; 1.64)	-18.67; 22.73	2.82	5.38
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept and Slope	-0.06 (0.012)	0.01 (-0.93; 0.87)	-21.90; 14.06	1.56	1.79
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept	0.05 (0.022)	0.00 (-1.58; 1.66)	-18.70; 22.71	2.82	5.37
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept and Slope	-0.04 (0.012)	0.03 (-0.89; 0.88)	-22.10; 14.07	1.55	1.77

Table 3-4. Summary statistics for differences between measured and predicted values of maternal weight in singleton pregnancies (n=65286 weight measurements). PPW refers to pre-pregnancy weight.

Method	Mean (SE)	Median (IQR)	Range	RMSE (Measured PPW)	RMSE (Predicted PPW)
Last Value Carried Forward	3.32 (0.014)	2.72 (0.91; 5.44)	-17.69; 36.29	4.87	4.87
Linear Interpolation, Proximal	-0.20 (0.006)	-0.03 (-0.99; 0.65)	-12.39; 9.38	1.46	1.46
Linear Interpolation, All	-0.30 (0.007)	-0.11 (-1.40; 0.94)	-14.01; 10.04	1.91	2.63
Linear Interpolation, Ln-Transformed	-0.19 (0.007)	-0.03 (-1.27; 1.00)	-13.75; 10.14	1.87	2.47
Quadratic Interpolation, All	0.01 (0.006)	0.11 (-0.83; 0.92)	-24.88; 10.87	1.59	1.63
Linear Mixed Effects, Random Intercept	-0.26 (0.012)	-0.40 (-2.22; 1.55)	-19.61; 21.40	3.18	6.11
Linear Mixed Effects, Random Intercept and Slope	-0.28 (0.008)	-0.15 (-1.44; 0.97)	-13.93; 10.13	1.96	2.74
Quadratic Mixed Effects, Random Intercept	0.08 (0.012)	0.02 (-1.71; 1.81)	-19.91; 20.84	3.05	5.98
Quadratic Mixed Effects, Random Intercept and Slope	0.05 (0.006)	0.11 (-0.85; 1.01)	-10.91; 9.62	1.57	1.84
Ln-Transformed Mixed Effects, Random Intercept	-0.57 (0.015)	-0.55 (-2.76; 1.79)	-22.80; 19.94	3.79	7.66
Ln-Transformed Mixed Effects, Random Intercept and Slope	-0.15 (0.008)	-0.06 (-1.31; 1.08)	-13.62; 11.23	1.95	2.68
Fractional Polynomial Mixed Effects, Random Intercept	0.00 (0.012)	-0.05 (-1.76; 1.71)	-19.96; 20.82	3.04	5.98
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept	0.01 (0.012)	-0.03 (-1.74; 1.72)	-19.94; 20.80	3.04	5.97
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept and Slope	-0.03 (0.006)	0.03 (-0.84; 0.83)	-11.99; 9.24	1.45	1.60
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept	0.01 (0.012)	-0.02 (-1.72; 1.72)	-19.93; 20.76	3.03	5.97
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept and Slope	-0.02 (0.006)	0.03 (-0.83; 0.84)	-12.22; 9.26	1.45	1.62

Table 3-5. Relationship between arithmetic difference in measured and predicted weights and gestational age of measurement, total number of unique measurements among twins (N pregnancies = 2002, N weights = 16338).

Method	Coefficient for GA	Estimate 20 weeks	Estimate 32 weeks	Estimate 40 weeks	Coefficient for Visits
Last Value Carried Forward	0.027 (0.025; 0.028)	3.35 (3.25; 3.46)	5.59 (5.44; 5.74)	7.08 (6.86; 7.31)	0.139 (0.104; 0.175)
Linear Interpolation, Proximal	0.010 (0.010; 0.011)	-0.73 (-0.78; -0.67)	0.14 (0.09; 0.18)	0.71 (0.64; 0.79)	-0.018 (-0.031; -0.004)
Linear Interpolation, All	0.024 (0.023; 0.025)	-1.53 (-1.60; -1.45)	0.47 (0.42; 0.53)	1.81 (1.71; 1.90)	-0.025 (-0.042; -0.008)
Linear Interpolation, Ln-Transformed	0.021 (0.020; 0.022)	-1.23 (-1.30; -1.16)	0.56 (0.50; 0.61)	1.75 (1.65; 1.84)	-0.018 (-0.035; -0.002)
Quadratic Interpolation, All	0.009 (0.008; 0.010)	-0.26 (-0.33; -0.20)	0.49 (0.43; 0.54)	0.99 (0.91; 1.07)	-0.012 (-0.028; 0.004)
Linear Mixed Effects, Random Intercept	0.026 (0.025; 0.028)	-1.57 (-1.66; -1.49)	0.65 (0.51; 0.79)	2.13 (1.89; 2.37)	0.035 (0.008; 0.062)
Linear Mixed Effects, Random Intercept and Slope	0.024 (0.023; 0.026)	-1.54 (-1.62; -1.47)	0.51 (0.45; 0.58)	1.89 (1.78; 1.99)	-0.009 (-0.026; 0.008)
Quadratic Mixed Effects, Random Intercept	0.011 (0.009; 0.013)	-0.32 (-0.40; -0.24)	0.61 (0.47; 0.75)	1.23 (1.00; 1.47)	0.014 (-0.011; 0.040)
Quadratic Mixed Effects, Random Intercept and Slope	0.009 (0.008; 0.010)	-0.26 (-0.32; -0.20)	0.52 (0.46; 0.57)	1.04 (0.95; 1.12)	-0.011 (-0.025; 0.004)
Ln-Transformed Mixed Effects, Random Intercept	0.021 (0.018; 0.023)	-1.27 (-1.36; -1.17)	0.47 (0.29; 0.65)	1.62 (1.33; 1.92)	0.033 (0.001; 0.066)
Ln-Transformed Mixed Effects, Random Intercept and Slope	0.022 (0.021; 0.023)	-1.21 (-1.29; -1.14)	0.65 (0.58; 0.72)	1.89 (1.78; 2.00)	0.000 (-0.018; 0.018)
Fractional Polynomial Mixed Effects, Random Intercept	0.003 (0.001; 0.005)	-0.10 (-0.18; -0.01)	0.13 (-0.01; 0.27)	0.28 (0.04; 0.52)	-0.019 (-0.045; 0.007)
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept	0.007 (0.005; 0.009)	-0.24 (-0.33; -0.16)	0.32 (0.18; 0.46)	0.70 (0.46; 0.94)	-0.004 (-0.030; 0.022)

Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept and Slope	0.005 (0.004; 0.006)	-0.25 (-0.31; -0.20)	0.18 (0.14; 0.23)	0.47 (0.40; 0.55)	-0.026 (-0.040; -0.012)
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept	0.005 (0.004; 0.007)	-0.16 (-0.25; -0.08)	0.29 (0.15; 0.44)	0.60 (0.36; 0.84)	-0.006 (-0.032; 0.020)
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept and Slope	0.004 (0.003; 0.004)	-0.16 (-0.22; -0.11)	0.15 (0.10; 0.20)	0.36 (0.28; 0.43)	-0.028 (-0.042; -0.014)

Table 3-6. Relationship between arithmetic difference in measured and predicted weights and gestational age of measurement, total number of unique measurements among singletons (N pregnancies = 7652, N weights = 65286).

Method	**Coefficient for GA**	**Estimate 20 weeks**	**Estimate 32 weeks**	**Estimate 40 weeks**	**Coefficient for Visits**
Last Value Carried Forward	0.023 (0.022; 0.024)	1.89 (1.83; 1.94)	3.83 (3.77; 3.89)	5.13 (5.04; 5.22)	0.094 (0.076; 0.113)
Linear Interpolation, Proximal	0.008 (0.008; 0.008)	-0.64 (-0.66; -0.61)	0.04 (0.03; 0.06)	0.50 (0.47; 0.53)	-0.029 (-0.036; -0.023)
Linear Interpolation, All	0.015 (0.014; 0.015)	-1.11 (-1.15; -1.08)	0.12 (0.09; 0.14)	0.93 (0.90; 0.97)	-0.034 (-0.042; -0.026)
Linear Interpolation, Ln-Transformed	0.013 (0.013; 0.014)	-0.93 (-0.96; -0.89)	0.19 (0.16; 0.21)	0.93 (0.89; 0.96)	-0.034 (-0.042; -0.026)
Quadratic Interpolation, All	0.008 (0.007; 0.008)	-0.40 (-0.43; -0.37)	0.24 (0.22; 0.26)	0.66 (0.63; 0.69)	-0.025 (-0.033; -0.017)
Linear Mixed Effects, Random Intercept	0.015 (0.015; 0.016)	-1.13 (-1.17; -1.08)	0.17 (0.10; 0.23)	1.03 (0.92; 1.14)	-0.012 (-0.028; 0.005)
Linear Mixed Effects, Random Intercept and Slope	0.015 (0.014; 0.015)	-1.11 (-1.15; -1.08)	0.13 (0.11; 0.15)	0.96 (0.92; 1.00)	-0.027 (-0.036; -0.019)
Quadratic Mixed Effects, Random Intercept	0.008 (0.007; 0.009)	-0.37 (-0.42; -0.33)	0.32 (0.26; 0.39)	0.79 (0.68; 0.89)	-0.019 (-0.036; -0.003)
Quadratic Mixed Effects, Random Intercept and Slope	0.008 (0.007; 0.008)	-0.37 (-0.40; -0.34)	0.27 (0.25; 0.29)	0.70 (0.67; 0.73)	-0.028 (-0.035; -0.021)

Ln-Transformed Mixed Effects, Random Intercept	0.008 (0.006; 0.009)	-0.93 (-0.98; -0.89)	-0.30 (-0.38; -0.22)	0.12 (-0.01; 0.26)	-0.059 (-0.079; -0.038)
Ln-Transformed Mixed Effects, Random Intercept and Slope	0.014 (0.013; 0.014)	-0.91 (-0.94; -0.87)	0.23 (0.21; 0.26)	0.99 (0.95; 1.03)	-0.026 (-0.034; -0.017)
Fractional Polynomial Mixed Effects, Random Intercept	0.007 (0.006; 0.008)	-0.36 (-0.40; -0.31)	0.20 (0.14; 0.27)	0.57 (0.47; 0.68)	-0.023 (-0.040; -0.007)
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept	0.006 (0.005; 0.007)	-0.33 (-0.37; -0.28)	0.20 (0.14; 0.26)	0.55 (0.44; 0.66)	-0.022 (-0.039; -0.006)
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept and Slope	0.006 (0.006; 0.006)	-0.35 (-0.38; -0.32)	0.15 (0.13; 0.17)	0.48 (0.45; 0.51)	-0.025 (-0.032; -0.019)
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept	0.006 (0.005; 0.007)	-0.29 (-0.33; -0.25)	0.19 (0.13; 0.25)	0.51 (0.40; 0.62)	-0.022 (-0.038; -0.006)
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept and Slope	0.005 (0.005; 0.006)	-0.31 (-0.34; -0.28)	0.14 (0.12; 0.16)	0.44 (0.41; 0.47)	-0.026 (-0.032; -0.019)

Table 3-7. Relationship between absolute difference in measured and predicted weights and gestational age of measurement, total number of unique measurements among twins (N pregnancies = 2002, N weights = 16338).

Method	**Coefficient for GA**	**Estimate 20 weeks**	**Estimate 32 weeks**	**Estimate 40 weeks**	**Coefficient for Visits**
Last Value Carried Forward	0.023 (0.022; 0.025)	3.72 (3.64; 3.81)	5.69 (5.55; 5.83)	7.01 (6.79; 7.22)	0.125 (0.094; 0.157)
Linear Interpolation, Proximal	-0.005 (-0.005; -0.004)	1.35 (1.31; 1.39)	0.97 (0.94; 1.00)	0.72 (0.67; 0.77)	0.003 (-0.006; 0.012)
Linear Interpolation, All	-0.010 (-0.011; -0.009)	2.14 (2.09; 2.20)	1.30 (1.26; 1.34)	0.74 (0.68; 0.80)	0.007 (-0.006; 0.021)
Linear Interpolation, Ln-Transformed	-0.008 (-0.008; -0.007)	1.96 (1.91; 2.02)	1.30 (1.26; 1.34)	0.86 (0.80; 0.92)	0.002 (-0.011; 0.015)

Quadratic Interpolation, All	-0.003 (-0.004; -0.003)	1.38 (1.34; 1.43)	1.11 (1.07; 1.14)	0.92 (0.87; 0.98)	0.001 (-0.010; 0.012)
Linear Mixed Effects, Random Intercept	0.000 (-0.001; 0.001)	2.43 (2.37; 2.49)	2.44 (2.36; 2.52)	2.45 (2.32; 2.57)	0.037 (0.018; 0.057)
Linear Mixed Effects, Random Intercept and Slope	-0.009 (-0.010; -0.008)	2.17 (2.11; 2.22)	1.41 (1.37; 1.45)	0.90 (0.84; 0.97)	0.010 (-0.004; 0.023)
Quadratic Mixed Effects, Random Intercept	0.006 (0.005; 0.007)	1.89 (1.84; 1.94)	2.38 (2.30; 2.46)	2.70 (2.58; 2.82)	0.027 (0.008; 0.046)
Quadratic Mixed Effects, Random Intercept and Slope	-0.003 (-0.003; -0.002)	1.40 (1.36; 1.44)	1.16 (1.12; 1.19)	0.99 (0.94; 1.04)	-0.005 (-0.015; 0.006)
Ln-Transformed Mixed Effects, Random Intercept	0.007 (0.006; 0.008)	2.37 (2.32; 2.43)	2.95 (2.85; 3.05)	3.33 (3.18; 3.49)	0.053 (0.031; 0.076)
Ln-Transformed Mixed Effects, Random Intercept and Slope	-0.006 (-0.007; -0.006)	1.99 (1.94; 2.04)	1.45 (1.41; 1.50)	1.10 (1.03; 1.16)	0.005 (-0.008; 0.018)
Fractional Polynomial Mixed Effects, Random Intercept	0.007 (0.006; 0.008)	1.79 (1.74; 1.84)	2.36 (2.28; 2.44)	2.75 (2.63; 2.87)	0.022 (0.003; 0.040)
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept	0.007 (0.006; 0.007)	1.80 (1.75; 1.85)	2.36 (2.28; 2.44)	2.73 (2.62; 2.85)	0.025 (0.006; 0.043)
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept and Slope	-0.002 (-0.002; -0.001)	1.22 (1.18; 1.26)	1.07 (1.04; 1.11)	0.97 (0.92; 1.02)	-0.004 (-0.014; 0.006)
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept	0.007 (0.006; 0.008)	1.79 (1.74; 1.84)	2.36 (2.28; 2.44)	2.75 (2.63; 2.87)	0.024 (0.005; 0.042)
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept and Slope	-0.002 (-0.002; -0.001)	1.20 (1.16; 1.23)	1.07 (1.04; 1.10)	0.98 (0.93; 1.03)	-0.004 (-0.014; 0.006)

Table 3-8. Relationship between absolute difference in measured and predicted weights and gestational age of measurement, total number of unique measurements among singletons (N pregnancies = 7652, N weights = 65286).

Method	Coefficient for GA	Estimate 20 weeks	Estimate 32 weeks	Estimate 40 weeks	Coefficient for Visits
Last Value Carried Forward	0.018 (0.018; 0.019)	2.60 (2.56; 2.64)	4.13 (4.08; 4.19)	5.16 (5.07; 5.24)	0.100 (0.085; 0.116)
Linear Interpolation, Proximal	-0.004 (-0.004; -0.003)	1.26 (1.24; 1.28)	0.97 (0.95; 0.98)	0.77 (0.75; 0.79)	0.009 (0.004; 0.013)
Linear Interpolation, All	-0.006 (-0.006; -0.006)	1.79 (1.77; 1.82)	1.28 (1.27; 1.30)	0.94 (0.92; 0.96)	0.007 (0.002; 0.013)
Linear Interpolation, Ln-Transformed	-0.005 (-0.005; -0.005)	1.71 (1.69; 1.74)	1.27 (1.26; 1.29)	0.98 (0.96; 1.00)	0.005 (0.000; 0.011)
Quadratic Interpolation, All	-0.003 (-0.003; -0.003)	1.32 (1.29; 1.34)	1.08 (1.06; 1.09)	0.92 (0.90; 0.94)	0.004 (-0.001; 0.009)
Linear Mixed Effects, Random Intercept	0.005 (0.005; 0.005)	2.09 (2.06; 2.11)	2.51 (2.47; 2.55)	2.79 (2.74; 2.85)	0.048 (0.037; 0.058)
Linear Mixed Effects, Random Intercept and Slope	-0.006 (-0.006; -0.005)	1.81 (1.78; 1.83)	1.35 (1.33; 1.36)	1.04 (1.02; 1.06)	0.009 (0.003; 0.015)
Quadratic Mixed Effects, Random Intercept	0.007 (0.007; 0.008)	1.82 (1.80; 1.85)	2.45 (2.41; 2.49)	2.86 (2.80; 2.92)	0.042 (0.031; 0.053)
Quadratic Mixed Effects, Random Intercept and Slope	-0.002 (-0.003; -0.002)	1.33 (1.31; 1.34)	1.12 (1.11; 1.13)	0.98 (0.96; 1.00)	0.003 (-0.002; 0.007)
Ln-Transformed Mixed Effects, Random Intercept	0.012 (0.011; 0.012)	2.16 (2.13; 2.19)	3.13 (3.09; 3.18)	3.79 (3.71; 3.86)	0.072 (0.059; 0.085)
Ln-Transformed Mixed Effects, Random Intercept and Slope	-0.004 (-0.004; -0.004)	1.73 (1.70; 1.75)	1.37 (1.35; 1.39)	1.13 (1.11; 1.16)	0.008 (0.002; 0.013)
Fractional Polynomial Mixed Effects, Random Intercept	0.008 (0.007; 0.008)	1.79 (1.77; 1.82)	2.43 (2.39; 2.47)	2.85 (2.79; 2.91)	0.043 (0.033; 0.054)
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept	0.008 (0.007; 0.008)	1.78 (1.76; 1.81)	2.43 (2.39; 2.47)	2.86 (2.80; 2.91)	0.043 (0.033; 0.054)
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept and Slope	-0.002 (-0.002; -0.002)	1.19 (1.17; 1.21)	1.02 (1.01; 1.03)	0.91 (0.89; 0.93)	0.006 (0.002; 0.010)
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept	0.008 (0.007; 0.008)	1.77 (1.75; 1.80)	2.42 (2.38; 2.46)	2.86 (2.80; 2.92)	0.044 (0.033; 0.054)

Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept and Slope	-0.002 (-0.002; -0.002)	1.18 (1.16; 1.20)	1.02 (1.01; 1.04)	0.92 (0.90; 0.94)	0.006 (0.002; 0.010)

3.8. Figures

Figure 3-1. Twin pregnancy sample selection.

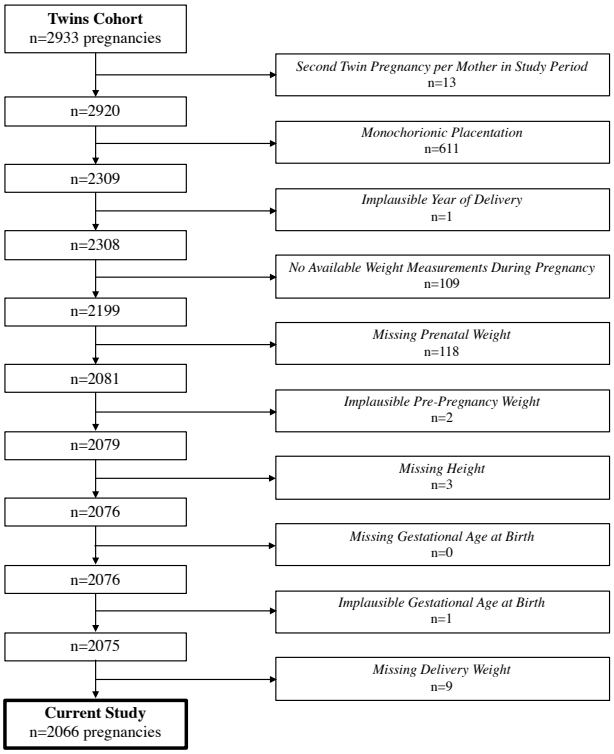


Figure 3-2. Singleton pregnancy sample selection.

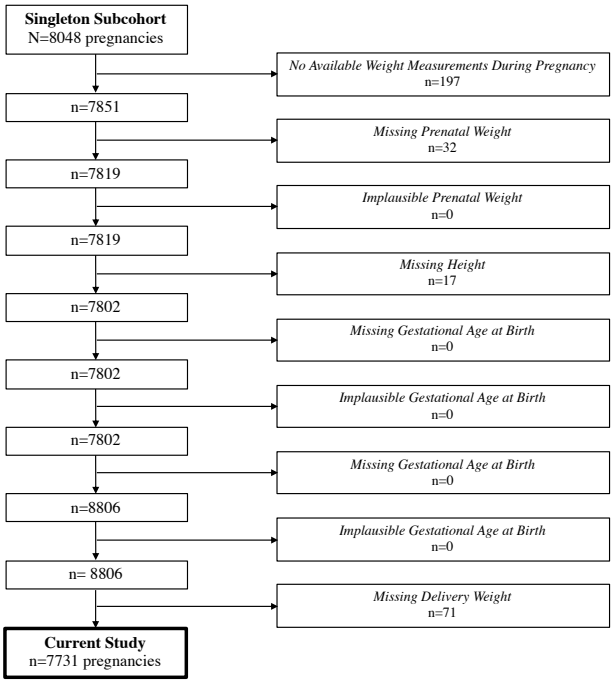


Figure 3-3. Pattern of weight measurements by gestational age in twin pregnancies; weights in red were used to fit interpolation models while weights in black were used to evaluate interpolation models.

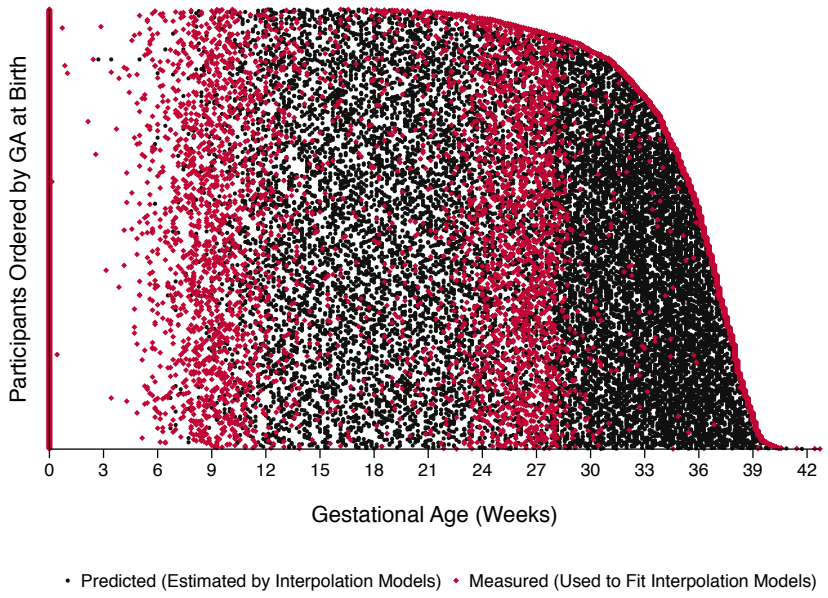


Figure 3-4. Pattern of weight measurements by gestational age for a random selection of singleton pregnancies (N=2066); weights in red were used to fit interpolation models while weights in black were used to evaluate interpolation models.

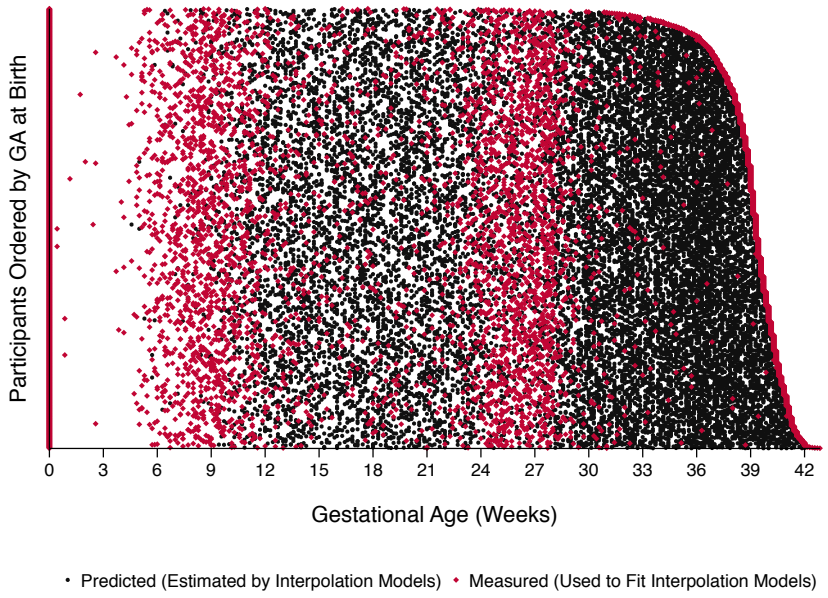


Figure 3-5. Arithmetic difference in measured and predicted maternal weight by gestational age of measurement and interpolation method for twin pregnancies.

Figure 3-5A. Individual-level last value carried forward.

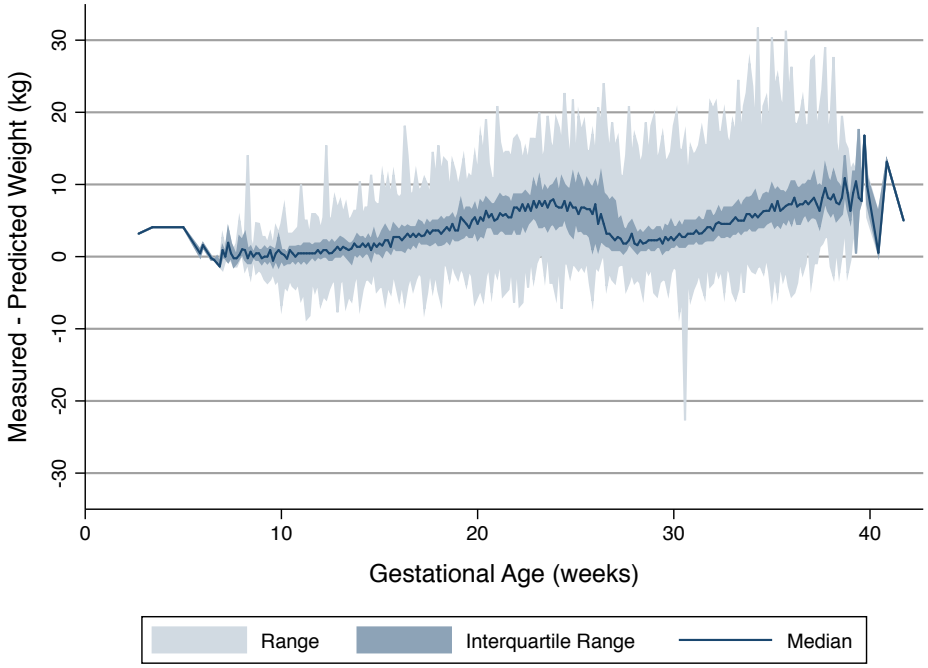


Figure 3-5B. Individual-level linear interpolation between most proximal measurements.

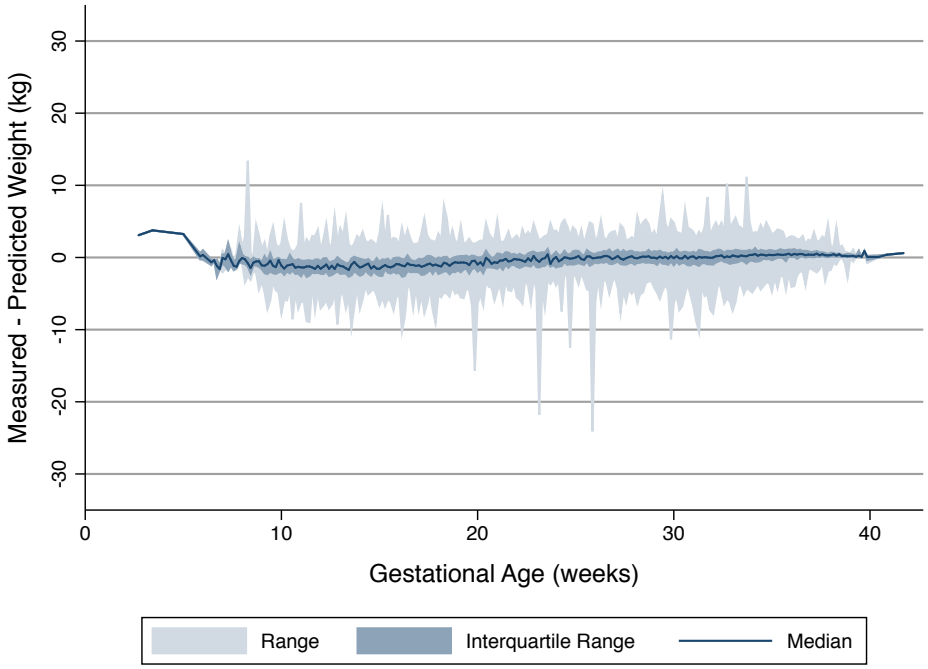


Figure 3-5C. Individual-level linear regression.

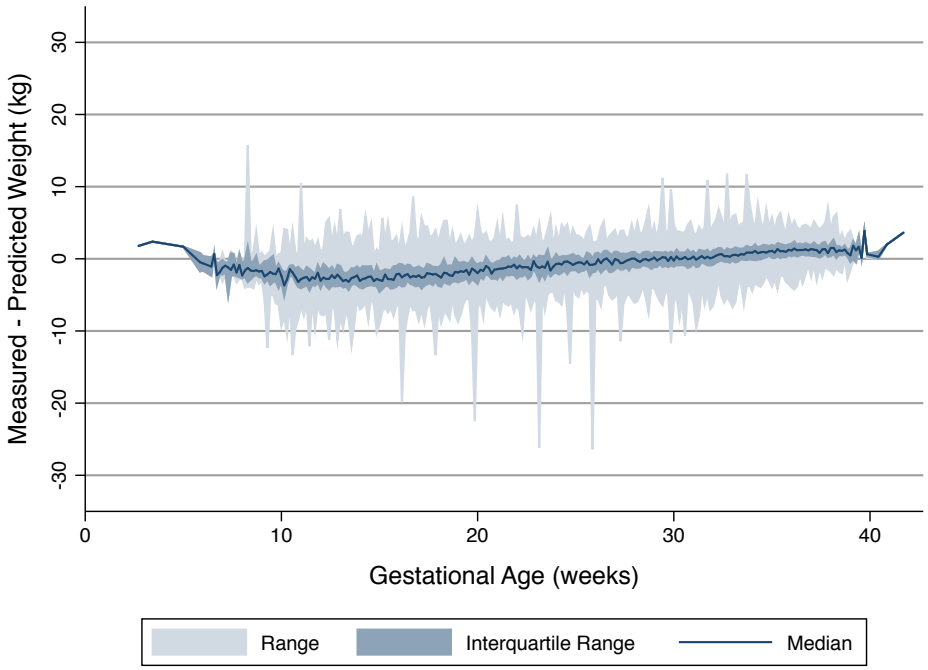


Figure 3-5D. Individual-level linear regression using ln-transformed weight measurements.

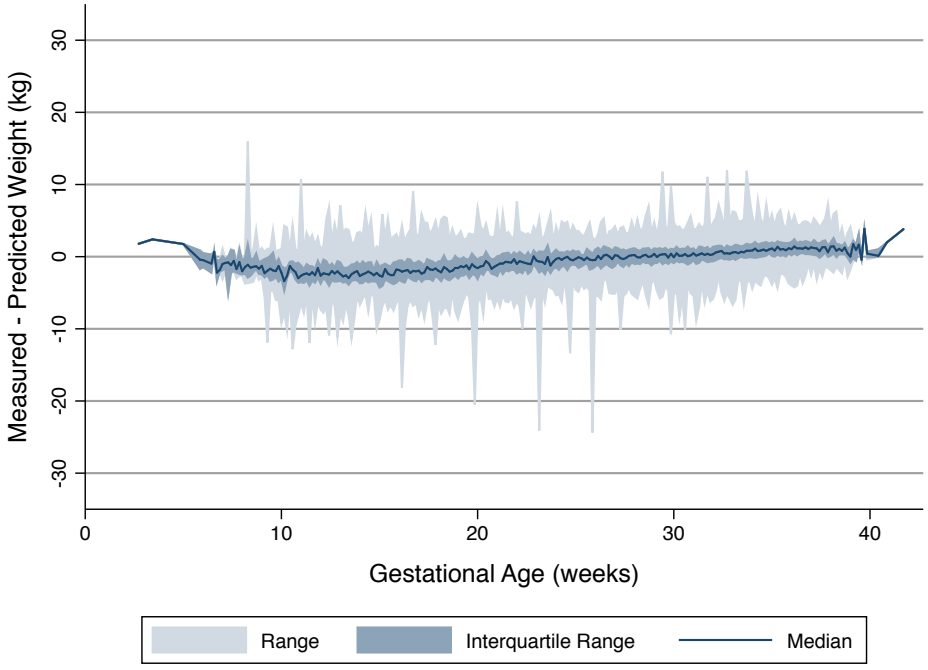


Figure 3-5E. Individual-level linear and quadratic regression.

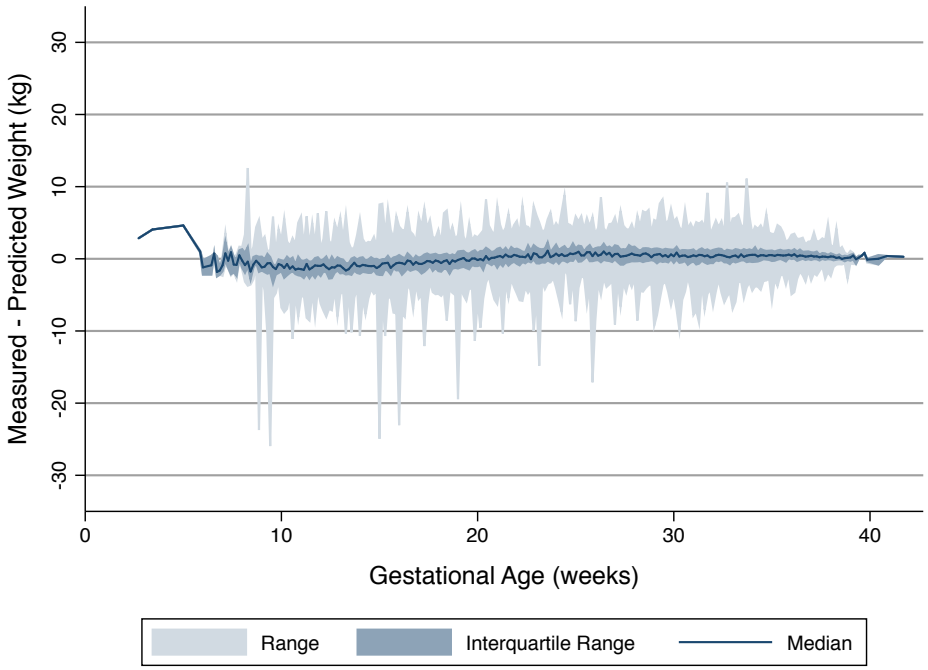


Figure 3-5F. Pooled linear regression with random intercepts.

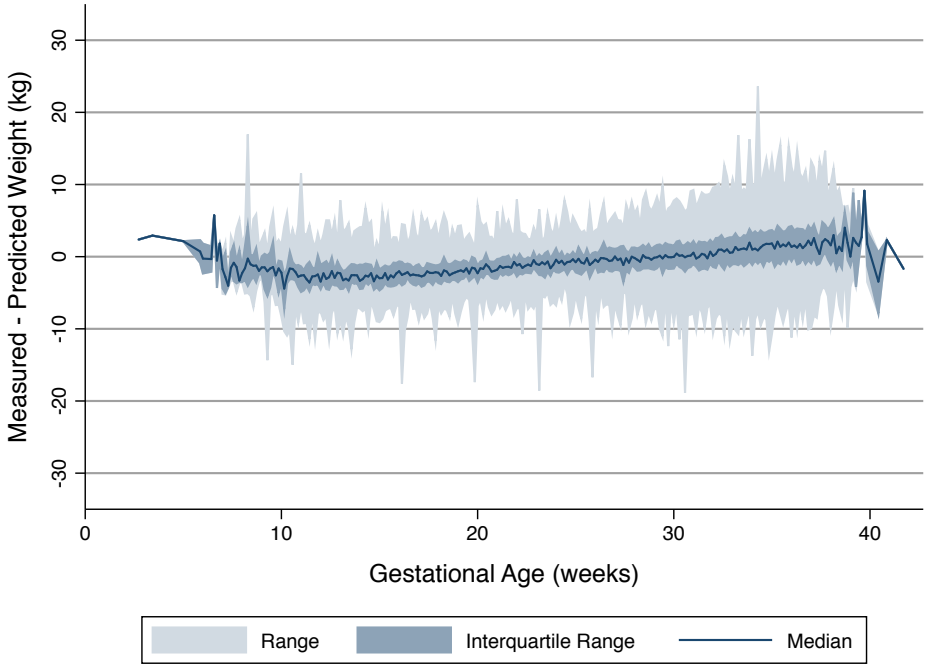


Figure 3-5G. Pooled linear regression with random intercepts and slopes.

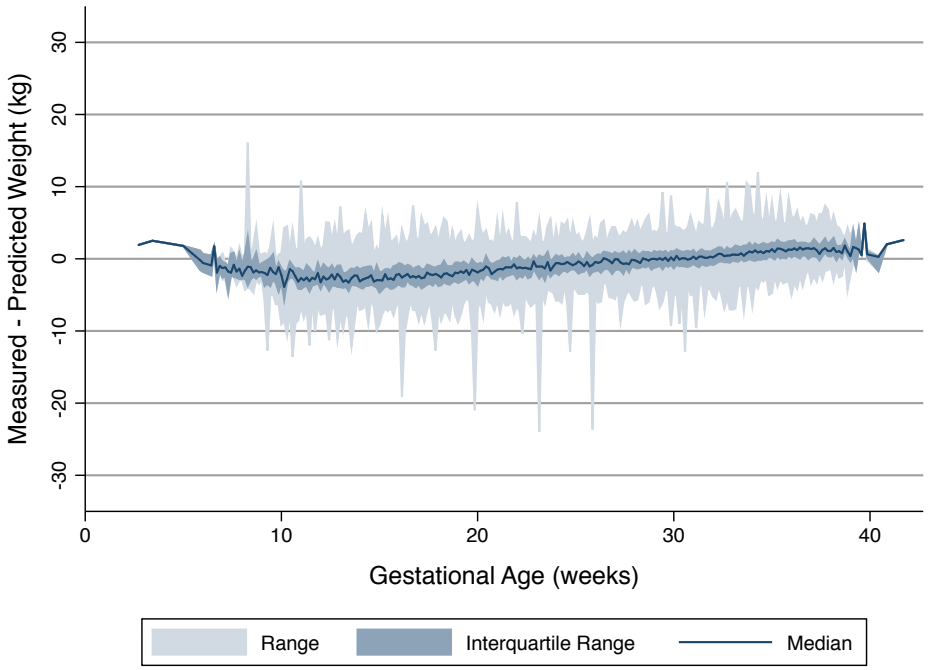


Figure 3-5H. Pooled linear and quadratic regression with random intercepts.

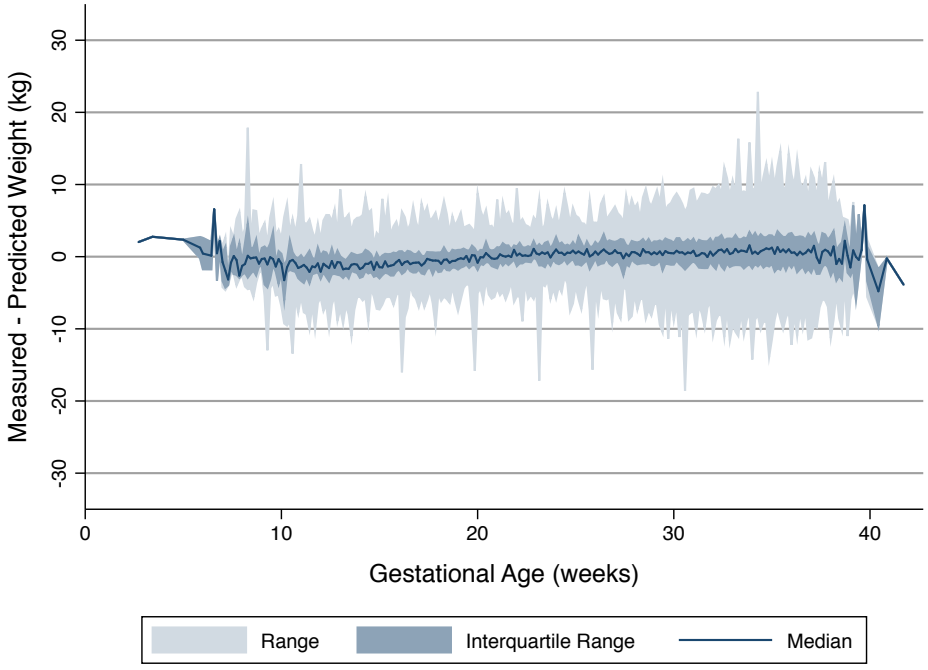


Figure 3-5I. Pooled linear and quadratic regression with random intercepts and slopes.

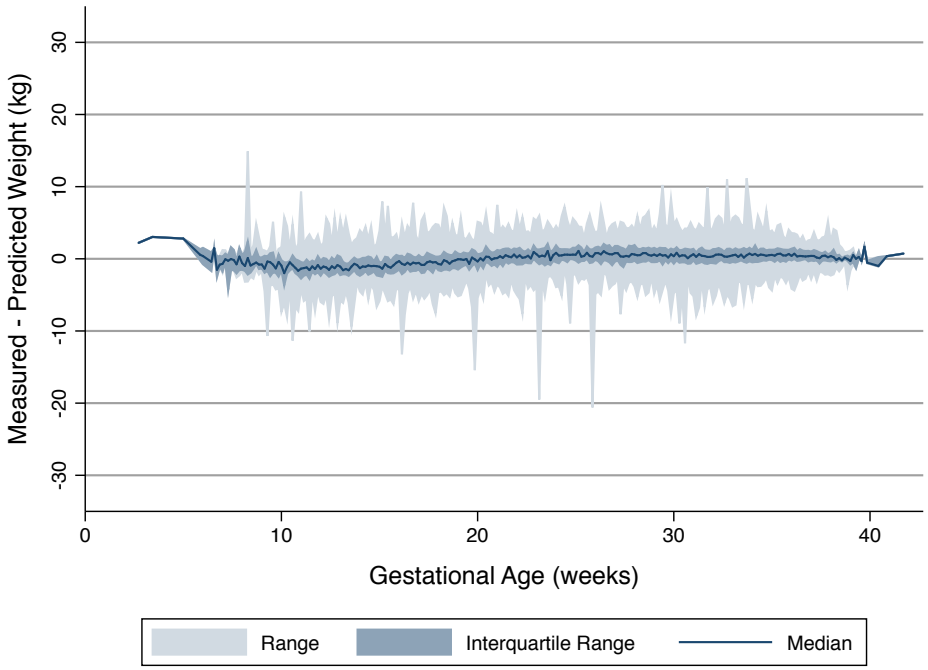


Figure 3-5J. Pooled linear regression using ln-transformed weight measurements with random intercepts.

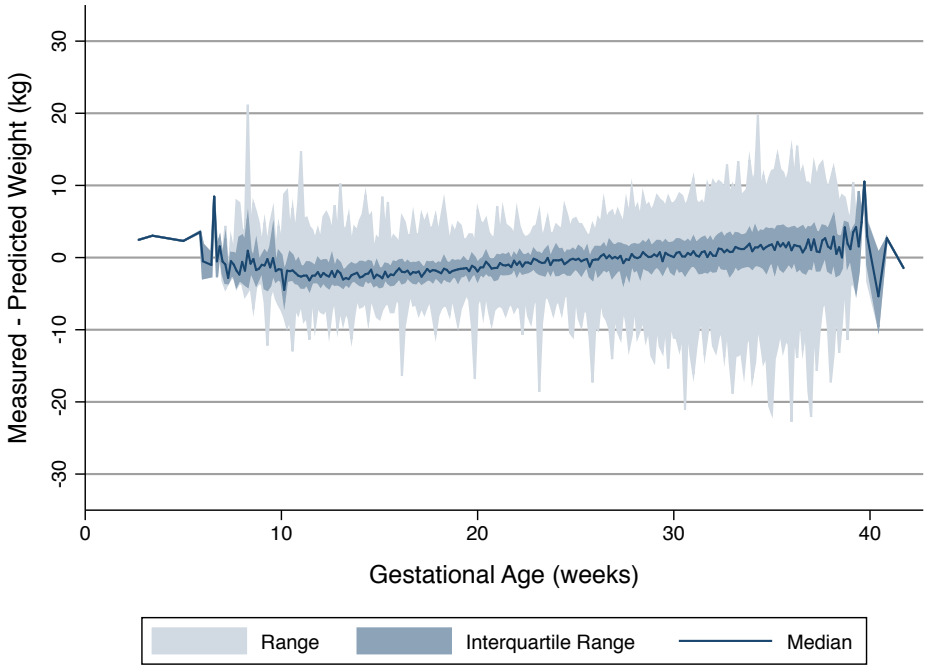


Figure 3-5K. Pooled linear regression using ln-transformed weight measurements with random intercepts and slopes.

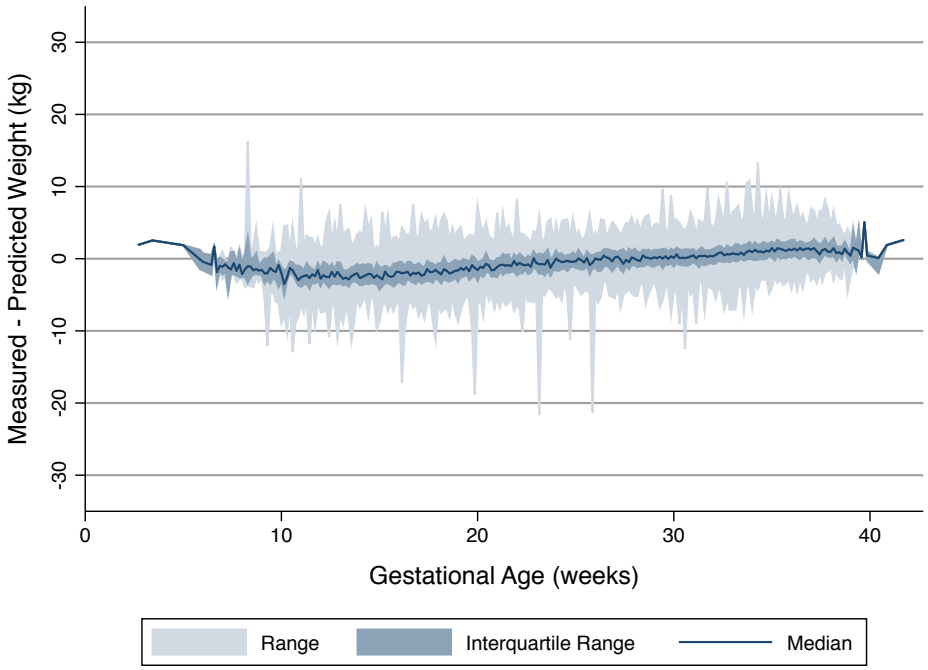


Figure 3-5L. Pooled regression using fractional polynomial terms and random intercepts.

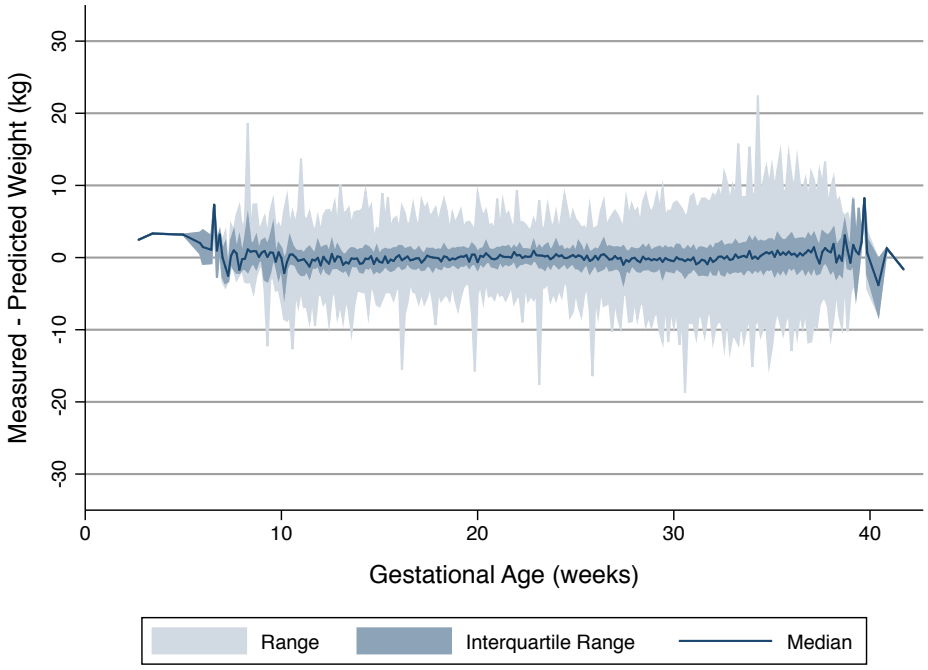


Figure 3-5M. Pooled regression using restricted cubic spline terms with four knots at quantiles and random intercepts.

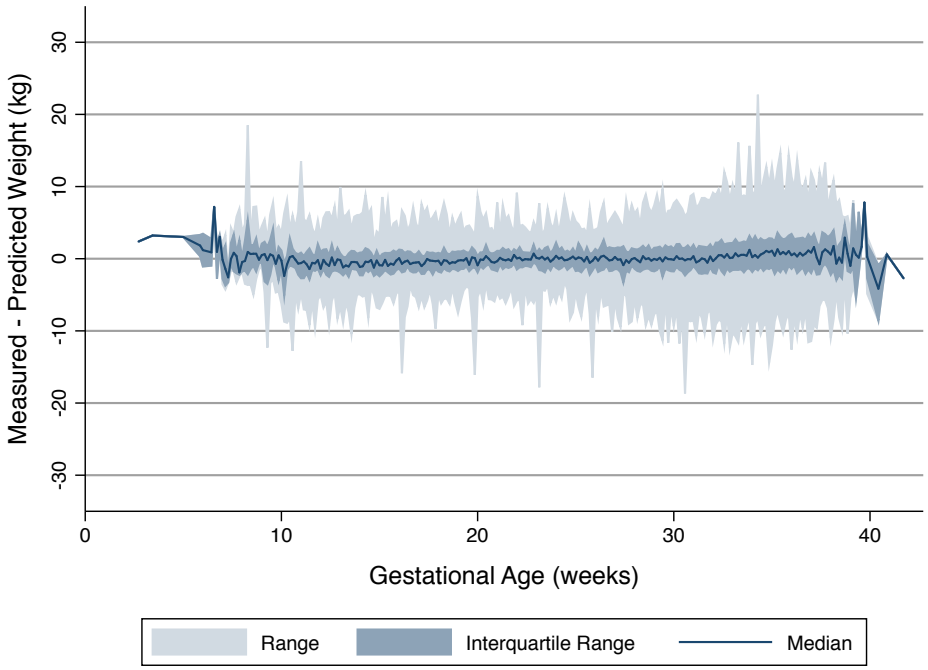


Figure 3-5N. Pooled regression using restricted cubic spline terms with four knots at trimesters and random intercepts.

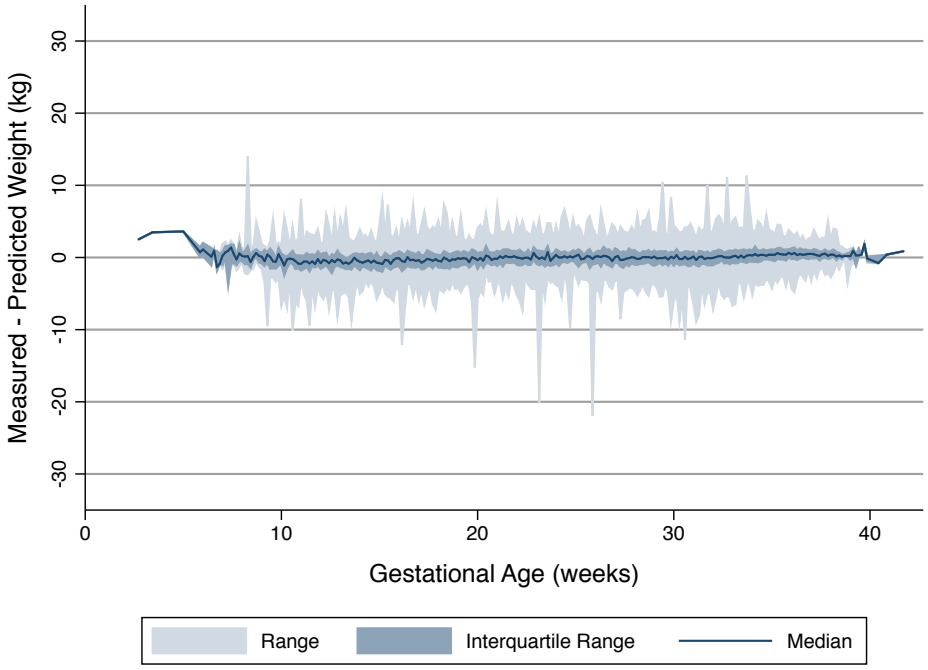


Figure 3-5O. Pooled regression using restricted cubic spline terms with four knots at quantiles and random intercepts and slopes.

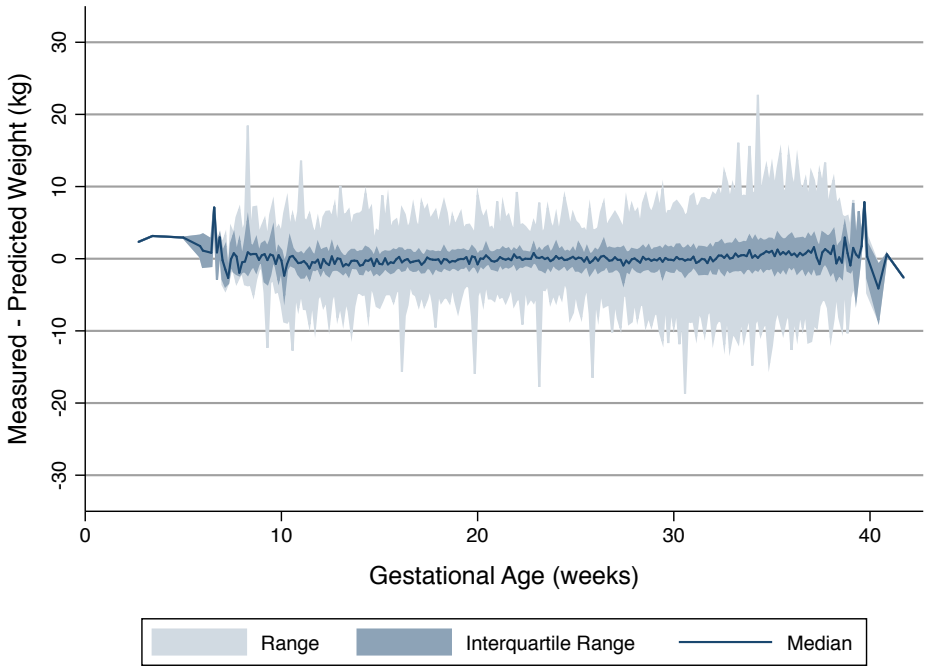


Figure 3-5P. Pooled regression using restricted cubic spline terms with four knots at trimesters and random intercepts and slopes.

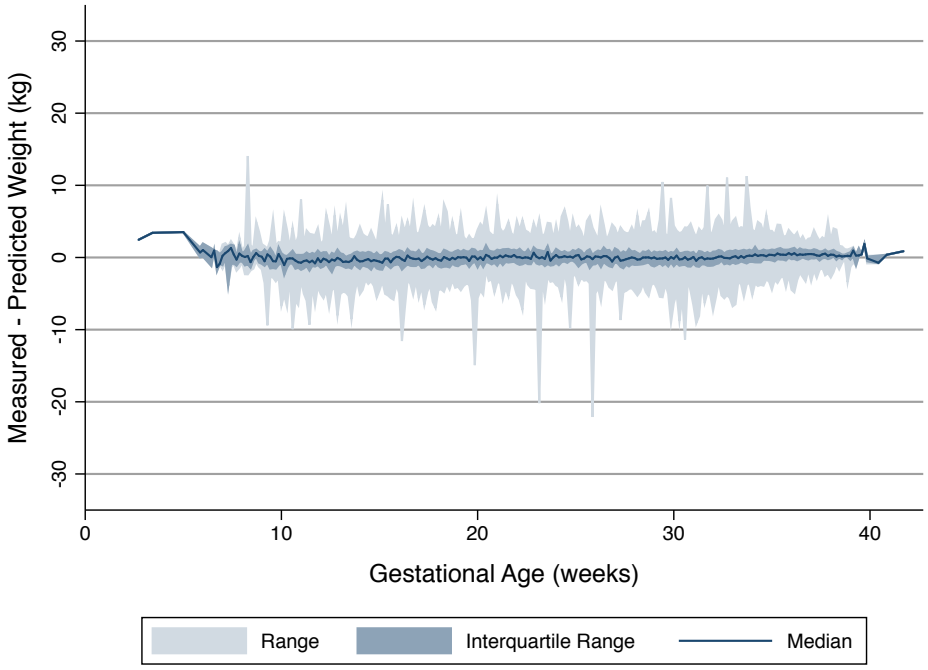


Figure 3-6. Arithmetic difference in measured and predicted maternal weight by gestational age of measurement and interpolation method for singleton pregnancies.

Figure 3-6A. Individual-level last value carried forward.

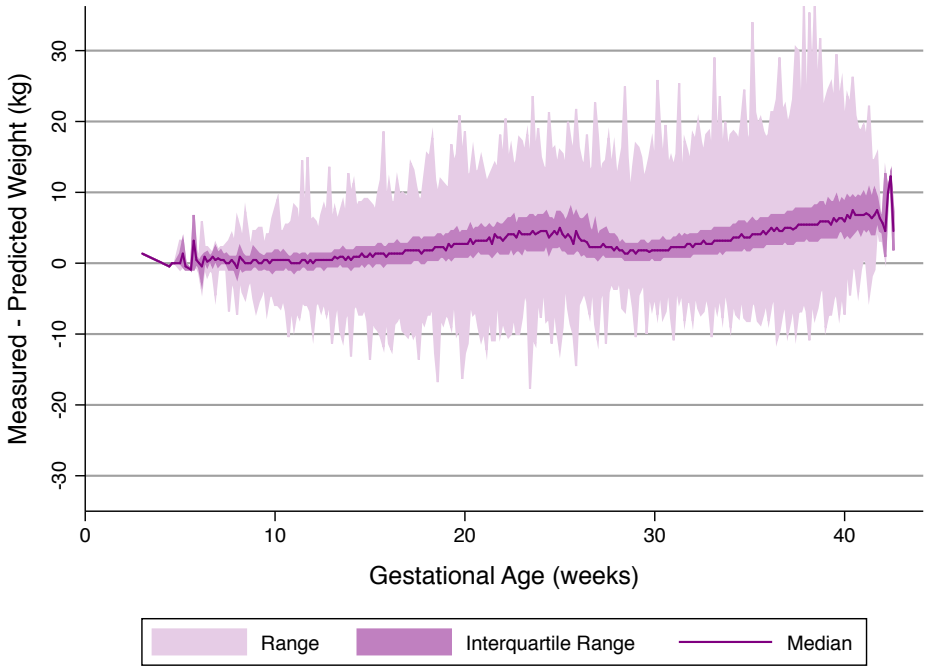


Figure 3-6B. Individual-level linear interpolation between most proximal measurements.

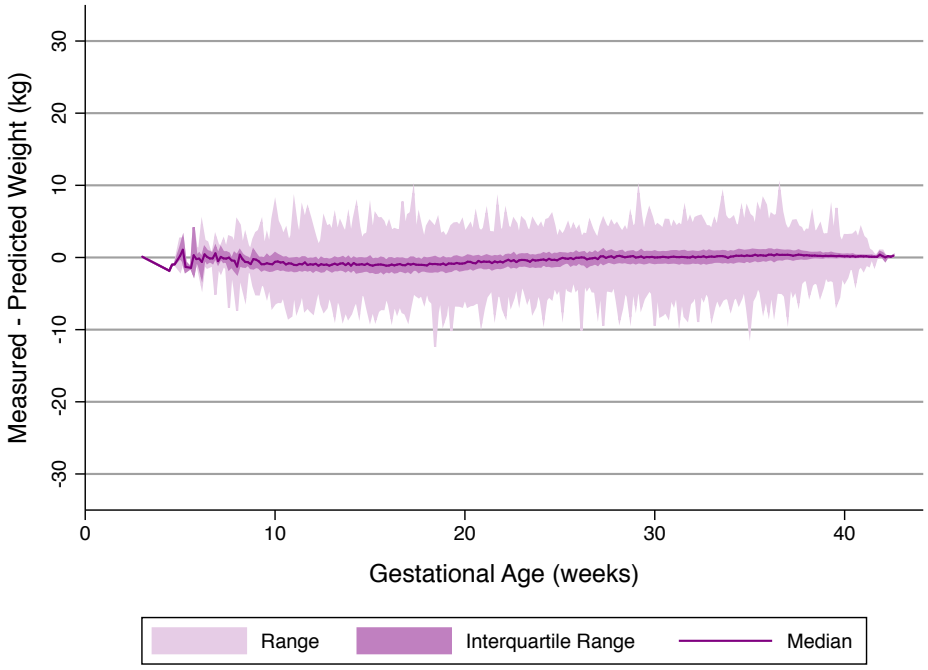


Figure 3-6C. Individual-level linear regression.

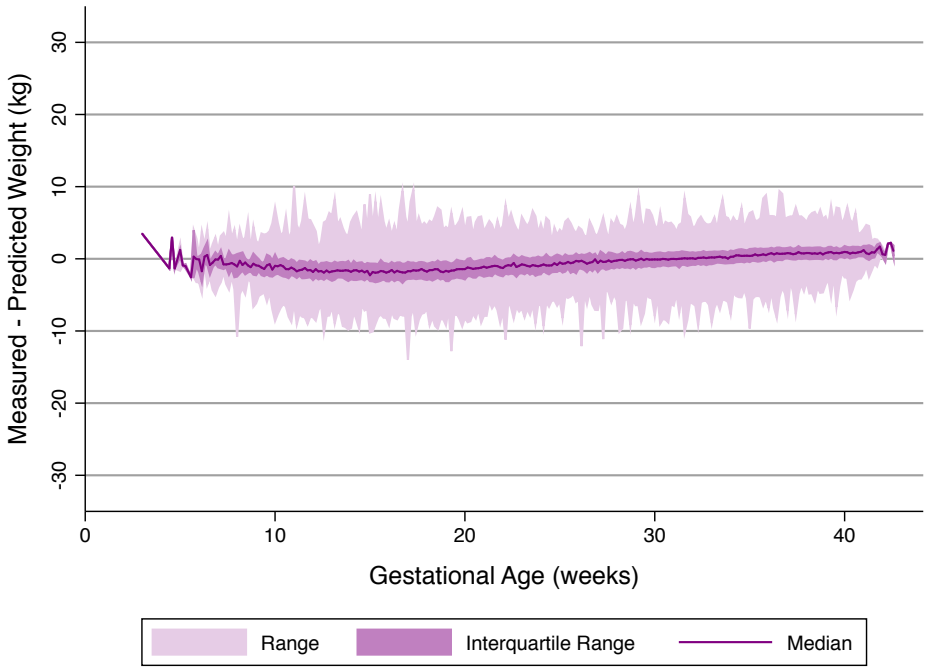


Figure 3-6D. Individual-level linear regression using ln-transformed weight measurements.

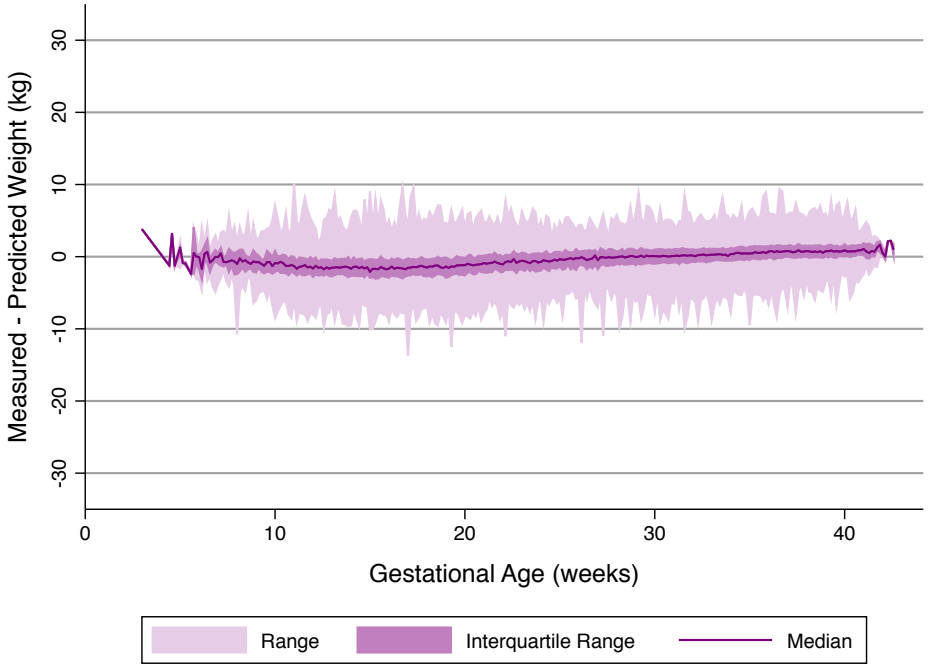


Figure 3-6E. Individual-level linear and quadratic regression.

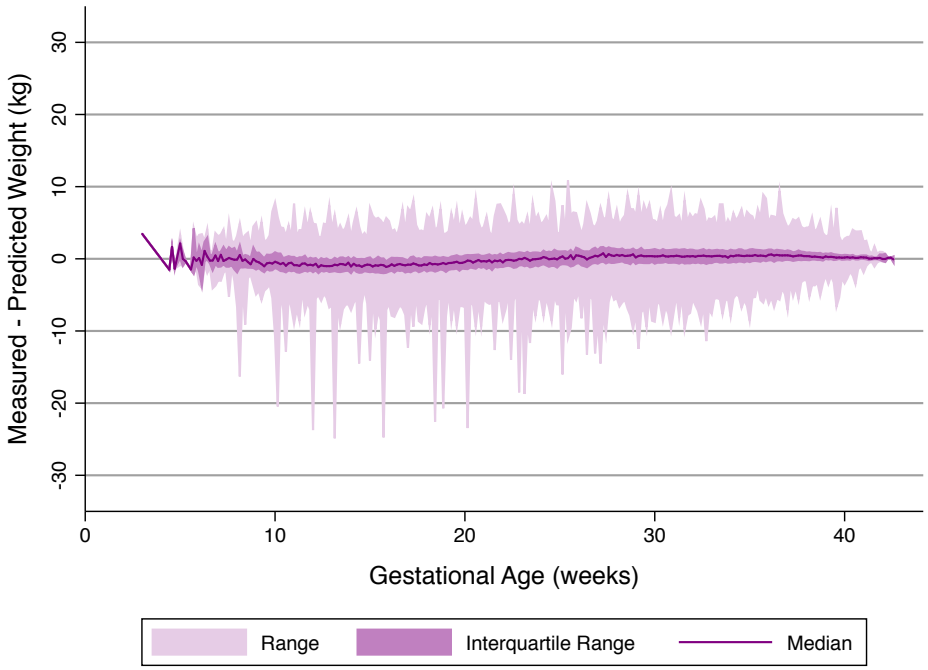


Figure 3-6F. Pooled linear regression with random intercepts.

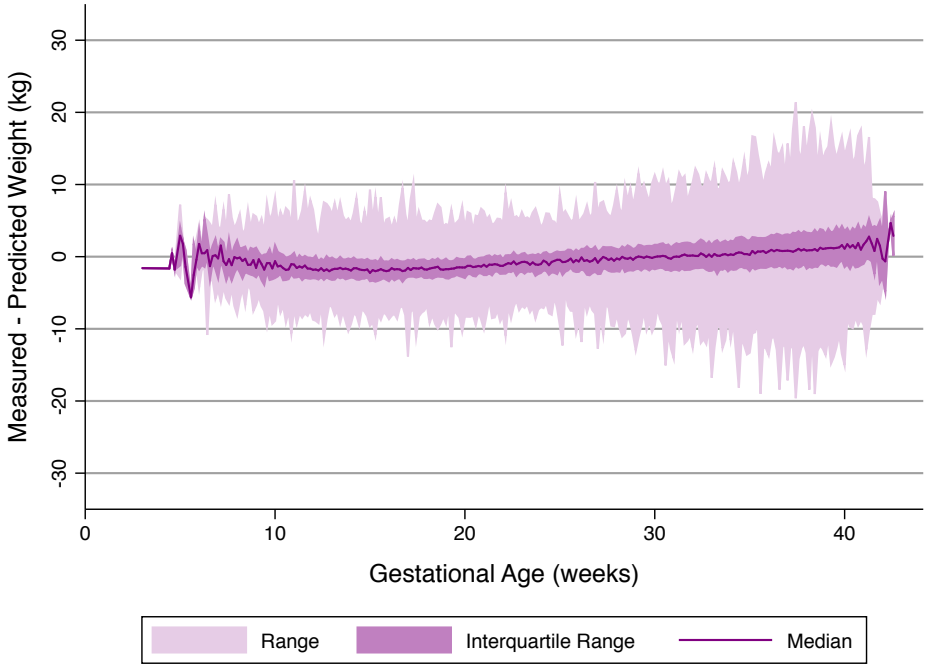


Figure 3-6G. Pooled linear regression with random intercepts and slopes.

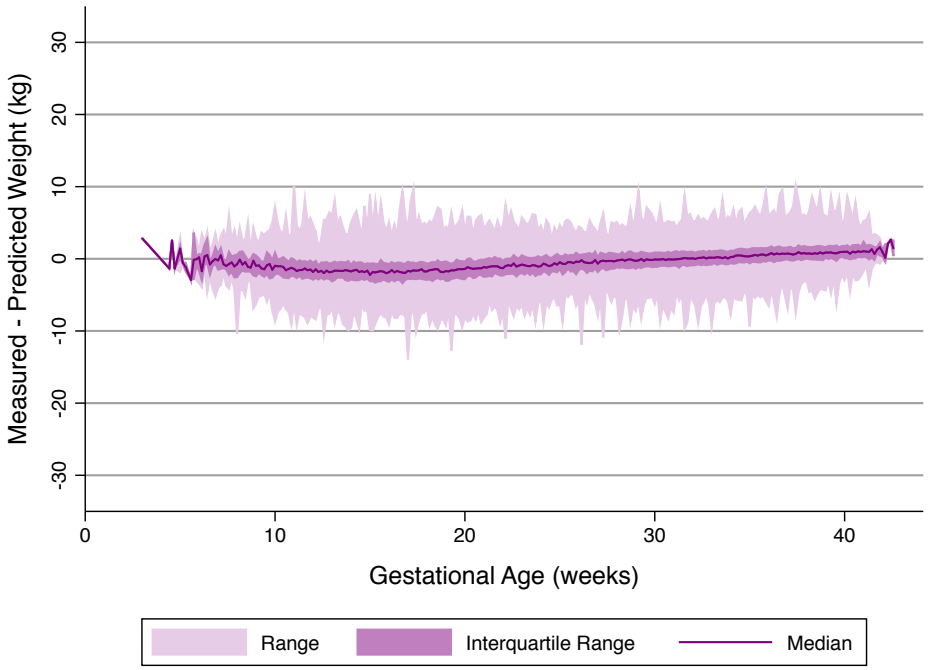


Figure 3-6H. Pooled linear and quadratic regression with random intercepts.

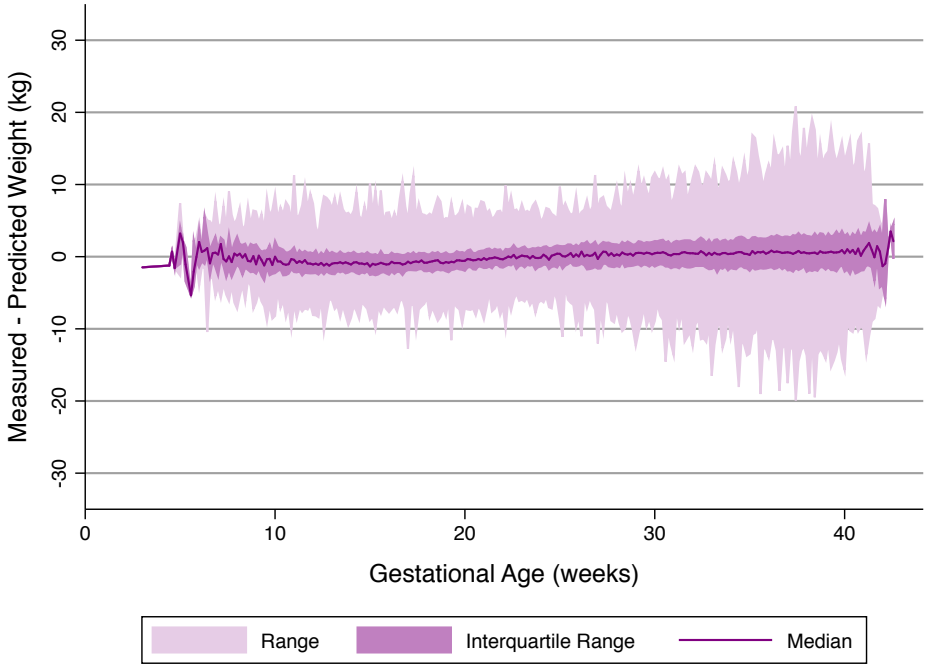


Figure 3-6I. Pooled linear and quadratic regression with random intercepts and slopes.

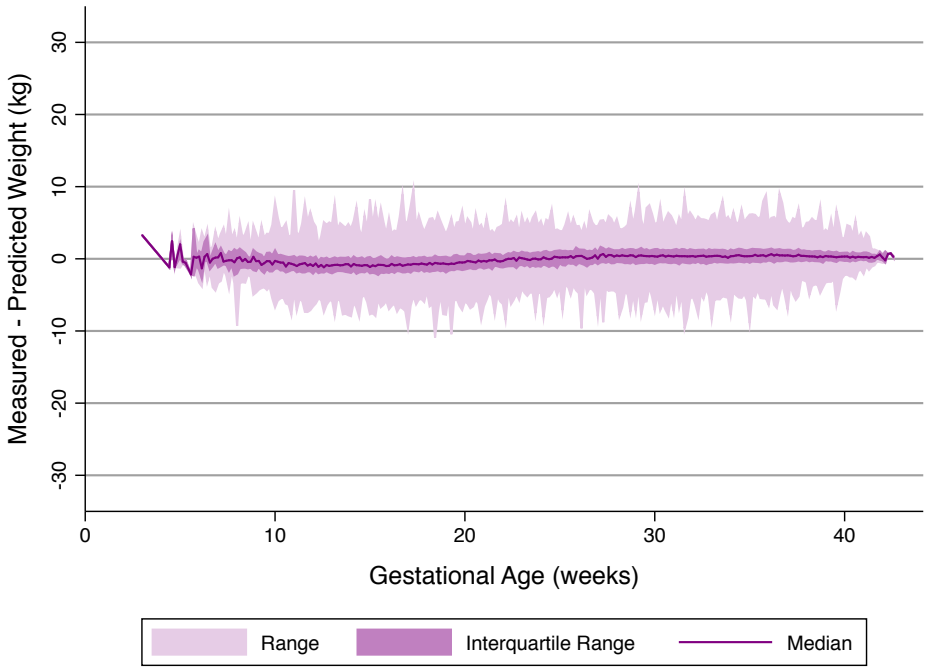


Figure 3-6J. Pooled linear regression using ln-transformed weight measurements with random intercepts.

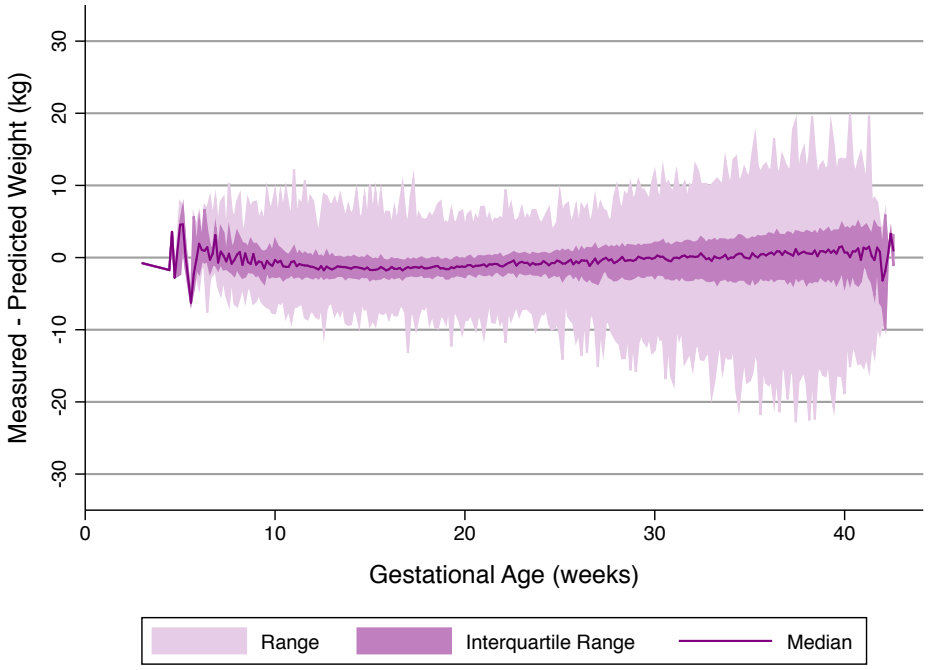


Figure 3-6K. Pooled linear regression using ln-transformed weight measurements with random intercepts and slopes.

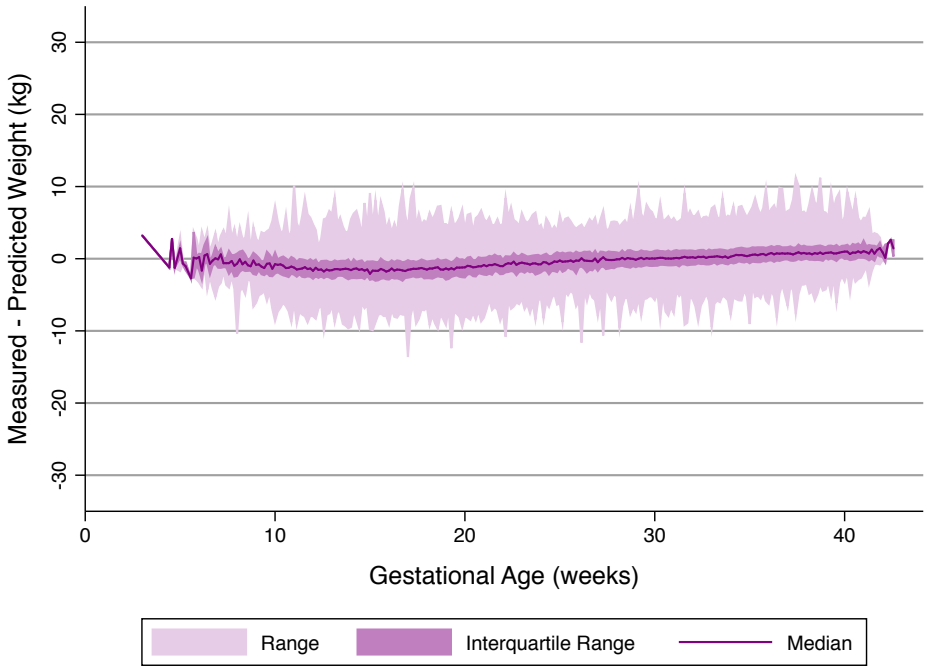


Figure 3-6L. Pooled regression using fractional polynomial terms and random intercepts.

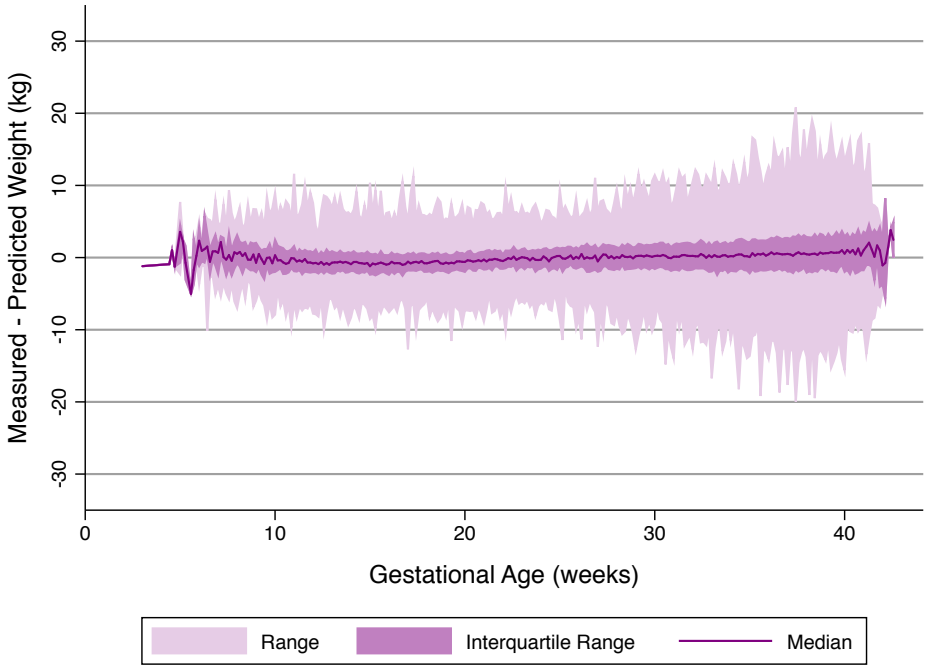


Figure 3-6M. Pooled regression using restricted cubic spline terms with four knots at quantiles and random intercepts.

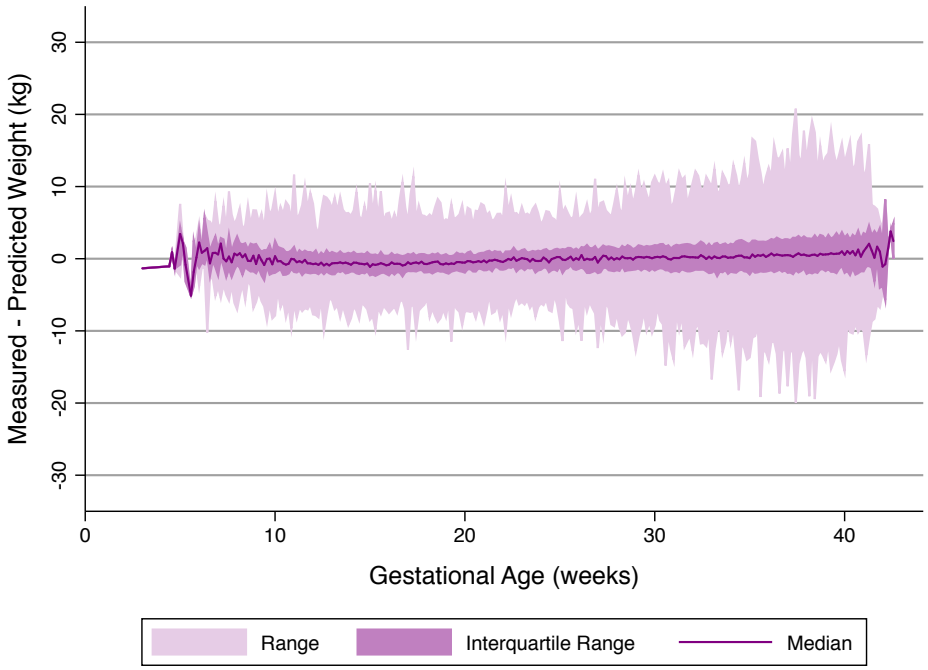


Figure 3-6N. Pooled regression using restricted cubic spline terms with four knots at trimesters and random intercepts.

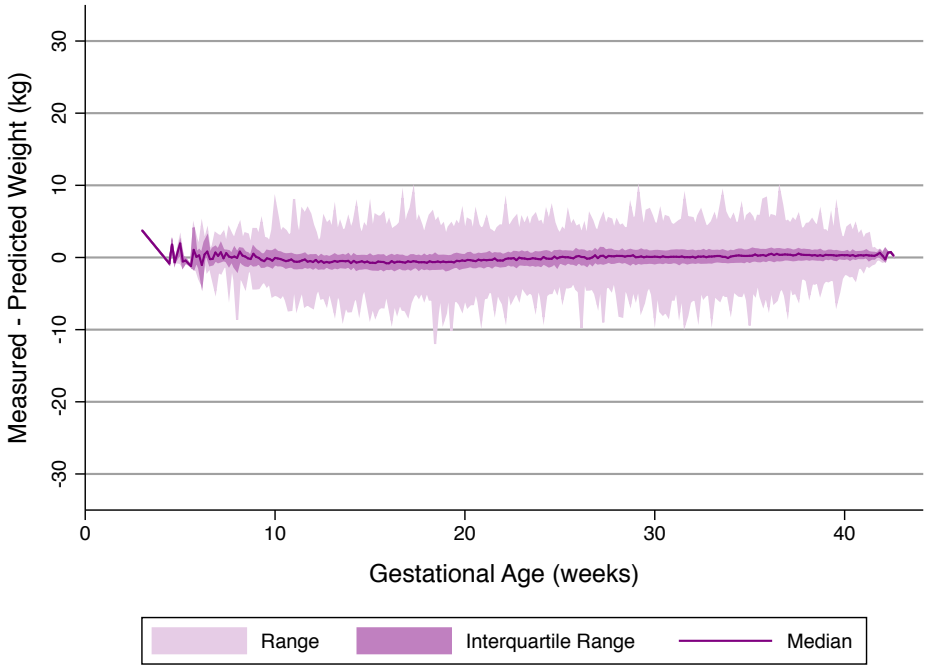


Figure 3-6O. Pooled regression using restricted cubic spline terms with four knots at quantiles and random intercepts and slopes.

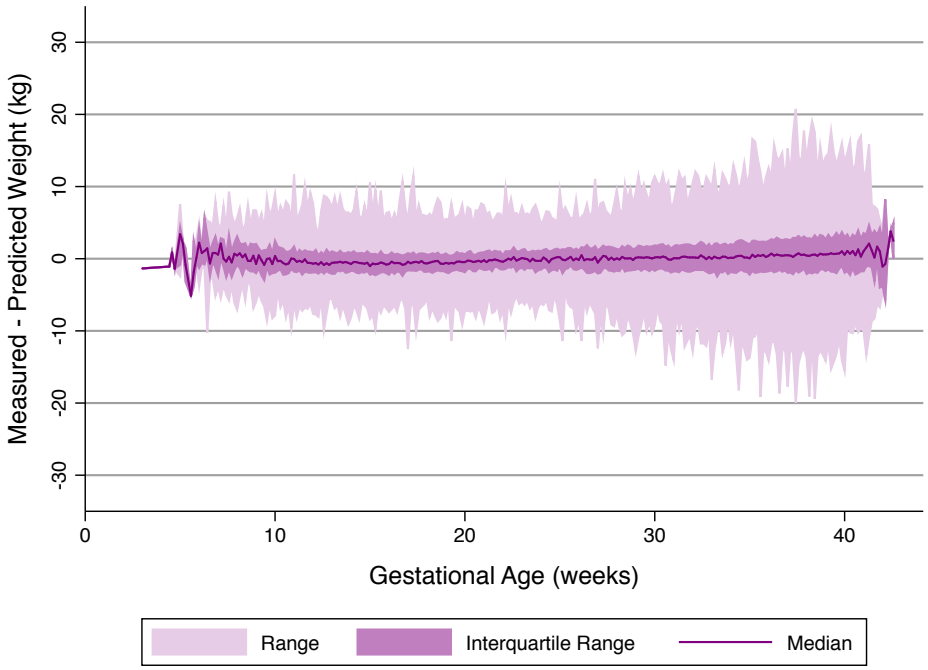
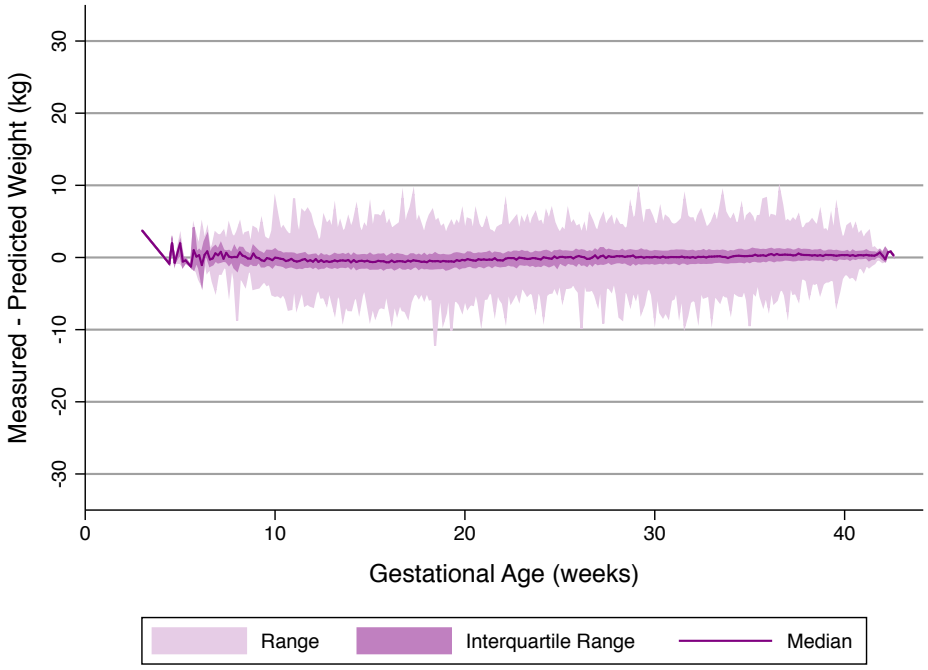


Figure 3-6P. Pooled regression using restricted cubic spline terms with four knots at trimesters and random intercepts and slopes.



3.9. Appendix

Table A3-1. Summary statistics for differences between measured and predicted values of maternal weight in twin pregnancies (n=24400 measurements) when measurements used to fit model are included in summary statistics for differences. PPW refers to pre-pregnancy weight.

Method	Mean (SE)	Median (IQR)	Range	RMSE (Measured PPW)	RMSE (Predicted PPW)
Last Value Carried Forward	3.08 (0.025)	1.81 (0.00; 5.44)	-22.68; 31.75	4.99	4.99
Linear Interpolation, Proximal	-0.22 (0.008)	0.00 (-0.62; 0.24)	-24.10; 13.41	1.33	1.33
Linear Interpolation, All	-0.38 (0.014)	-0.12 (-1.68; 1.05)	-26.41; 15.77	2.17	3.14
Linear Interpolation, Ln-Transformed	-0.24 (0.013)	-0.03 (-1.44; 1.09)	-24.38; 15.96	2.04	2.77
Quadratic Interpolation, All	0.06 (0.010)	0.06 (-0.61; 0.85)	-25.95; 12.58	1.52	1.58
Linear Mixed Effects, Random Intercept	-0.31 (0.021)	-0.52 (-2.38; 1.55)	-23.28; 23.61	3.25	5.56
Linear Mixed Effects, Random Intercept and Slope	-0.36 (0.014)	-0.21 (-1.73; 1.11)	-23.98; 16.09	2.26	3.21
Quadratic Mixed Effects, Random Intercept	0.10 (0.019)	0.07 (-1.65; 1.82)	-24.40; 22.84	2.97	5.19
Quadratic Mixed Effects, Random Intercept and Slope	0.07 (0.010)	0.11 (-0.80; 0.98)	-20.60; 14.91	1.56	1.90
Ln-Transformed Mixed Effects, Random Intercept	-0.23 (0.023)	-0.35 (-2.46; 1.88)	-24.43; 22.98	3.64	6.5
Ln-Transformed Mixed Effects, Random Intercept and Slope	-0.18 (0.014)	-0.07 (-1.49; 1.20)	-21.69; 16.32	2.16	2.97
Fractional Polynomial Mixed Effects, Random Intercept	-0.01 (0.019)	-0.03 (-1.64; 1.64)	-24.73; 22.45	2.92	5.16
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept	0.02 (0.019)	-0.04 (-1.65; 1.66)	-24.52; 22.73	2.92	5.16
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept and Slope	-0.04 (0.009)	0.01 (-0.77; 0.72)	-21.90; 14.06	1.41	1.65

Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept	0.03 (0.019)	-0.02 (-1.62; 1.67)	-24.54; 22.71	2.92	5.16
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept and Slope	-0.03 (0.009)	0.02 (-0.75; 0.72)	-22.10; 14.07	1.39	1.63

Table A3-2. Summary statistics for differences between measured and predicted values of maternal weight in singleton pregnancies (n=95567 measurements) when measurements used to fit model are included in summary statistics for differences. PPW refers to pre-pregnancy weight.

Method	**Mean (SE)**	**Median (IQR)**	**Range**	**RMSE (Measured PPW)**	**RMSE (Predicted PPW)**
Last Value Carried Forward	2.27 (0.011)	1.36 (0.00; 4.08)	-17.69; 36.29	4.03	4.03
Linear Interpolation, Proximal	-0.13 (0.004)	0.00 (-0.50; 0.29)	-12.39; 9.38	1.20	1.20
Linear Interpolation, All	-0.20 (0.006)	-0.03 (-1.23; 0.95)	-14.01; 10.04	1.81	2.50
Linear Interpolation, Ln-Transformed	-0.13 (0.006)	0.02 (-1.13; 0.98)	-13.75; 10.14	1.77	2.35
Quadratic Interpolation, All	0.01 (0.005)	0.03 (-0.63; 0.72)	-24.88; 10.87	1.39	1.46
Linear Mixed Effects, Random Intercept	-0.17 (0.010)	-0.32 (-2.13; 1.63)	-19.61; 21.40	3.21	5.77
Linear Mixed Effects, Random Intercept and Slope	-0.19 (0.006)	-0.07 (-1.29; 1.00)	-13.93; 10.13	1.87	2.60
Quadratic Mixed Effects, Random Intercept	0.05 (0.010)	0.01 (-1.74; 1.80)	-19.91; 20.84	3.11	5.65
Quadratic Mixed Effects, Random Intercept and Slope	0.03 (0.005)	0.07 (-0.76; 0.85)	-10.91; 9.62	1.43	1.73
Ln-Transformed Mixed Effects, Random Intercept	-0.36 (0.012)	-0.41 (-2.60; 1.95)	-23.67; 29.45	3.85	7.22
Ln-Transformed Mixed Effects, Random Intercept and Slope	-0.09 (0.006)	0.00 (-1.17; 1.06)	-13.62; 11.23	1.85	2.54
Fractional Polynomial Mixed Effects, Random Intercept	0.00 (0.010)	-0.04 (-1.77; 1.73)	-19.96; 20.82	3.10	5.65

Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept	0.01 (0.010)	-0.03 (-1.76; 1.74)	-19.94; 20.80	3.09	5.65
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept and Slope	-0.02 (0.004)	0.02 (-0.68; 0.67)	-11.99; 9.24	1.29	1.46
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept	0.01 (0.010)	-0.03 (-1.74; 1.74)	-19.93; 20.76	3.09	5.64
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept and Slope	-0.01 (0.004)	0.02 (-0.68; 0.69)	-12.22; 9.26	1.30	1.49

Table A3-3. Summary statistics for differences between measured and predicted weights in twin pregnancies when pregnancies with outlying weight measurements are excluded.

	Main N=16338		**Conditional %tile, Linear** N=16278		**Conditional %tile, Quadratic** N=16305		**Residual, Linear** N=15804		**Residual, Quadratic** N=16211	
Method	**RMSE**	**Range**	**RMSE**	**Range**	**RMSE**	**Range**	**RMSE**	**Range**	**RMSE**	**Range**
Last Value Carried Forward	6.09	-22.68; 31.75	6.09	-8.62; 31.75	6.09	-22.68; 31.75	6.00	-22.68; 31.75	6.10	-22.68; 31.75
Linear Interpolation, Proximal	1.62	-24.10; 13.41	1.58	-12.51; 6.68	1.61	-24.10; 11.18	1.61	-12.51; 13.41	1.62	-24.10; 13.41
Linear Interpolation, All	2.30	-26.41; 15.77	2.25	-14.55; 7.87	2.30	-26.41; 11.80	2.26	-14.55; 15.77	2.3	-26.41; 15.77
Linear Interpolation, Ln-Transformed	2.17	-24.38; 15.96	2.11	-13.37; 8.20	2.16	-24.38; 11.95	2.13	-13.37; 15.96	2.16	-24.38; 15.96
Quadratic Interpolation, All	1.75	-25.95; 12.58	1.72	-25.95; 8.82	1.74	-25.95; 11.12	1.74	-25.95; 12.58	1.74	-25.95; 12.58
Linear Mixed Effects, Random Intercept	3.20	-18.87; 23.61	3.17	-14.99; 23.61	3.19	-18.86; 23.63	3.15	-18.89; 23.61	3.20	-18.89; 23.58
Linear Mixed Effects, Random Intercept and Slope	2.37	-23.98; 16.09	2.32	-13.61; 11.99	2.36	-23.98; 11.99	2.32	-13.61; 16.07	2.36	-23.98; 16.09
Quadratic Mixed Effects, Random Intercept	2.87	-18.60; 22.84	2.83	-14.26; 22.83	2.86	-18.58; 22.85	2.83	-18.62; 22.86	2.87	-18.61; 22.82

Quadratic Mixed Effects, Random Intercept and Slope	1.70	-20.60; 14.91	1.65	-11.32; 7.97	1.69	-20.59; 11.19	1.70	-11.73; 15.07	1.69	-20.56; 14.86
Ln-Transformed Mixed Effects, Random Intercept	3.54	-22.74; 21.19	3.51	-22.80; 19.72	3.53	-22.73; 19.75	3.51	-22.78; 21.12	3.55	-22.80; 21.20
Ln-Transformed Mixed Effects, Random Intercept and Slope	2.26	-21.69; 16.32	2.22	-12.95; 13.40	2.25	-21.68; 13.39	2.22	-12.95; 16.29	2.26	-21.69; 16.31
Fractional Polynomial Mixed Effects, Random Intercept	2.81	-18.76; 22.45	2.78	-15.15; 22.44	2.80	-18.75; 22.46	2.80	-18.71; 22.72	2.81	-18.78; 22.42
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept	2.82	-18.67; 22.73	2.78	-14.72; 22.72	2.81	-18.65; 22.74	2.79	-18.68; 22.76	2.82	-18.68; 22.71
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept and Slope	1.56	-21.90; 14.06	1.47	-9.50; 7.79	1.55	-21.88; 11.31	1.56	-11.44; 14.23	1.54	-21.92; 13.98
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept	2.82	-18.70; 22.71	2.78	-14.79; 22.70	2.80	-18.68; 22.72	2.78	-18.71; 22.74	2.82	-18.71; 22.69
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept and Slope	1.55	-22.10; 14.07	1.46	-9.61; 7.81	1.53	-22.11; 11.21	1.54	-11.43; 14.23	1.53	-22.10; 13.98

Table A3-4. Summary statistics for differences between measured and predicted weights in singleton pregnancies when pregnancies with outlying weight measurements are excluded.

	Main N=65286		**Conditional %tile, Linear** N=65120		**Conditional %tile, Quadratic** N=65119		**Residual, Linear** N=64124		**Residual, Quadratic** N=64781	
Method	**RMSE**	**Range**	**RMSE**	**Range**	**RMSE**	**Range**	**RMSE**	**Range**	**RMSE**	**Range**
Last Value Carried Forward	4.87	-17.69; 36.29	4.85	-17.69; 36.29	4.86	-17.69; 36.29	4.85	-17.69; 36.29	4.87	-17.69; 36.29
Linear Interpolation, Proximal	1.46	-12.39; 9.38	1.45	-12.39; 9.38	1.45	-12.39; 9.38	1.46	-12.39; 9.38	1.45	-12.39; 9.38

Linear Interpolation, All	1.91	-14.01; 10.04	1.90	-14.01; 10.04	1.91	-14.01; 10.04	1.9	-14.01; 10.04	1.91	-14.01; 10.04
Linear Interpolation, Ln-Transformed	1.87	-13.75; 10.14	1.86	-13.75; 10.14	1.86	-13.75; 10.14	1.86	-13.75; 10.14	1.86	-13.75; 10.14
Quadratic Interpolation, All	1.59	-24.88; 10.87	1.58	-24.88; 9.38	1.58	-24.88; 10.87	1.59	-24.88; 9.81	1.58	-24.88; 10.87
Linear Mixed Effects, Random Intercept	3.18	-19.61; 21.40	3.16	-19.61; 21.40	3.17	-19.61; 21.40	3.17	-19.62; 21.39	3.18	-19.63; 21.38
Linear Mixed Effects, Random Intercept and Slope	1.96	-13.93; 10.13	1.96	-13.93; 10.13	1.96	-13.93; 10.12	1.95	-13.93; 10.10	1.96	-13.93; 10.11
Quadratic Mixed Effects, Random Intercept	3.05	-19.91; 20.84	3.04	-19.91; 20.85	3.05	-19.91; 20.84	3.05	-19.91; 20.85	3.05	-19.93; 20.83
Quadratic Mixed Effects, Random Intercept and Slope	1.57	-10.91; 9.62	1.56	-10.91; 9.62	1.56	-10.97; 9.61	1.57	-10.65; 9.68	1.56	-10.91; 9.63
Ln-Transformed Mixed Effects, Random Intercept	3.79	-22.80; 19.94	3.78	-22.80; 19.93	3.78	-22.80; 19.93	3.78	-22.80; 19.94	3.79	-22.82; 19.93
Ln-Transformed Mixed Effects, Random Intercept and Slope	1.95	-13.62; 11.23	1.94	-13.62; 11.24	1.94	-13.62; 11.22	1.93	-13.63; 11.05	1.94	-13.62; 11.19
Fractional Polynomial Mixed Effects, Random Intercept	3.04	-19.96; 20.82	3.03	-19.96; 20.83	3.04	-19.96; 20.83	3.04	-19.96; 20.83	3.04	-19.98; 20.81
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept	3.04	-19.94; 20.80	3.02	-19.93; 20.80	3.03	-19.94; 20.80	3.04	-19.93; 20.81	3.04	-19.95; 20.78
Cubic Splines Mixed Effects, Knots at Quantiles, Random Intercept and Slope	1.45	-11.99; 9.24	1.44	-11.98; 9.24	1.43	-12.19; 9.25	1.46	-11.63; 9.22	1.44	-12.00; 9.23
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept	3.03	-19.93; 20.76	3.02	-19.93; 20.77	3.03	-19.93; 20.76	3.03	-19.93; 20.78	3.03	-19.95; 20.75
Cubic Splines Mixed Effects, Knots at Trimesters, Random Intercept and Slope	1.45	-12.22; 9.26	1.44	-12.22; 9.27	1.43	-12.49; 9.28	1.46	-11.82; 9.25	1.44	-12.21; 9.26

Connecting Manuscripts 1 and 2

In Manuscript 1, I emphasize that advanced statistical methods, including survival analyses, require estimates of gestational weight gain between measurements collected at prenatal visits. In Manuscript 2, I use the best interpolation method identified for both twins and singletons to estimate gestational weight gain between all prenatal visits. I then examine the impact of time-varying gestational weight gain on gestational age at birth using a time-to-event framework and incorporating competing risks for type of delivery.

4. Manuscript 2

4.1. Title

Investigating the Relationship Between Maternal Weight Gain and Gestational Duration in Twin and Singleton Pregnancies

4.2. Abstract:

Gestational weight gain predicts adverse maternal and infant health outcomes and is rarely studied in twin pregnancies. In singletons, both low and high gestational weight gain are correlated with preterm birth. In lieu of cumulative measures often used in studies of this relationship, such as total or rate of weight gain, the Institute of Medicine has recommended capturing patterns or trajectories of gestational weight gain in future research. We aimed to investigate the relationship between time-varying gestational weight gain and gestational age at birth in twin pregnancies and compare to that in singletons delivered in the same study population. We used serial weight measurements abstracted from medical charts and routinely collected maternal and pregnancy characteristics for a cohort of twin and case-cohort of singleton pregnancies delivered from 1998 and 2014 at Magee Women's Hospital in Pittsburgh, Pennsylvania. Gestational weight gain was expressed as time-varying z-scores calculated using pre-pregnancy BMI-stratified twin- and singleton-specific charts, while gestational age at birth was assessed as a time-to-event outcome, with delivery following spontaneous onset of labour, induced labour, and prelabour Caesarean considered competing risks. We analyzed the relationship using an extension of the Cox proportional hazards model that allows both non-linear and time-dependent effects of continuous covariates. We found a similar u-shaped relationship between gestational weight gain and gestational age at birth as previously described in singletons; we additionally observe that this is due to increased risk of early preterm spontaneous delivery among pregnancies with low gestational weight gain, and increased risk of late preterm delivery with no labour among pregnancies with high gestational weight gain. Our findings have

implications for Institute of Medicine guidelines for gestational weight gain in twin pregnancies, which are currently provisional.

4.3. Introduction

Gestational weight gain (GWG), or maternal weight gain during pregnancy, is a potentially modifiable pregnancy characteristic[37] correlated with maternal and infant outcomes[2,44,67]. Previous research among singleton pregnancies indicate that both low and high GWG are associated with increased risk of preterm birth[1,38]. Twin pregnancies have increased in recent years[4,5], and experience greater GWG[49] and higher incidence of poor perinatal outcomes, including preterm birth[4,17]. Despite this, the contribution of GWG to perinatal health in twin pregnancies, and its relationship to that in singletons, is understudied[58].

Several studies have aimed to quantify the effect of GWG on preterm birth in singletons, however, most failed to account for the inherent correlation between total GWG and gestational duration[83–85]. Specifically, longer gestational duration generally yields greater total GWG. Since rate of weight gain is higher in the second and third trimesters, studies that conceptualize GWG as average rate of change across the entire pregnancy experience similar methodological shortcomings[3]. Studies of cumulative measures of GWG and preterm birth produce biased estimates and are therefore difficult to interpret.

Studies that account for the link between total GWG and gestational duration find modest u-shaped relationships between both relatively low and high GWG and preterm birth among singleton pregnancies. A large meta-analysis that examined combined effects of pre-pregnancy BMI and total GWG, conceptualized as z-score using an internal reference, found that low GWG, was associated with 1.15 to 1.82-fold increased odds of preterm birth, while high GWG was correlated with 1.23 to 2.94-fold increased odds of preterm birth, compared to pregnancies with normal weight BMI and average GWG[38]. A study among women with normal weight and overweight pre-pregnancy BMI found u-

shaped relationships of similar magnitude and noted that effects were attenuated when using z-score charts to account for gestational duration compared to cumulative total measures[53].

Recently, researchers have recommended investigating preterm birth, and more broadly gestational age (GA) at birth, within a time-to-event framework[27]. Instead of examining preterm birth as a binary outcome, time-to-event analyses capture the continuum of gestational duration, from extreme, severe, and moderate prematurity[17] to late preterm, early to late term, and post term[22]. Additionally, time-to-event analyses easily incorporate time-varying exposures, namely GWG per day of GA, as well as competing risks, such as deliveries following spontaneous versus induced onset of labour. Preserving original measures of both GWG and GA at birth may increase both precision and relevance of effect estimates[27].

We are aware of two studies that have investigated GWG and preterm birth in a time-to-event framework[27,86]. One conceptualized GWG in kilograms linearly interpolated between measurements as the exposure of interest[27], while the other investigated area under the GWG curve (termed "pound-days"), where GWG was estimated between visits using restricted cubic splines and random intercepts and slopes[86]. Both studies found that high GWG was associated with lower relative hazard of delivery[27,86], while the latter made additional conclusions about the pattern of GWG. Specifically, this study noted that earlier accumulated pound-days decreased hazard of delivery, though first trimester GWG was not considered in the analysis[86]. Importantly, neither study incorporated non-linear relationships between time-varying GWG and time-to-delivery, despite the previously described u-shaped relationship between total GWG and preterm birth. Neither study investigated types of delivery as competing risks, and either grouped all types of delivery or focused on relationships in spontaneous deliveries. Importantly, neither study included twin pregnancies in the analysis.

We aimed to investigate the relationship between time-varying GWG and GA at birth in twin pregnancies and compare it to that in singletons. Since twins are less common than singletons, and are

typically delivered at earlier gestational ages, we were especially interested in similarities in the shape of the relationship between GWG and GA at birth, as well as its magnitude across GA. We hypothesized that twin and singleton pregnancies differ on the range of GWG associated with minimal risk of delivery (i.e. higher in twins) as well as the range of GAs at which effects are strongest (i.e. earlier in twins).

4.4. Methods

We analyzed data assembled for a cohort of diamniotic twin and case-cohort of singleton pregnancies delivered from 1998-2013 at Magee-Women's Hospital in Pittsburgh, Pennsylvania. Parent studies selected all twin pregnancies delivered from 1998-2013, and a random sample of singleton pregnancies within six strata of pre-pregnancy BMI (underweight, normal weight, overweight, obese class I, obese class II, obese class III). Pregnancies resulting in preterm birth were also over-selected in the parent singleton case cohort, though we examined gestational age at birth continuously in the current study. Serial weight measurements were abstracted from clinical charts, while pre-pregnancy/delivery weight and maternal/pregnancy characteristics were obtained from the Magee Obstetric Maternal and Infant Database and supplemented with vital statistics information collected by the state of Pennsylvania. Details of both twin cohort and singleton case-cohort have been previously described.[47,69]

We assessed all available pregnancies for inclusion in the current study. We excluded the second twin pregnancy per mother and twin pregnancies with monochorionic placentation. We additionally excluded any pregnancies with missing/implausible maternal pre-pregnancy weight, height, delivery weight, or gestational age at birth, no available serial weight measurements, presence of any gestational weight gain measurements that could not be converted to gestational age-standardized weight gain z-scores using twin- and singleton-specific charts, or missing key covariates. Since comparing the shape/strength of relationship in twins to that among singleton pregnancies was a

secondary objective of this study, we restricted singleton analyses to women with normal weight pre-pregnancy BMI. Twin and singleton pregnancies were analyzed in separate statistical models.

Our primary exposure of interest was time-varying GWG, which we calculated in kilograms as the delivery or prenatal weight minus the pre-pregnancy weight. For gestational ages at which weight was not measured (i.e. days between prenatal visits), we estimated GWG by pooling all pregnancies within plurality and using linear regression with restricted cubic splines for gestational age, knots at trimesters, and random intercepts and slopes for pregnancy to obtain estimates through interpolation; we previously found this interpolation method to most accurately and precisely estimate GWG between measurements.

We then converted time-varying GWG to GA-standardized z-scores using plurality- and pre-pregnancy BMI-specific charts. We standardized GWG because a key assumption of our statistical model is that the shape of the relationship between exposure and outcome is constant, although the strength can vary over the time scale[87]. We anticipated that time-varying GWG as an arithmetic measure may violate this assumption; for example, a weight gain of 10 kg accumulated by the second versus third trimester may have different implications for gestational duration. Conversely, we expected that a GA-standardized GWG z-score at or near zero may consistently be associated with minimal risk for delivery across GA, with the exception of at/after term, where GWG z-score at or near zero may be associated with optimally-timed delivery. We calculated GA-standardized GWG z-scores by comparing natural log-transformed GWG using twin- and singleton-specific charts stratified by pre-pregnancy BMI.[47,88] Since twin- and singleton-specific charts were developed within pre-pregnancy BMI categories, we anticipated that effect measure modification by pre-pregnancy BMI, which has been observed in studies of GWG and preterm birth, may be avoided using this strategy. We tested this and other assumptions in sensitivity analysis.

Our primary outcome was gestational duration, which we analyzed as time-to-delivery in days of GA. First, we defined an event as any type of delivery. Second, we considered delivery following spontaneous onset of labour, induction, and deliveries without labour (i.e. prelabour Caesarean) as competing risks; specifically, we assumed ongoing pregnancies were at risk for any delivery, while occurrence of one type precluded risk for the other two types of delivery. In these models, we defined an event as type of delivery of interest, for example spontaneous onset of labour, and censored pregnancies at occurrence of the other types of delivery, for example induced deliveries or deliveries without labour; we repeated this process for each type of delivery.

We investigated relationships using a flexible extension of the Cox proportional hazards model[87], otherwise known as the product method, which incorporates both non-linear and time-dependent effects of GWG on time-to-delivery; this model allowed us to simultaneously quantify the shape of the relationship across GWG z-score and the strength of the relationship across GA. In particular, we could discern whether the relationship between GWG and gestational duration was u-shaped, as previously described in singletons, and whether it was more prominent at certain gestational ages. We considered 20 weeks' GA as time zero, and adjusted models for a pre-defined list of covariates, namely year of delivery (continuous; linear/quadratic terms), parity (nulliparous/primiparous or multiparous), maternal age (continuous; linear/quadratic terms, with quadratic term centred at 25), pre-pregnancy body mass index (continuous; linear/quadratic terms, with quadratic term centred at 21.7 kg/m^2), race (white, non-Hispanic black, Hispanic/other), education (less than high school, high school, some college/associates, college graduate), insurance (private/other, Medicaid/self-pay), marital status (married/unmarried), smoking status (ever/never), pre-existing hypertension, diabetes, polycystic ovary syndrome, and use of assisted reproductive technology (yes/no). We generated bootstrapped 95% confidence intervals with 100 repetitions . We fit the following model, where $\lambda(GA|GWG, C_i)/ \lambda_0(GA)$ is relative hazard of delivery, $\beta(GA)$ is a function for

the strength of the relationship or time-dependent effect, $r(GWG_{GA})$ is a function for the shape of the relationship or non-linear effect, and $\Sigma\alpha_i C_i$ are functions for covariate effects:

$$\lambda(GA|GWG,\ C_i) = \lambda_0(GA)e^{\beta(GA)r(GWG_{GA}) + \sum \alpha_i C_i}$$

We compared fit of models with time-varying GWG in kilograms as the exposure, with/without non-linear and/or time-dependent effects, with different number of knots and/or degrees of splines, and incorporating effect measure modification of pre-pregnancy BMI by multiplying GWG by pre-pregnancy BMI and adjusting for this using either linear/quadratic terms centred on 21.7 kg/m^2, which is the midpoint of the normal weight pre-pregnancy BMI category, or functions for non-linear/time dependent effects generated by the product method. We assessed whether these modifications substantially improved model fit by comparing model fit statistics, including Akaike information criteria (AIC), which represents model deviance adjusted for number of parameters[89]. For AIC, a lower value indicated better model fit, and differences of more than 4 to 10 are typically interpreted as being meaningful[90]. For both twin and singleton pregnancies, we censored at median GA (36.3 weeks' GA for twins and 38.9 weeks' GA for singletons) to confirm that the shape of the relationship was similar in early versus late pregnancy. We additionally censored at optimal gestation for twin and singleton pregnancies, or 37 and 38 weeks' respectively[24–26], in sensitivity analyses.

Analyses were performed with Stata 14.2[78], R 3.5.1[91], and Microsoft Excel 16.26. Ethical approval was obtained from University of Pittsburgh and University of British Columbia for parent studies, and from McGill University Faculty of Medicine Institutional Review Board for the current study.

4.5. Results

We retained 2021 twin and 1852 singleton pregnancies in the primary analysis (Figures 4-1 and 4-2). We excluded 243 twin pregnancies with missing/implausible pre-pregnancy weight, height, delivery weight, gestational age, and/or serial weight measurements; we additionally excluded 9

pregnancies with GA at birth/fetal death on or prior to 20 weeks' GA, 1 pregnancy with at least one serial weight (either measured or imputed) on or after 20 weeks' gestational age that could not be converted to z-score using existing twin-specific chart (i.e. GWG less than or equal to the constant added prior to log-transformation in the authors' formula[47]), and 35 pregnancies with missing covariates. Overall, we retained greater than 85% of the parent dichorionic twin cohort .

Table 4-1 displays the characteristics of women with twin pregnancies stratified by quartiles of total GWG z-score calculated using the twin-specific chart. Generally, women in the lowest total GWG z-score quartile were less educated, more frequently of non-Hispanic black/other race, less frequently married, more frequently insured by Medicaid, less frequently nulliparous, and more frequently classified as ever-smoker. Women in the lowest quartile were also more likely experience spontaneous delivery and stillbirth. Conversely, pregnancies in the highest total gestational weight gain z-score were more frequently nulliparous and more likely to be classified as having pre-existing hypertension and diabetes. Overall, deliveries following spontaneous onset of labour were most frequent (56.9%), while no labour and deliveries following induction of labour were less common (26.7% and 16.5%, respectively). Median total gestational weight gain ranged from 9.5 kg in the lowest to 24.9 kg in the highest quartile, while median and interquartile range of gestational age at birth appeared similar across quartiles.

Table 4-2 displays the characteristics of normal weight pre-pregnancy BMI women with singleton pregnancies stratified by quartiles of total GWG z-score calculated using a singleton-specific chart. Crude relationships observed between covariates and GWG z-score quartile were similar to those described in twins, with the exception of stillbirth and pre-existing hypertension, both of which were rare across quartiles. In singletons, spontaneous deliveries were most frequent (56.4%), followed by induced deliveries (37.6%). Deliveries without labour were rare in singleton pregnancies (6.0%).

Figure 4-3 displays relationships between time-varying gestational weight gain z-score and relative hazard of delivery (any, spontaneous, induced, and no labour) at 28 to 38 weeks' gestational age for twin pregnancies. When all types of delivery were considered, we observed an increased relative hazard of delivery at earlier gestational ages among women with low GWG, and an increased relative hazard of delivery at later gestational ages among women with high GWG. Among twins, the lowest hazard of any type of delivery was correlated with a z-score of 0.9, which corresponds to several values of GWG depending on pre-pregnancy BMI and GA by which GWG occurred. For example, at 30 weeks' GA, z-score of 0.9 corresponds to approximately 20 kg in underweight/normal weight/overweight women, and 17 kg among women with obese pre-pregnancy BMI. Similarly, at 36 weeks' GA, a z-score of 0.9 corresponds to approximately 25 kg in underweight/normal weight women, 26 kg in overweight women, 23 kg in obese women.

When only deliveries following spontaneous onset of labour were considered and induced/no labour deliveries were censored at birth, we observed an increased relative hazard of delivery at 28-34 weeks among women with low GWG, which gradually attenuated as gestational age increased. Relative hazard of spontaneous delivery increased monotonically at z-scores at or less than 1.1; this corresponds to 20-24 kg among underweight/normal weight women, 19-25 kg among overweight women, and 18-23 kg among obese women gained between 28-34 weeks' GA, respectively.

When only deliveries without labour were considered, we observed a monotonically increased relative hazard of delivery from 32-38 weeks, which was most precisely estimated at 36 weeks. When only deliveries following induction of labour were considered, we observed a shape similar to that for deliveries without labour, but relationships were statistically null. Relative hazard by GWG for deliveries following spontaneous onset of labour at 38 weeks, induced deliveries and deliveries without labour at 28-30 weeks, as well as all delivery types at 40 weeks not shown due to the paucity of events at these gestational ages in twin pregnancies.

Relationships in singleton pregnancies exhibited both similarities and differences when compared to twin pregnancies (Figure 4-4). Similar to twins, low GWG was correlated with increased relative hazard of spontaneous delivery at 30-32 weeks' GA among singleton pregnancies. Conversely, both relatively low and high GWG were associated with increased hazard of induced delivery at 36-38 weeks' GA in singletons. Minimum relative hazard of induced delivery was observed at a z-score of -0.3, or approximately 13-14 kg of GWG at 36-38 weeks' GA, respectively, among women with normal weight pre-pregnancy BMI. Inferences regarding the relationship between GWG and time-to-delivery without labour were unclear due to the rarity of this outcome, particularly at early gestational ages. Relationships between GWG and relative hazard of any delivery at 28 weeks' GA, induced/no labour delivery at 30-32 weeks' GA, and no labour delivery at 34-36 weeks' GA are not shown due to paucity of events in singleton pregnancies.

Among twins, model fit improved when using GWG z-score versus GWG in kg and incorporating both non-linear and time-dependent effects (Table 4A-1). Models with 1 knot/2 degrees or 2 knots/3 degrees fit better than those with 2 knots/2 degrees and 1 knot/3 degrees. Since the model with 2 knots/3 degrees produced similar shapes of effect at common z-scores as the model with 1 knot/2 degrees yet produced implausible shapes of effect at extreme z-scores, we chose to model relationships using splines with 1 knot/2 degrees. Incorporating effect modification by pre-pregnancy BMI appeared important when investigating GWG in kg, but less important when examining GWG z-score as the exposure of interest. Furthermore, addition of pre-pregnancy BMI as a covariate did not improve model fit when examining GWG z-score, likely because z-score is already calculated within strata of pre-pregnancy BMI. Model fit in singletons followed similar patterns, although AIC was lowest when examining GWG in kg and incorporating pre-pregnancy BMI as both a covariate and effect modifier. Since singleton analyses was already restricted to normal weight pre-pregnancy BMI, we did not include precise BMI as a covariate or effect modifier in primary analyses. Model fit

improved with the addition of all covariates that may confound the relationship between GWG z-score and time-to-delivery in both twin and singleton pregnancies. Additionally, censoring pregnancies at either median GA or 37 weeks for twins (n=805 or 39.8% censored) and 38 weeks for singletons (n=1123 or 60.6% censored) did not substantially change the shapes of observed effects.

4.6. Discussion

We found considerable evidence for increased risk of preterm birth among twin pregnancies with relatively low GWG, and limited evidence of increased risk of preterm birth among those with relatively high GWG, which mimics the u-shaped relationship commonly reported in singletons. Our findings appear to suggest two separate phenomena: (1) Increased risk of early spontaneous preterm delivery among twin pregnancies with low GWG and (2) Increased risk of no labour preterm delivery among those with high GWG. We find evidence for both non-linear and time-dependent effects of GWG on GA at birth among twins.

We used innovative methods to study GWG as a time-varying covariate, allowed effects to vary by GA, and treated of delivery as competing risks; we are unaware of any studies that have taken this approach, particularly among twin pregnancies. While our methods are novel, our findings are consistent with prior research in singletons , yet strengthen the evidence base for GWG in twin pregnancies. Specifically, our results suggest minimum hazard of induced delivery among singleton pregnancies of normal weight women that gain 13-14 kg by 36-38 weeks' GA; this is squarely within the range of 11.5 to 16 kg total GWG recommended by current IOM guidelines[49]. Conversely, our results suggest that GWG in the range of 20 to 25 kg between 28 and 36 weeks' GA is associated with minimal hazard of delivery among twin pregnancies in normal weight women; this is on the higher end of current provisional guidelines, which recommend 17 to 25 kg total GWG for twin pregnancies of women with normal weight pre-pregnancy BMI[49]. Our results are consistent with research that

suggests a u-shaped relationship between GWG and GA; we found analogous relationships between GWG and GA at birth in twins and singletons.

Results of our study also strengthen the rationale for studying both GWG and GA at birth as continuous measures. We studied GWG as a time-varying continuous covariate rather than examining cumulative total or rate of GWG or as a ratio of adequacy with current guidelines. Our exposure conceptualization, namely GWG for GA z-score calculated using twin- and singleton-specific charts, facilitates meaningful interpretation of effects associated with one-unit increases in GWG across gestational age. GWG is typically conceptualized in kilograms, although a one-kilogram increase is not equally meaningful in the first versus second/third trimesters.

Previous research suggests increased risk of spontaneous delivery among pregnancies with low GWG and increased risk of induced delivery among pregnancies with high GWG when examining these outcomes separately in singleton pregnancies.[44] Our study corroborates these findings in twin pregnancies while also considering types of delivery together in as competing risks. Furthermore, our study identifies potential critical periods for effects of GWG; specifically, relatively low GWG confers higher risk of spontaneous delivery at 28 to 32 weeks' GA, while relatively high GWG presents higher risk of delivery with no labour only at 34-36 weeks' GA. This may have important implications for the clinical monitoring of GWG in both twin and singleton pregnancies.

Our study had many strengths, including exposure and outcome definitions that maintained precision of GWG and GA at birth instead of using categorical measures. We also leveraged a new extension of the Cox proportional hazards model to evaluate both non-linear and time-dependent effects of GWG, which followed from similar research in singletons employing a time-to-event framework[27]. Instead of producing a single effect estimate in the context of a time-to-event framework, our model allowed us to simultaneously examine the strength and shape effects. We analyzed a wealth of data from two large cohorts of twin and singleton pregnancies from the same study population,

including prenatal weight measurements abstracted from medical charts. Our data sources and approach allowed us to compare results in twin and singleton pregnancies.

Although our statistical approach is advanced in comparison to other studies, the extended Cox proportional hazards method is limited in that it assumes a single shape of effect across gestational age. For example, if the relationship between GWG and relative hazard of delivery was linear at 28 weeks' GA and bimodal at 32 weeks' GA, the model may inaccurately specify the relationship as a mix of these two functions. We feel this is unlikely to have occurred, particularly since our findings mirror those found in previous studies in singletons. We additionally addressed this limitation by converting arithmetic GWG to GA-specific z-scores to improve model stability. Given the increase in parameters needed, we lacked power to fully incorporate effect modification by pre-pregnancy BMI; however, we calculated GWG z-scores using BMI-stratified charts, thus reducing the need for such parameters. Our approach also necessitated an estimate for GWG per day of GA, which we estimated using available pre-pregnancy, delivery, and prenatal weights. Interpolating between measurements produces some error; however, the method we employed produced error that was non-differential with respect to both GA and total number of measurements, which would likely bias results towards the null. It is possible that estimates of gestational age differed with respect to accuracy/precision among twins and singletons in our study population. For example, gestational age estimates were less precise in twins than in singletons, this may comparatively bias effect estimates among twins towards the null; if gestational age estimates were less accurate in twins than in singletons, this may comparatively bias effect estimates for twins either towards or away from the null. However, prior research in in vitro fertilization pregnancies note that ultrasound-based gestational age estimation performs with similar accuracy and precision in twin and singleton pregnancies[82]. Our model simultaneously estimated several parameters, including coefficients for covariates and spline terms for non-linear/time-dependent effects of GWG, which is statistically more complex than typical approaches. However, our strategy

was informed by varying effects of GWG among different subtypes of preterm birth, within different strata of pre-pregnancy BMI, and at different ranges of GWG and GA that has been reported in previous studies. Therefore, we feel our approach more accurately reflects the complexity of the relationship between GWG and preterm birth.

We conclude that both low and high GWG are associated with GA at birth among twin pregnancies. Additionally, we highlight that increased risk for early spontaneous preterm delivery among pregnancies with low GWG and increased risk for late preterm birth without labour jointly contribute to this relationship. Our results mirror the relationship found in singletons, while also strengthening the evidence base for GWG guidelines in twin pregnancies.

4.7. Tables

Table 4-1. Sample characteristics by total GWG z-score in twin pregnancies. Continuous variables presented as median (25th percentile; 75th percentile), categorical variables presented as number (percentage), and covariates at the level of the mother/pregnancy unless otherwise stated.

Characteristic	**1 (-9.3;-0.8)**	**2 (-0.8;-0.2)**	**3 (-0.2;0.4)**	**4 (0.4;3.0)**	**All**
N	506	505	505	505	2021
Maternal Age (years)[a]	30.4 (6.5)	30.6 (5.7)	30.8 (5.6)	30.2 (5.8)	30.5 (5.9)
Maternal Education					
<High School	39 (7.7)	24 (4.8)	21 (4.2)	28 (5.5)	112 (5.5)
High School/GED	117 (23.1)	102 (20.2)	83 (16.4)	108 (21.4)	410 (20.3)
Some College/Associates	116 (22.9)	116 (23.0)	95 (18.8)	114 (22.6)	441 (21.8)
College Graduate	234 (46.2)	263 (52.1)	306 (60.6)	255 (50.5)	1058 (52.4)
Race					
Non-Hispanic White	360 (71.1)	394 (78.0)	430 (85.1)	414 (82.0)	1598 (79.1)
Non-Hispanic Black	117 (23.1)	91 (18.0)	59 (11.7)	79 (15.6)	346 (17.1)
Hispanic	3 (0.6)	6 (1.2)	2 (0.4)	6 (1.2)	17 (0.8)
Other	26 (5.1)	14 (2.8)	14 (2.8)	6 (1.2)	60 (3.0)
Married					
Yes	328 (64.8)	347 (68.7)	382 (75.6)	355 (70.3)	1412 (69.9)
No	178 (35.2)	158 (31.3)	123 (24.4)	150 (29.7)	609 (30.1)
Insurance					
Private/Other	296 (58.5)	326 (64.6)	336 (66.5)	318 (63.0)	1276 (63.1)
Medicaid/Self-Pay	210 (41.5)	179 (35.4)	169 (33.5)	187 (37.0)	745 (36.9)
Parity					
Nulliparous	188 (37.2)	211 (41.8)	259 (51.3)	282 (55.8)	940 (46.5)
Primiparous	183 (36.2)	166 (32.9)	145 (28.7)	133 (26.3)	627 (31.0)
Multiparous	135 (26.7)	128 (25.3)	101 (20.0)	90 (17.8)	454 (22.5)
Pre-pregnancy BMI Category					
Underweight	15 (3.0)	21 (4.2)	21 (4.2)	7 (1.4)	64 (3.2)
Normal Weight	244 (48.2)	225 (44.6)	232 (45.9)	259 (51.3)	960 (47.5)
Overweight	130 (25.7)	130 (25.7)	138 (27.3)	119 (23.6)	517 (25.6)
Obese	117 (23.1)	129 (25.5)	114 (22.6)	120 (23.8)	480 (23.8)
Pre-existing Diabetes					
Yes	14 (2.8)	9 (1.8)	14 (2.8)	20 (4.0)	57 (2.8)
No	492 (97.2)	496 (98.2)	491 (97.2)	485 (96.0)	1964 (97.2)
Pre-existing Hypertension					
Yes	29 (5.7)	32 (6.3)	21 (4.2)	31 (6.1)	113 (5.6)

No	477 (94.3)	473 (93.7)	484 (95.8)	474 (93.9)	1908 (94.4)
Pre-existing Polycystic Ovarian Syndrome					
Yes	15 (3.0)	15 (3.0)	14 (2.8)	15 (3.0)	59 (2.9)
No	491 (97.0)	490 (97.0)	491 (97.2)	490 (97.0)	1962 (97.1)
Ever Smoker					
Yes	90 (17.8)	70 (13.9)	55 (10.9)	70 (13.9)	285 (14.1)
No	416 (82.2)	435 (86.1)	450 (89.1)	435 (86.1)	1736 (85.9)
Interpregnancy Interval					
<18 months	188 (37.2)	211 (41.8)	259 (51.3)	282 (55.8)	940 (46.5)
>=18 months	60 (11.9)	55 (10.9)	47 (9.3)	31 (6.1)	193 (9.5)
Nulliparous	130 (25.7)	130 (25.7)	91 (18.0)	92 (18.2)	443 (21.9)
missing	128 (25.3)	109 (21.6)	108 (21.4)	100 (19.8)	445 (22.0)
Assisted Reproductive Technology					
Yes	140 (27.7)	139 (27.5)	176 (34.9)	148 (29.3)	603 (29.8)
No	366 (72.3)	366 (72.5)	329 (65.1)	357 (70.7)	1418 (70.2)
Number of Unique Weight Measurements	12.0 (9.0;14.0)	12.0 (10.0;15.0)	12.0 (10.0;14.0)	12.0 (10.0;14.0)	12.0 (10.0;14.0)
Gestational Age at Birth (weeks)	36.4 (33.9;38.0)	36.6 (34.0;38.0)	36.3 (34.0;37.7)	36.0 (34.0;37.3)	36.3 (34.0;37.9)
Ultrasound					
Yes	506 (100.0)	504 (99.8)	505 (100.0)	505 (100.0)	2020 (100.0)
No	0 (0.0)	1 (0.2)	0 (0.0)	0 (0.0)	1 (0.0)
Gestational Age at First Ultrasound (weeks)	12.3 (8.4;18.0)	11.7 (8.0;18.1)	11.9 (8.0;18.2)	12.3 (7.9;18.4)	12.0 (8.1;18.3)
missing	24 (4.7)	26 (5.1)	24 (4.8)	25 (5.0)	99 (4.9)
Type of Delivery					
Spontaneous	314 (62.1)	278 (55.0)	285 (56.4)	272 (53.9)	1149 (56.9)
Induced	77 (15.2)	94 (18.6)	69 (13.7)	93 (18.4)	333 (16.5)
No Labor	115 (22.7)	133 (26.3)	151 (29.9)	140 (27.7)	539 (26.7)
Delivery Route (Twin A)					
Vaginal	248 (49.0)	234 (46.3)	219 (43.4)	207 (41.0)	908 (44.9)
Caesarean Section	258 (51.0)	271 (53.7)	286 (56.6)	298 (59.0)	1113 (55.1)
Delivery Route (Twin B)					
Vaginal	218 (43.1)	218 (43.2)	197 (39.0)	186 (36.8)	819 (40.5)
Caesarean Section	288 (56.9)	287 (56.8)	308 (61.0)	319 (63.2)	1202 (59.5)
Stillbirth (Twin A)					
Yes	6 (1.2)	2 (0.4)	0 (0.0)	1 (0.2)	9 (0.4)
No	500 (98.8)	503 (99.6)	505 (100.0)	504 (99.8)	2012 (99.6)
Stillbirth (Twin B)					

Yes	7 (1.4)	2 (0.4)	3 (0.6)	4 (0.8)	16 (0.8)
No	499 (98.6)	503 (99.6)	502 (99.4)	501 (99.2)	2005 (99.2)
Total Gestational Weight Gain (kg)	9.5 (5.9;12.7)	15.4 (12.7;17.7)	19.5 (16.8;21.3)	24.9 (22.7;28.1)	16.8 (12.2;21.8)
Total Gestational Weight Gain (z-score)	-1.3 (-1.7; -1.0)	-0.5 (-0.6; -0.3)	0.1 (-0.1; 0.2)	0.8 (0.6; 1.1)	-0.2 (-0.8; 0.4)

[a]*Mean(SE)*

Table 4-2. Sample characteristics by total GWG z-score in singleton pregnancies. Continuous variables presented as median (25th percentile; 75th percentile), categorical variables presented as number (percentage).

Characteristic	**1 (-7.9;-0.8)**	**2 (-0.8;-0.2)**	**3 (-0.2;0.5)**	**4 (0.5;3.9)**	**All**
N	464	464	462	462	1852
Maternal Age (years)[a]	28.4 (6.7)	29.3 (6.4)	29.1 (6.1)	28.6 (6.0)	28.8 (6.3)
Maternal Education					
<High School	58 (12.5)	43 (9.3)	32 (6.9)	34 (7.4)	167 (9.0)
High School/GED	123 (26.5)	96 (20.7)	93 (20.1)	89 (19.3)	401 (21.7)
Some College/Associates	100 (21.6)	85 (18.3)	96 (20.8)	127 (27.5)	408 (22.0)
College Graduate	183 (39.4)	240 (51.7)	241 (52.2)	212 (45.9)	876 (47.3)
Race					
Non-Hispanic White	331 (71.3)	365 (78.7)	379 (82.0)	374 (81.0)	1449 (78.2)
Non-Hispanic Black	98 (21.1)	62 (13.4)	44 (9.5)	68 (14.7)	272 (14.7)
Hispanic	10 (2.2)	4 (0.9)	2 (0.4)	4 (0.9)	20 (1.1)
Other	25 (5.4)	33 (7.1)	37 (8.0)	16 (3.5)	111 (6.0)
Married					
Yes	267 (57.5)	312 (67.2)	316 (68.4)	282 (61.0)	1177 (63.6)
No	197 (42.5)	152 (32.8)	146 (31.6)	180 (39.0)	675 (36.4)
Insurance					
Private/Other	250 (53.9)	280 (60.3)	299 (64.7)	283 (61.3)	1112 (60.0)
Medicaid/Self-Pay	214 (46.1)	184 (39.7)	163 (35.3)	179 (38.7)	740 (40.0)
Parity					
Nulliparous	186 (40.1)	219 (47.2)	244 (52.8)	268 (58.0)	917 (49.5)
Primiparous	166 (35.8)	165 (35.6)	131 (28.4)	117 (25.3)	579 (31.3)
Multiparous	112 (24.1)	80 (17.2)	87 (18.8)	77 (16.7)	356 (19.2)
Pre-existing Diabetes					
Yes	2 (0.4)	0 (0.0)	5 (1.1)	7 (1.5)	14 (0.8)
No	462 (99.6)	464 (100.0)	457 (98.9)	455 (98.5)	1838 (99.2)
Pre-existing Hypertension					
Yes	2 (0.4)	3 (0.6)	6 (1.3)	2 (0.4)	13 (0.7)

No	462 (99.6)	461 (99.4)	456 (98.7)	460 (99.6)	1839 (99.3)
Ever Smoker					
Yes	99 (21.3)	88 (19.0)	73 (15.8)	73 (15.8)	333 (18.0)
No	365 (78.7)	376 (81.0)	389 (84.2)	389 (84.2)	1519 (82.0)
Number of Unique Weight Measurements	11.0 (9.0;13.0)	12.0 (10.0;14.0)	12.0 (10.0;14.0)	12.0 (10.0;15.0)	12.0 (10.0;14.0)
Gestational Age at Birth (weeks)	38.5 (35.8;39.7)	39.0 (36.4;40.0)	38.9 (36.3;40.0)	38.9 (36.3;39.9)	38.9 (36.3;39.9)
Ultrasound					
Yes	456 (98.3)	455 (98.1)	455 (98.5)	458 (99.1)	1824 (98.5)
No	8 (1.7)	9 (1.9)	7 (1.5)	4 (0.9)	28 (1.5)
Gestational Age at First Ultrasound (weeks)	17.1 (10.3;19.3)	15.9 (10.7;18.9)	16.9 (10.0;19.0)	13.9 (9.3;19.1)	16.0 (10.0;19.0)
missing	29 (6.3)	31 (6.7)	24 (5.2)	25 (5.4)	109 (5.9)
Type of Delivery					
Spontaneous	270 (58.2)	257 (55.4)	269 (58.2)	248 (53.7)	1044 (56.4)
Induced	172 (37.1)	173 (37.3)	167 (36.1)	184 (39.8)	696 (37.6)
No Labor	22 (4.7)	34 (7.3)	26 (5.6)	30 (6.5)	112 (6.0)
Delivery Route					
Vaginal	375 (80.8)	363 (78.2)	350 (75.8)	346 (74.9)	1434 (77.4)
Caesarean Section	89 (19.2)	101 (21.8)	112 (24.2)	116 (25.1)	418 (22.6)
Stillbirth					
Yes	1 (0.2)	1 (0.2)	0 (0.0)	2 (0.4)	4 (0.2)
No	463 (99.8)	463 (99.8)	462 (100.0)	460 (99.6)	1848 (99.8)
Total Gestational Weight Gain (kg)	9.1 (6.8;10.4)	12.7 (11.8;13.6)	15.9 (14.5;16.8)	20.4 (19.1;24.0)	14.1 (10.9;17.7)
Total Gestational Weight Gain (z-score)	-1.4 (-1.8;-1.0)	-0.5 (-0.6;-0.3)	0.1 (0.0;0.3)	1.0 (0.7;1.5)	-0.2 (-0.8;0.5)

[a]*Mean(SE)*

4.8. Figures

Figure 4-1. Twin pregnancy sample selection.

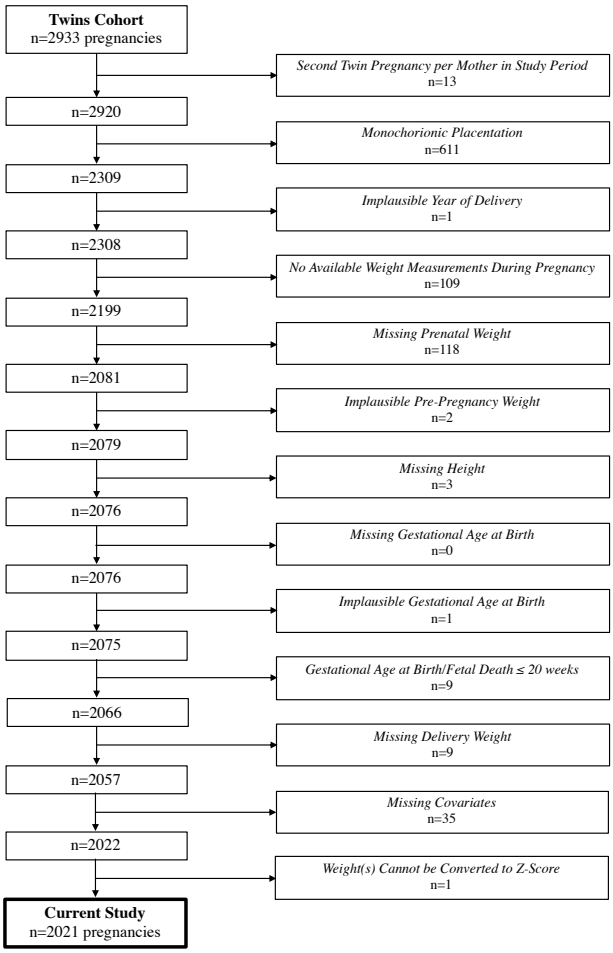


Figure 4-2. Singleton pregnancy sample selection.

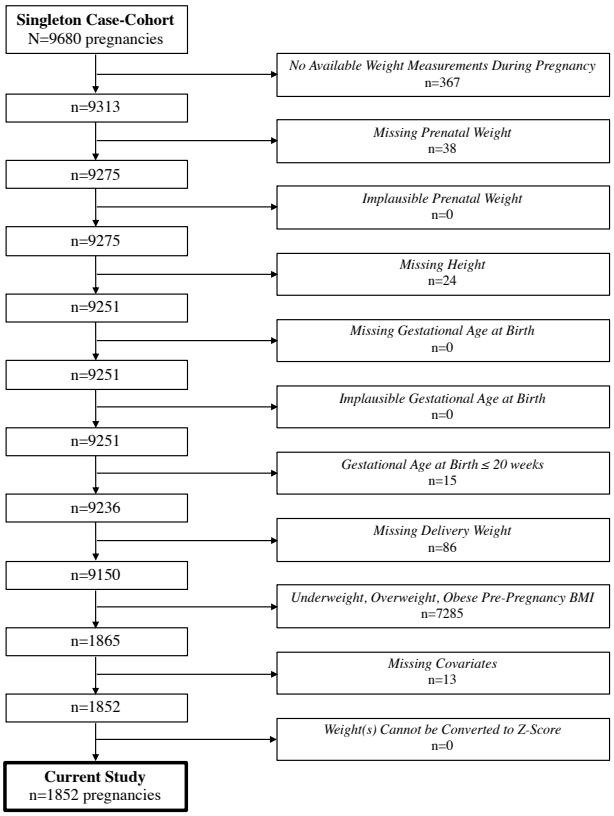


Figure 4-3. Relationship between time-varying GWG z-score and relative hazard of delivery by GA and type of delivery in twin pregnancies.

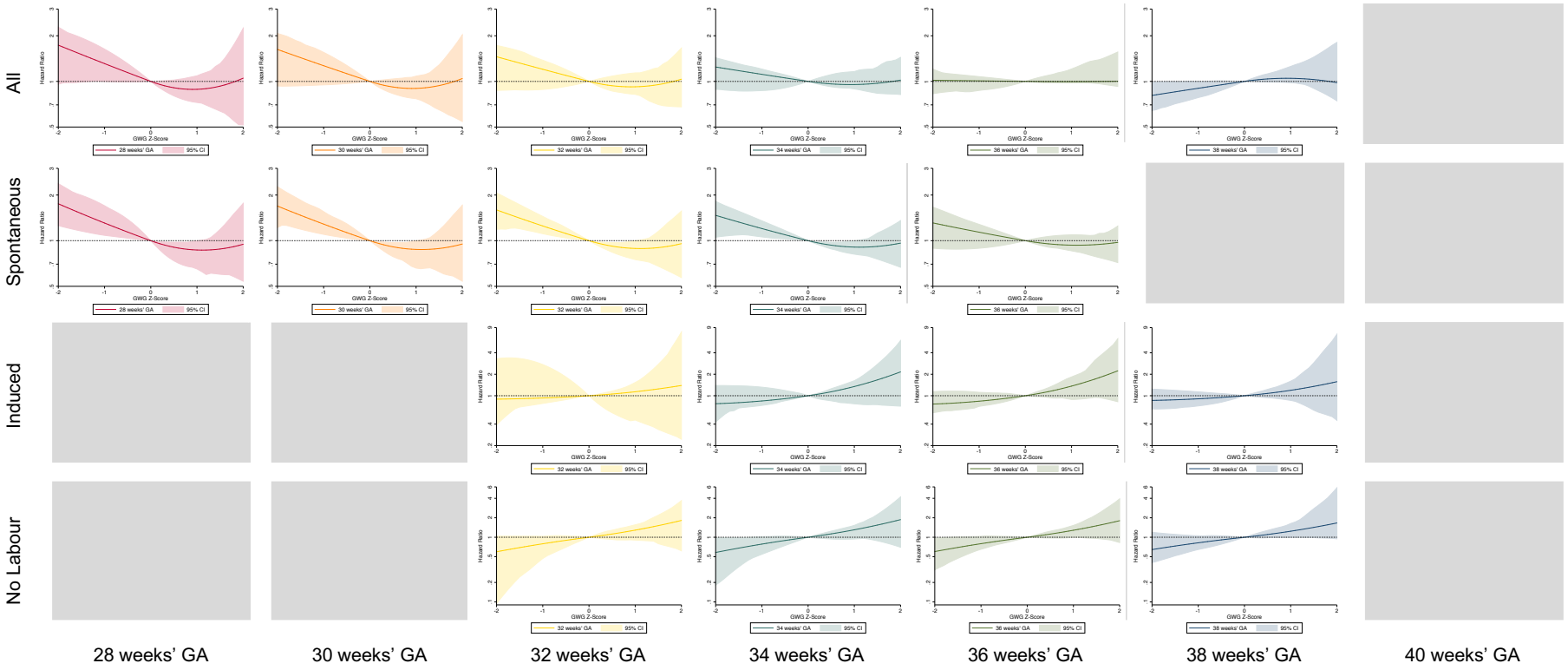
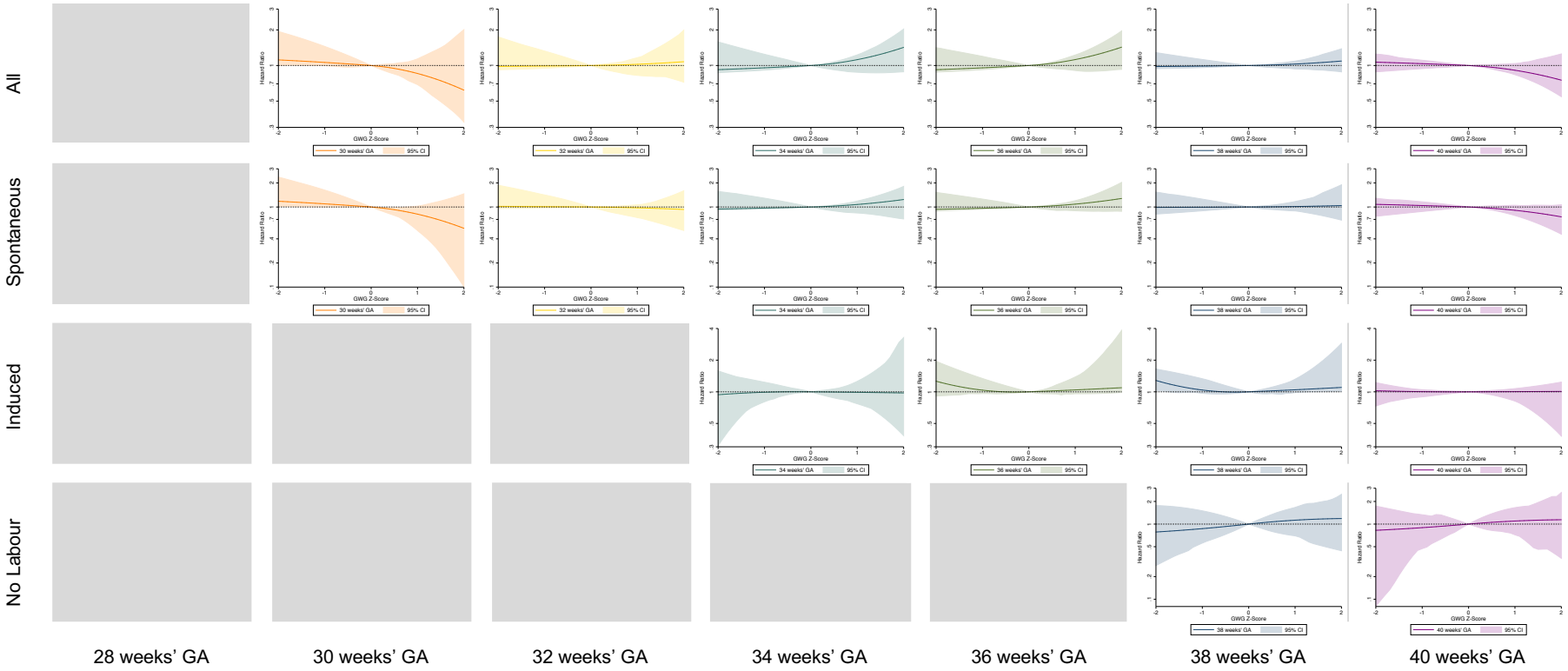


Figure 4-4. Relationship between time-varying GWG z-score and relative hazard of delivery by GA and type of delivery in singleton pregnancies.



4.9. Appendix

Figure A4-1. Model fit statistics by addition/modification of parameters.

Model	Parameters	Deviance	AIC	Change in AIC	Parameters	Deviance	AIC	Change in AIC
Exposure								
GWG (kg)	7	26709.16	26723.16	REF	7	24126.8	24140.8	REF
GWG (z-score)	7	26704.46	26718.46	-4.7	7	24140.08	24154.08	13.28
Covariate = Pre-Pregnancy BMI								
None	7	26704.46	26718.46	REF	7	24140.08	24154.08	REF
Linear/Quadratic	9	26704.46	26722.46	4.00	9	24136.66	24154.66	0.58
Product Method (No TD, No NL)	8	26704.46	26720.46	2.00	8	24136.78	24152.78	-1.30
Product Method (TD, No NL)	11	26700.28	26722.28	3.82	11	24120.76	24142.76	-11.32
Product Method (No TD, NL)	10	26702.44	26722.44	3.98	10	24136.56	24156.56	2.48
Product Method (TD and NL)	14	26695.06	26723.06	4.60	14	24119.26	24147.26	-6.82
Time-Dependent and Non-Linear Effects								
No TD, No NL	1	26732.28	26734.28	REF	1	24166.06	24168.06	REF
TD, No NL	3	26722.94	26728.94	-5.34	4	24136.66	24144.66	-23.40
No TD, NL	4	26708.44	26716.44	-17.84	3	24153.3	24159.3	-8.76
TD and NL	7	26704.46	26718.46	-15.82	7	24140.08	24154.08	-13.98
Splines Knots/Degrees								
1 Knot, 2 Degrees	7	26704.46	26718.46	REF	7	24140.08	24154.08	REF
1 Knot, 3 Degrees	9	26711.14	26729.14	10.68	9	24137.22	24155.22	1.14
2 Knots, 2 Degrees	9	26702.98	26720.98	2.52	9	24135.3	24153.3	-0.78
2 Knots, 3 Degrees	11	26688.32	26710.32	-8.14	11	24137.1	24159.1	5.02
Effect Modification (GWG in kg)								
None	7	26709.16	26723.16	REF	7	24126.8	24140.8	REF

Linear/Quadratic	11	26687.2	26709.2	-13.96	11	24124.54	24146.54	5.74
Product Method (TD and NL)	21	26672.82	26714.82	-8.34	21	24093.44	24135.44	-5.36
Effect Modification (GWG in z-score)								
None	7	26704.46	26718.46	REF	7	24140.08	24154.08	REF
Linear/Quadratic	11	26694.86	26716.86	-1.60	11	24135.32	24157.32	3.24
Product Method (TD and NL)	21	26669.32	26711.32	-7.14	21	24098.14	24140.14	-13.94
Covariate = Confounders								
None	7	26704.46	26718.46	REF	7	24140.08	24154.08	REF
Including All Confounders	24	26621.1	26669.1	-49.36	22	24022.86	24066.86	-87.22

Connecting Manuscripts 2 and 3

In Manuscript 2, I examine the relationship between gestational weight gain and gestational age at birth in a time-to-event framework for both twin and singleton pregnancies. I find similarities in the shape of the relationship between gestational weight gain and spontaneous delivery by plurality, and differences in the relationship between gestational weight gain and induced delivery by plurality. In Manuscript 3, I combine information on both twin and singleton pregnancies to examine whether increased incidence of gestational diabetes in twin pregnancies can be attributed to increased gestational weight gain. Similar to Manuscript 2, I leverage what is known about the relationship between gestational weight gain and gestational diabetes in singleton pregnancies to investigate the role of this exposure among twins.

5. Manuscript 3

5.1. Title

Increased Risk of Gestational Diabetes in Twin Pregnancies is Not Primarily Mediated by Gestational Weight Gain.

5.2. Abstract

Gestational diabetes mellitus, or glucose intolerance during pregnancy is up to three times more common in twin versus singleton pregnancies. The reason for this increased risk is unclear. Although gestational weight gain is a known modifiable cause of gestational diabetes, and gestational weight gain is higher among twins, the extent to which increased gestational weight gain explains the relationship between plurality and gestational diabetes is unknown. We evaluated the extent to which increased risk of gestational diabetes in twin pregnancies is mediated by increased gestational weight gain. We leveraged previously-collected serial weights and glucose screening/diagnostic data abstracted from medical charts for 1397 twin and 2622 singleton pregnancies with normal or overweight pre-pregnancy body mass index delivered between 1998 and 2013 at Magee Women's Hospital in Pittsburgh, Pennsylvania. We used causal mediation analyses to estimate natural indirect and direct effects, or those mediated and not mediated by gestational weight gain, respectively. We found that although odds of gestational diabetes were higher among twin pregnancies [marginal total effect = 2.83 (95% CI 1.55; 5.17) for normal weight and 2.10 (95% CI 1.17; 3.75 for overweight pre-pregnancy body mass index, there is limited evidence that this relationship is mediated by GWG [natural indirect effect = 1.21 (95% CI 0.95; 1.54) for normal weight and 1.06 (95% CI 0.90; 1.24 for overweight pre-pregnancy body mass index], and more evidence of mediation via other mechanisms [natural direct effect = 2.34 (95% CI 1.25; 4.39) for normal weight and 1.99 (95% CI 1.10; 3.61) for overweight pre-pregnancy body mass index]. We conclude that, while twin pregnancies experience nearly 200% increased risk of gestational diabetes relative to singletons, only approximately 10% of

this is mediated by gestational weight gain. We recommend research that investigates alternative mechanisms for the observed relationship between plurality and gestational diabetes, as well as the role of gestational weight gain in causing gestational diabetes specifically among twins.

5.3. Introduction

Gestational diabetes mellitus (GDM) describes glucose intolerance caused by insufficient insulin and/or excessive blood glucose first diagnosed during pregnancy[28]. GDM is associated with adverse perinatal outcomes in both twin and singleton pregnancies, including preeclampsia, Caesarean delivery, preterm birth, large for gestational age, neonatal hypoglycemia, jaundice, and neonatal intensive care unit admissions, although evidence for adverse effects of GDM in singletons is more conclusive than that in twins[31–33]. Nonetheless, diagnostic criteria are applied across pluralities[92].

Gestational weight gain (GWG) is a potentially modifiable risk factor for GDM. Observational studies in singleton pregnancies have found that excessive maternal weight gain prior to glucose screening, and particularly in the first trimester, is correlated with 40% to 130% increased risk for GDM[54–56]. Furthermore, systematic reviews of randomized controlled trials for the prevention of GDM conclude that diet- and exercise-based interventions decrease both average GWG by 0.89 kg and risk of GDM by 15% to 44%[39,40]. In mothers with overweight/obese pre-pregnancy body mass index (BMI), who are at higher risk for GDM, physical activity interventions decreased GWG by 1.14 kg on average, and risk of GDM by 29%[41]. Taken together, these studies may support a causal relationship between early GWG and risk of GDM, although it is plausible that exercise- and diet-based interventions themselves decrease risk for GDM and that GWG is a surrogate measure of this effect.

Twin pregnancies may experience both higher GWG and 100%-200% increased risk for GDM compared to singletons[18,19,93], yet the role of pregnancy weight gain in the increased risk of GDM in twins is unclear. Studies investigating the relationship between GWG and GDM in twin pregnancies are scarce, and those that exist typically examine total GWG as the exposure of interest[59,61,63,64]. Total

GWG is an inappropriate measure in this context, since GDM is often diagnosed in mid-pregnancy, and diet interventions that limit subsequent weight gain are generally recommended[94]. Investigating total GWG and GDM may introduce bias due to reverse causation, since a woman's total weight gain may be the result of her diagnosis rather than its cause. Given that twin pregnancies experience higher GWG and GDM, it is possible that increased GWG may explain some of the increased risk of GDM compared to singleton pregnancies. However, the role of GWG in conferring additional risk of GDM in twin pregnancies, and how the relationship between GWG and GDM relates to that in singletons, has not been adequately studied.

We aimed to investigate the extent to which any increased risk of GDM in twin compared to singleton pregnancies is explained by higher mid-pregnancy weight gain.

5.4. Methods

Study Population

We analyzed previously-collected data for a cohort of diamniotic twin pregnancies and a subcohort of singleton pregnancies delivered at Magee Women's Hospital in Pittsburgh, Pennsylvania[47,69]. Parent cohorts included all twin pregnancies delivered from 1998-2013 and a stratified random sample of singleton pregnancies within pre-pregnancy BMI groups underweight, normal weight, overweight, obese I, obese II, and obese III delivered from 1998-2011. In the singleton subcohort, pregnancies with underweight, overweight, and obese pre-pregnancy BMI were oversampled relative to normal weight women. Since effect modification of the relationship between GWG and GDM by pre-pregnancy BMI was expected, we restricted analyses to twin and singleton pregnancies of normal weight and overweight women.

Data Sources

Maternal and pregnancy characteristics, including self-reported pre-pregnancy weight, were collected from the hospital-maintained Magee Obstetric Maternal and Infant (MOMI) database

supplemented by linked vital statistics information collected by the state of Pennsylvania. Serial maternal weights as well as glucose screening and tolerance test information were abstracted from medical charts.

Inclusion/Exclusion Criteria

We excluded the second twin/singleton pregnancy per mother, pregnancies resulting in delivery or fetal death prior to 24 weeks' gestational age (GA), as well as all twin pregnancies with monochorionic placentation, due to their high-risk nature[70]. We additionally excluded all pregnancies with no available weight measurements prior to glucose screening/tolerance test (or prior to 26 weeks' GA in absence of any glucose screening/tolerance tests) missing pre-pregnancy weight or height, missing or implausible GA at birth, positive or missing diagnosis of pre-existing diabetes, and missing values for key covariates. We included twin and singleton pregnancies with either normal and overweight pre-pregnancy body mass index (BMI) in the current study, and analyzed pregnancies separately within each stratum, since effect modification of the relationship between GWG and GDM by pre-pregnancy BMI is hypothesized.

Gestational Diabetes

In this study population, gestational diabetes was generally determined by 50g glucose screening and 100g glucose tolerance tests. Glucose screening was generally conducted between 24-32 weeks' GA; if blood glucose is ≥135 mg/dL results were considered abnormal, and the diagnostic glucose tolerance test was conducted. Glucose tolerance test thresholds were ≥95 mg/dL for fasting blood glucose value, and ≥180 mg/dL, ≥155 mg/dL, and ≥140 mg/dL for one-, two-, and three-hour blood glucose values, respectively. According to Carpenter and Coustan criteria, GDM is diagnosed when at least two of four thresholds are met or exceeded[95,96]; we followed these criteria where possible.

GDM was occasionally diagnosed at glucose screening, specifically, when blood glucose was ≥200 mg/dL. Additionally, some pregnancies had inconclusive, incomplete, or missing glucose

screening and tolerance test data. For this reason, and to leverage all available pregnancies, GDM was determined by glucose tolerance test, glucose screening test, and International Classification of Disease (ICD-9) code for physician diagnosis in decreasing priority. Our detailed algorithm used to identify GDM from this information is available in the appendix.

Gestational Weight Gain

GWG was calculated by subtracting the prenatal weight measurement of interest by pre-pregnancy weight in kilograms. For our primary analysis, we used the antenatal weight measured at or most recently before the first glucose screening or tolerance test conducted between 20-32 weeks' GA. By definition, pregnancies with no glucose screening/tolerance tests conducted between 20-32 weeks' GA (i.e. all tests conducted <20 or >32 weeks' GA), or pregnancies with no glucose screening/tolerance tests conducted were omitted from primary analysis. Additionally, glucose screening or tolerance tests conducted before 20 weeks' GA were not considered in GWG measurement selection, although the pregnancies themselves were included if any additional glucose screening/tolerance test(s) occurred within the 20-32 weeks' GA window. Our exposure definition was consistent with previous research on this topic while also including pregnancies with only glucose tolerance tests conducted between 20-32 weeks' GA. This is because pregnancies at high risk of GDM may bypass glucose screening and proceed immediately to glucose tolerance test. GWG was measured in kilograms, and analyses adjusted for precise GA of measurement; this is also consistent with previous studies[67].

Statistical Analysis

We first examined crude relationships between GWG and GDM separately in both twins and singletons using both density plots and logistic regression, where GWG was the exposure and GDM was the outcome. We then analyzed twin and singleton pregnancies together and investigated the extent to which any increased risk of GDM in twin compared to singleton pregnancies is explained by

greater GWG using causal mediation analysis[97]. Briefly, causal mediation analyses assess the extent to which an observed relationship between an exposure and outcome can be attributed to mechanisms that include the mediator of interest. We considered plurality the exposure, GWG the mediator of interest, and GDM the outcome; a causal diagram of our proposed relationships is displayed in Figure 5-3.

Causal mediation analyses estimate both natural and controlled effects and assume: (1) no unmeasured confounding between exposure (plurality) and outcome (GDM) (U_1 in Figure 5-3); (2) no unmeasured confounding between mediator (GWG) and outcome (GDM) (U_2 in Figure 5-3). If exposure-mediator interaction is present, causal mediation models additionally assume: (3) no unmeasured exposure-mediator confounding (U_3 in Figure 5-3); (4) no measured/unmeasured mediator-outcome confounding that is caused by exposure (U_4 in Figure 5-3). In our case, the natural indirect effect (NIE) quantifies relative change in odds of GDM if GWG was set to the value it would take if each pregnancy was twin versus singleton, and plurality was set to singleton for all pregnancies. Similarly, the natural direct effect (NDE) quantifies the relative change in odds of GDM if all pregnancies were twin versus singleton, where GWG is set to the value it would have if plurality were singleton for each individual pregnancy. In contrast, the controlled direct effect (CDE) estimates the relative odds of GDM if all pregnancies were twin versus singleton, where GWG is set to a constant value. If no effect modification of the relationship between GWG and GDM by plurality is assumed, the natural and controlled direct effects are equivalent; however, we allowed for interaction between plurality and GWG in all models, regardless of statistical significance, due to biological plausibility of effect modification by plurality. Specifically, GWG of dichorionic twin pregnancies includes an additional fetus and placenta; thus, equivalent mid-pregnancy GWG in singletons and twins may have different implications for risk of GDM. We plotted CDE by across a range of values to which GWG was set.

We estimated the effect of plurality on GWG using linear regression, and the effect of plurality on GDM while holding GWG constant using logistic regression; thus, all effect estimates are displayed as odds ratios (ORs). Since GDM is relatively rare, ORs can generally be interpreted as relative changes in risk of GDM, although we use OR nomenclature in this manuscript. We adjusted models for covariates that may confound relationships between plurality, GWG, and/or GDM, including parity (continuous; linear/quadratic terms), maternal age (continuous; linear/quadratic terms), delivery year (continuous; linear/quadratic terms), marital status, ever smoker, insurance, and maternal race (Hispanic/other grouped due to small cell size). We adjusted for GA at which GWG was measured in days within all models, including covariate- adjusted and unadjusted models. We planned to control for pre-existing hypertension, however, this was not possible since there were no cases of GDM among pregnancies with normal weight pre-pregnancy BMI and pre-existing hypertension. We separately analyzed pregnancies with normal weight and overweight pre-pregnancy BMI.

Sensitivity Analyses

In primary analyses, we considered both glucose screening and glucose tolerance tests conducted between 20-32 weeks' GA, and calculated GWG using the prenatal weight measurement conducted at or most recently prior to the first test. In sensitivity analyses, we both broadened and restricted these criteria. First, we considered all glucose screening and tolerance tests, regardless of timing, and took the corresponding GWG measurement at or most recently prior to the first glucose screening or tolerance test. If no glucose screening or tolerance tests were conducted, we took the most recent GWG measurement on or prior to 26 weeks' GA. Broadening exposure assessment allowed us to include all pregnancies that were eligible for this study but relied more heavily on ICD-9 codes physician diagnosis of GDM in lieu of glucose screening or tolerance test information. Second, we limited our exposure definition by only considering glucose screening tests conducted between 20-32 weeks' GA, which is a criterion used in prior research[67]; this excluded pregnancies with only glucose

tolerance test information. We took the GWG measurement at or most recently prior to the first glucose screening test conducted between 20-32 GA for this alternative exposure definition.

We conducted additional sensitivity analyses restricting to nulliparous women and omitting pregnancies with inconclusive glucose screening and/or tolerance tests. Since the twin cohort and singleton case-cohort included pregnancies up to 2013 and 2011, respectively, we further restricted analyses to pregnancies that were delivered during years common to both parent studies. Lastly, we fit models that did not incorporate interaction between plurality and GWG.

Statistical analyses were conducted using Stata 14.2[78], in particular the paramed command[98]. Ethical approval for the parent studies was obtained from the University of Pittsburgh and University of British Columbia, and additionally from the McGill University Faculty of Medicine Institutional Review Board for the current study.

5.5. Results

Overall, we included 1397 twins and 2622 singletons that met criteria for either primary or sensitivity analyses within the current study (Figures 5-1 and 5-2). For our primary exposure definition, we excluded an additional 442 twins and 845 singletons, or approximately one third of each sample.

Table 5-1 displays maternal and pregnancy characteristics by both plurality and GWG quartile within plurality. Women with twin pregnancies were older and more frequently classified as non-Hispanic white race, college graduate, married, privately insured, and nulliparous. Twin pregnancies were less likely to be classified as ever smoker or pre-existing hypertension but were more likely to experience both preeclampsia and GDM during pregnancy. Additionally, twin pregnancies had slightly lower GA at glucose screening or tolerance test compared to singletons and higher mean GWG, even before accounting for GA of GWG measurement.

Maternal age was similar across GWG quartiles for both singletons and twins. Maternal education was more frequently college graduate in second/third GWG quartiles for singletons, and additionally the fourth quartile for twins. Percent non-Hispanic white race and nulliparous were higher in upper quartiles, while percent black/Hispanic race and multiparous were higher in lower GWG quartiles. Pre-existing hypertension was more frequent in the first GWG quartile, while preeclampsia and ever smoker were more frequent in the first and fourth GWG quartiles, among both twin and singleton pregnancies. As expected, GWG was higher when the first glucose screening/tolerance test was conducted at a greater GA. GDM was increased in the highest but also the lowest GWG quartiles, likely because higher pre-pregnancy BMI is a risk factor for GDM, and these pregnancies tend to gain less weight.

Crude relationships between GWG and GDM by plurality and pre-pregnancy BMI are displayed in Figures 5-4 and 5-5. Generally, mid-pregnancy GWG was increased among pregnancies with GDM in both twins and singletons and for both normal weight and overweight women. Average mid-pregnancy GWG was 11.4 kg in twins and 9.1 kg in singletons, while overall prevalence of GDM was 5.9% and 3.1%, respectively. When the relationship between GWG and GDM was evaluated separately in twins and singletons, relative odds of GDM increased by factors of 1.03 (95% 0.98; 1.08 CI) per 1-kg increase in GWG among both pluralities in models adjusted for GA of GWG measurement and pre-pregnancy BMI.

Effects estimated by causal mediation analysis are displayed in Table 5-2. In this study population and within normal weight pre-pregnancy BMI stratum, OR for marginal total effect of a twin versus singleton pregnancy on GDM was 2.83 (95% CI 1.55; 5.17). In other words, the odds of GDM diagnosis if all pregnancies were twin are nearly three times the odds of GDM diagnosis if all pregnancies were singleton through mechanisms both mediated and not mediated by GWG.

Both the NIE and NDE are ratios of two odds. The NIE compares the odds of GDM if all pregnancies were singleton, but each pregnancy gained the amount of weight it would have gained if it were twin, versus the odds of GDM if all pregnancies were singleton, and each pregnancy gained the amount it would have gained if it were singleton; the OR for the NIE was 1.21 (95% CI 0.95; 1.54). The NDE compares the odds of GDM if each pregnancy were twin but gained the amount of weight it would have gained if it were singleton versus the odds of GDM if each pregnancy were singleton and gained the amount of weight it would have gained if it were singleton; the OR for the NDE was 2.34 (95% CI 1.25; 4.39). Since we allowed interaction between plurality and GWG, the CDE changes depending on the value at which GWG is set and therefore cannot be expressed as a single estimate (Figures 5-6 and 5-7). Magnitudes of coefficients for effect modification were small and the CDE was similar across a range of GWG; thus, CDE can be approximated by the NDE.

Among women with overweight pre-pregnancy BMI, the odds of GDM were 2.10 (95% CI 1.17; 3.75)-fold higher in twin versus singleton pregnancies. The OR for the NIE was 1.06 (95% CI 0.90; 1.24), which accounted for an even smaller proportion of the relative increase in odds of GDM among twins versus singletons than that of normal weight pre-pregnancy BMI women.

Effect estimates were similar throughout sensitivity analyses among women with both normal weight and overweight pre-pregnancy BMI, although confidence intervals widened as sample size decreased. In analyses that used the broader definition of exposure, the marginal total effect decreased from 2.83 to 2.36; however, the NIE was similar in magnitude to that estimated in primary analysis. In analyses that examined blood glucose as a continuous outcome, blood glucose ranged from 7.29 to 8.57 mg/dL higher in twin versus singleton pregnancies among women with overweight and normal weight pre-pregnancy BMI, respectively. Only 18% of this increase in normal-weight pre-pregnancy BMI women and 2% of the increase in overweight pre-pregnancy BMI women could be attributed to mechanisms involving GWG.

5.6. Discussion

Overall, our findings suggest a limited role of GWG in mediating the relationship between plurality and GDM. Odds of GDM increased almost 200% in twins versus singletons pregnancies, and only about 10% of this increase can be attributed to increased GWG among twin pregnancies. This suggests that the relationship between plurality and GDM primarily mediated by causal pathways that do not involve GWG.

Our effects are similar to those reported in prior observational research but suggest a lesser effect of GWG on GDM than that estimated in experimental studies. One observational study estimated up to two-fold, while our study suggests two to three-fold, increased risk of GDM in twin versus singleton pregnancies; however, it is notable that this observational study controlled for GWG when quantifying the relationship between plurality and GDM[19]. Thus, their observed effect may be more analogous to our NDE, which was of a similar magnitude. The marginal total effect of plurality on GDM was greater in normal weight versus overweight women, which suggests effect measure modification by pre-pregnancy BMI. It is important to highlight that effect measure modification is scale-dependent[99]; since pre-pregnancy BMI is also associated with GDM[100], baseline risk for GDM may be higher among overweight women with both twin and singleton pregnancies. Relative odds of GDM associated with increases in GWG may be smaller due to this increased baseline risk. Our marginal total effects approach that of a more recent observational study, which estimated a 3.5-fold increase in GDM among twin pregnancies[93]. A Cochrane review of RCTs aimed at preventing GDM in singletons uncover a 0.89 kg decrease in GWG and corresponding 15% decreased risk of GDM[39]. This is equivalent to a 39% risk reduction in GDM when scaled to the mean difference in GWG among twins versus singletons, which is greater than the 20% increased risk of GDM estimated in our study.

It is plausible that GWG does not mediate a substantial proportion of the relationship between plurality and GDM. We hypothesize that increased risk of GDM in twin pregnancies is mediated by

other physiological mechanisms, such as hormonal, epigenetic, or metabolic changes that occur only within or to a greater extent in multiple pregnancies. Alternatively, unmeasured confounding that either biases the effect between plurality and GDM away from the null and/or biases the effect between GWG and GDM towards the null may affect our inferences.

Our statistical model assumes no unmeasured confounding between, plurality and GDM, GWG and GDM, or plurality and GWG, as well as no measured/unmeasured confounding between GWG and GDM that is caused by plurality (U_1 to U_4 in Figure 5-3). Since we observed minimal interaction between plurality and GWG, the latter two assumptions may be relaxed as they are not required to estimate CDE or NDE when exposure-mediator interaction is not present. An unmeasured confounder, such as a genetic factor, that causes higher likelihood of twin pregnancy and increased risk for GDM may bias the total effect away from the null, while another unmeasured confounder, such as a metabolic factor, that both increases GWG and decreases the risk for GDM in twin pregnancies may bias the NIE towards the null. It has been estimated that up to 46% of GWG may be due to genetic factors[51], but we are unaware of any concurrent effects on either plurality or GDM. Use of assisted reproductive technology was unavailable/incomplete for singletons, and therefore, we were not able to adjust for this covariate; use of assisted reproductive technology increases risk of multiple birth[101] as well as risk of GDM in both twin and singleton pregnancies[102,103]; thus, our marginal total effect estimate may be biased away from the null. Additionally, there are conceivably many unmeasured common causes of GWG and GDM, including socioeconomic factors that impact health during pregnancy, as well as other measures of overall health both during and before pregnancy. Unmeasured common causes of GWG and GDM may bias the NIE and/or NDE either towards or away from the null but have no implications for the observed relationship between plurality and GDM. However, it is notable that crude relationships between GWG and GDM were similar by plurality, despite the diverse sociodemographic profiles observed in twin and singleton pregnancies. GA at which glucose

screening/tolerance test was performed, and therefore GWG at which exposure was measured, was somewhat lower in twins; since twins are at higher risk for GDM, this may constitute a measured confounder of GWG and GDM that is caused by plurality (i.e. U_4 in Figure 4-3). Average difference in GA at GWG measurement was only 0.6 weeks, but we cannot rule out residual confounding by this characteristic, since identifiability conditions under exposure-mediator interactions prohibit both measured and unmeasured mediator-outcome caused by exposure. However, substantial interaction between plurality and GWG was not observed; thus, violation of this assumption may be acceptable in the current study.

Strengths of our study include concurrent analysis of twin and singleton pregnancies from the same population. In the parent studies, twin cohort and singleton case-cohort pregnancies were purposefully selected to represent all twin and singleton pregnancies delivered at Magee Women's Hospital, which enabled us to reliably estimate the effect of plurality on GDM. Parent studies collected detailed data, including serial weight measurements and glucose screening/tolerance test data abstracted from medical charts, for large samples of both twin and singleton pregnancies. We used serial weight gain measurements to calculate GWG at or prior to glucose screening and tolerance tests; this is an important improvement on prior research, which largely examines the relationship between total GWG and GDM. Furthermore, we used causal mediation methods which incorporate exposure-mediator interaction and non-linear effects.

Our study had limitations, including our restriction to normal weight and overweight pre-pregnancy BMI. Effect modification of the relationship between GWG and GDM by pre-pregnancy BMI is difficult to incorporate and interpret in causal mediation analyses. Some studies find evidence for effect modification of the relationship between GWG and GDM by pre-pregnancy BMI[57], while others do not[54]. Since variation in recommended total GWG by pre-pregnancy BMI is well-established[49], we restricted to normal weight/overweight pre-pregnancy BMI to preliminary assess

plausibility of this biological mechanism. Next steps for this field of research include examining these effects across pre-pregnancy BMI strata. Our data were observational, and thus, it is difficult to infer causality; this is particularly true for twin pregnancies, which are typically excluded from RCTs of interventions for GDM[13]. Furthermore, causal inference in observational studies of GWG is difficult because GWG itself is not a well-defined exposure. Specifically, GWG may refer to several different types of change maternal physiology, such as changes in fat/protein mass, total body water, or fetal/placental size[50]. Since several different unmodifiable (genetics) and modifiable (diet/exercise) causes of GWG exist, and changes in GWG caused by these mechanisms may result in different increases/decreases in risk for perinatal outcomes, the consistency assumption needed for causal inference may be violated[52]. Further studies are needed to discern whether interventions that equivalently increase or decrease GWG observe similar effects on maternal and child health among both twins and singletons. To our knowledge, our study is the first to examine mid-pregnancy GWG and its relationship with GDM concurrently in twin and singleton pregnancies.

In conclusion, we find that only approximately 10% of the nearly 200% increased risk of GDM among twins is mediated by GWG in normal weight pre-pregnancy BMI pregnancies. Although GWG is higher on average in twin pregnancies, this did not appear to confer substantially increased risk for GDM, which may in part alleviate concerns about effects of higher GWG on maternal morbidity in twin pregnancies. We recommend future research that focuses on other mechanisms that mediate the relationship between plurality and GDM in addition to GWG. More broadly, we encourage additional research on causal relationships between GWG and maternal morbidity in twin pregnancies and its similarities to that observed in singleton pregnancies to inform future GWG guidelines.

5.7. Tables

Table 5-1. Study characteristics by plurality and singleton/twin-specific quartiles of GWG measured at or before glucose screening test. Continuous variables represented as median (25th percentile; 75th percentile) and categorical variables represented as number (%) unless otherwise stated.

Plurality	**Singletons**					**Twins**					**All**
GWG Quartile (Range)	**1 (-6.35; 5.90)**	**2 (5.90; 8.16)**	**3 (8.16; 11.34)**	**4 (11.79; 25.85)**	**All (-6.35; 25.85)**	**1 (-4.54; 8.16)**	**2 (8.39; 10.89)**	**3 (10.89; 14.06)**	**4 (14.06; 29.48)**	**All (-4.54; 29.48)**	**All (-6.35; 29.48)**
Number	675	644	677	626	2622	359	355	346	337	1397	4019
Maternal Age (years)[a]	28.9 (6.1)	29.7 (6.0)	29.8 (6.0)	29.0 (6.1)	29.4 (6.0)	30.6 (5.9)	31.5 (6.1)	31.3 (5.4)	30.7 (6.0)	31.0 (5.9)	30.0 (6.0)
Maternal Education											
<High School	57 (8.4)	43 (6.7)	41 (6.1)	51 (8.1)	192 (7.3)	16 (4.5)	11 (3.1)	6 (1.7)	20 (5.9)	53 (3.8)	245 (6.1)
High School/GED	154 (22.8)	117 (18.2)	132 (19.5)	148 (23.6)	551 (21.0)	59 (16.4)	55 (15.5)	66 (19.1)	70 (20.8)	250 (17.9)	801 (19.9)
Some College/ Associates	161 (23.9)	122 (18.9)	144 (21.3)	164 (26.2)	591 (22.5)	87 (24.2)	70 (19.7)	61 (17.6)	64 (19.0)	282 (20.2)	873 (21.7)
College Graduate	303 (44.9)	362 (56.2)	360 (53.2)	263 (42.0)	1288 (49.1)	197 (54.9)	219 (61.7)	213 (61.6)	183 (54.3)	812 (58.1)	2100 (52.3)
Maternal Race/ethnicity											
Non-Hispanic White	494 (73.2)	523 (81.2)	551 (81.4)	499 (79.7)	2067 (78.8)	282 (78.6)	297 (83.7)	299 (86.4)	287 (85.2)	1165 (83.4)	3232 (80.4)
Non-Hispanic Black	132 (19.6)	77 (12.0)	91 (13.4)	111 (17.7)	411 (15.7)	53 (14.8)	45 (12.7)	34 (9.8)	43 (12.8)	175 (12.5)	586 (14.6)
Hispanic	12 (1.8)	9 (1.4)	3 (0.4)	4 (0.6)	28 (1.1)	4 (1.1)	2 (0.6)	5 (1.4)	2 (0.6)	13 (0.9)	41 (1.0)

Other	37 (5.5)	35 (5.4)	32 (4.7)	12 (1.9)	116 (4.4)	20 (5.6)	11 (3.1)	8 (2.3)	5 (1.5)	44 (3.1)	160 (4.0)
Married											
Yes	424 (62.8)	475 (73.8)	467 (69.0)	379 (60.5)	1745 (66.6)	261 (72.7)	276 (77.7)	261 (75.4)	241 (71.5)	1039 (74.4)	2784 (69.3)
No	251 (37.2)	169 (26.2)	210 (31.0)	247 (39.5)	877 (33.4)	98 (27.3)	79 (22.3)	85 (24.6)	96 (28.5)	358 (25.6)	1235 (30.7)
Insurance											
Private/ Other	399 (59.1)	415 (64.4)	422 (62.3)	389 (62.1)	1625 (62.0)	229 (63.8)	248 (69.9)	238 (68.8)	218 (64.7)	933 (66.8)	2558 (63.6)
Medicaid/ Self-Pay	276 (40.9)	229 (35.6)	255 (37.7)	237 (37.9)	997 (38.0)	130 (36.2)	107 (30.1)	108 (31.2)	119 (35.3)	464 (33.2)	1461 (36.4)
Parity											
Nulliparous	277 (41.0)	277 (43.0)	296 (43.7)	315 (50.3)	1165 (44.4)	161 (44.8)	166 (46.8)	176 (50.9)	185 (54.9)	688 (49.2)	1853 (46.1)
Primiparous	252 (37.3)	251 (39.0)	242 (35.7)	179 (28.6)	924 (35.2)	120 (33.4)	115 (32.4)	104 (30.1)	87 (25.8)	426 (30.5)	1350 (33.6)
Multiparous	146 (21.6)	116 (18.0)	139 (20.5)	132 (21.1)	533 (20.3)	78 (21.7)	74 (20.8)	66 (19.1)	65 (19.3)	283 (20.3)	816 (20.3)
Pre-Pregnancy BMI											
Normal Weight	327 (48.4)	357 (55.4)	392 (57.9)	284 (45.4)	1360 (51.9)	189 (52.6)	241 (67.9)	249 (72.0)	235 (69.7)	914 (65.4)	2274 (56.6)
Overweight	348 (51.6)	287 (44.6)	285 (42.1)	342 (54.6)	1262 (48.1)	170 (47.4)	114 (32.1)	97 (28.0)	102 (30.3)	483 (34.6)	1745 (43.4)
Pre-existing Hypertension											
Yes	9 (1.3)	14 (2.2)	15 (2.2)	14 (2.2)	52 (2.0)	14 (3.9)	7 (2.0)	9 (2.6)	12 (3.6)	42 (3.0)	94 (2.3)
No	666 (98.7)	630 (97.8)	662 (97.8)	612 (97.8)	2570 (98.0)	345 (96.1)	348 (98.0)	337 (97.4)	325 (96.4)	1355 (97.0)	3925 (97.7)
Preeclampsia											
Yes	20 (3.0)	37 (5.7)	35 (5.2)	43 (6.9)	135 (5.1)	53 (14.8)	63 (17.7)	61 (17.6)	83 (24.6)	260 (18.6)	395 (9.8)
No	655 (97.0)	607 (94.3)	642 (94.8)	583 (93.1)	2487 (94.9)	306 (85.2)	292 (82.3)	285 (82.4)	254 (75.4)	1137 (81.4)	3624 (90.2)

Ever Smoker											
Yes	113 (16.7)	86 (13.4)	95 (14.0)	111 (17.7)	405 (15.4)	47 (13.1)	35 (9.9)	39 (11.3)	47 (13.9)	168 (12.0)	573 (14.3)
No	562 (83.3)	558 (86.6)	582 (86.0)	515 (82.3)	2217 (84.6)	312 (86.9)	320 (90.1)	307 (88.7)	290 (86.1)	1229 (88.0)	3446 (85.7)
Glucose Tolerance/Screening Test Results											
Confirmed No	550 (81.5)	547 (84.9)	588 (86.9)	544 (86.9)	2229 (85.0)	264 (73.5)	276 (77.7)	271 (78.3)	265 (78.6)	1076 (77.0)	3305 (82.2)
Confirmed Yes	15 (2.2)	16 (2.5)	15 (2.2)	19 (3.0)	65 (2.5)	16 (4.5)	17 (4.8)	13 (3.8)	25 (7.4)	71 (5.1)	136 (3.4)
Inconclusive	25 (3.7)	25 (3.9)	23 (3.4)	23 (3.7)	96 (3.7)	18 (5.0)	15 (4.2)	23 (6.6)	21 (6.2)	77 (5.5)	173 (4.3)
missing	85 (12.6)	56 (8.7)	51 (7.5)	40 (6.4)	232 (8.8)	61 (17.0)	47 (13.2)	39 (11.3)	26 (7.7)	173 (12.4)	405 (10.1)
Gestational Diabetes											
Yes	22 (3.3)	19 (3.0)	16 (2.4)	24 (3.8)	81 (3.1)	21 (5.8)	18 (5.1)	17 (4.9)	27 (8.0)	83 (5.9)	164 (4.1)
No	653 (96.7)	625 (97.0)	661 (97.6)	602 (96.2)	2541 (96.9)	338 (94.2)	337 (94.9)	329 (95.1)	310 (92.0)	1314 (94.1)	3855 (95.9)
GA of Glucose Screening (Primary)	26.7 (25.3; 27.9)	26.7 (25.6; 27.9)	27.0 (25.9; 28.1)	27.4 (26.1; 28.6)	27.0 (25.7; 28.1)	26.0 (24.6; 27.3)	26.3 (24.9; 27.6)	26.6 (25.4; 28.0)	26.9 (25.9; 28.1)	26.4 (25.1; 27.7)	26.9 (25.6; 28.0)
missing	256 (37.9)	228 (35.4)	191 (28.2)	170 (27.2)	845 (32.2)	132 (36.8)	116 (32.7)	103 (29.8)	91 (27.0)	442 (31.6)	1287 (32.0)
GA of Glucose Screening (Alternate Version 1)	26.6 (25.0; 27.9)	26.7 (25.6; 27.9)	27.1 (25.7; 28.3)	27.6 (26.1; 28.7)	27.0 (25.6; 28.1)	25.8 (24.2; 27.1)	26.1 (24.6; 27.6)	26.6 (25.3; 28.0)	26.9 (25.6; 28.1)	26.3 (24.9; 27.7)	26.7 (25.4; 28.1)
missing	225 (33.3)	218 (33.9)	178 (26.3)	143 (22.8)	764 (29.1)	119 (33.1)	113 (31.8)	96 (27.7)	81 (24.0)	409 (29.3)	1173 (29.2)
GA of Glucose Screening (Alternate Version 2)	26.7 (25.3; 27.9)	26.7 (25.6; 27.9)	27.0 (25.9; 28.1)	27.4 (26.1; 28.6)	27.0 (25.7; 28.1)	26.0 (24.6; 27.3)	26.3 (24.9; 27.4)	26.6 (25.4; 27.9)	26.9 (25.9; 28.0)	26.4 (25.1; 27.7)	26.7 (25.6; 28.0)

missing	264 (39.1)	236 (36.6)	200 (29.5)	188 (30.0)	888 (33.9)	139 (38.7)	120 (33.8)	107 (30.9)	98 (29.1)	464 (33.2)	1352 (33.6)
GA of GWG Measurement (Primary)	25.3 (23.6; 27.0)	25.8 (24.3; 27.3)	26.1 (24.7; 27.6)	26.9 (25.3; 28.3)	26.1 (24.4; 27.6)	24.6 (22.9; 26.3)	25.1 (23.9; 26.7)	25.9 (24.6; 27.3)	26.4 (25.0; 27.7)	25.6 (24.0; 27.1)	25.9 (24.3; 27.4)
missing	256 (37.9)	228 (35.4)	191 (28.2)	170 (27.2)	845 (32.2)	132 (36.8)	116 (32.7)	103 (29.8)	91 (27.0)	442 (31.6)	1287 (32.0)
GA of GWG Measurement (Alternate Version 1)	24.6 (22.7; 26.1)	25.1 (23.7; 26.3)	25.6 (24.1; 27.1)	26.1 (24.7; 27.9)	25.3 (23.9; 27.0)	24.3 (22.4; 25.6)	24.6 (23.4; 26.0)	25.4 (24.1; 26.6)	25.7 (24.3; 27.3)	25.0 (23.6; 26.3)	25.1 (23.7; 26.9)
missing	0 (0.0)	0 (0.0)	1 (0.1)	0 (0.0)	1 (0.0)	2 (0.6)	1 (0.3)	0 (0.0)	2 (0.6)	5 (0.4)	6 (0.1)
GA of GWG Measurement (Alternate Version 2)	25.3 (23.6; 27.0)	25.7 (24.3; 27.2)	26.1 (24.7; 27.6)	26.9 (25.3; 28.3)	26.1 (24.4; 27.6)	24.6 (22.9; 26.4)	25.0 (23.9; 26.7)	25.9 (24.6; 27.3)	26.4 (25.0; 27.7)	25.6 (24.0; 27.1)	25.9 (24.3; 27.4)
missing	264 (39.1)	236 (36.6)	200 (29.5)	188 (30.0)	888 (33.9)	139 (38.7)	120 (33.8)	107 (30.9)	98 (29.1)	464 (33.2)	1352 (33.6)
GWG (kg) Primary[a]	3.6 (2.2)	7.2 (0.7)	9.9 (0.9)	15.0 (3.0)	9.1 (4.6)	5.6 (2.1)	9.8 (0.7)	12.4 (0.9)	17.3 (2.8)	11.4 (4.6)	9.9 (4.7)
missing	256 (37.9)	228 (35.4)	191 (28.2)	170 (27.2)	845 (32.2)	132 (36.8)	116 (32.7)	103 (29.8)	91 (27.0)	442 (31.6)	1287 (32.0)
GWG (kg) Alternate Version 1[a]	3.5 (2.3)	7.2 (0.8)	9.8 (1.1)	14.8 (3.1)	8.7 (4.6)	5.4 (2.5)	9.6 (1.3)	12.1 (1.7)	16.9 (3.6)	10.9 (4.8)	9.5 (4.8)
missing	0 (0.0)	0 (0.0)	1 (0.1)	0 (0.0)	1 (0.0)	2 (0.6)	1 (0.3)	0 (0.0)	2 (0.6)	5 (0.4)	6 (0.1)
GWG (kg) Alternate Version 2[a]	3.6 (2.2)	7.2 (0.7)	9.9 (0.9)	14.9 (3.0)	9.0 (4.5)	5.6 (2.2)	9.8 (0.7)	12.4 (0.9)	17.3 (2.9)	11.4 (4.6)	9.9 (4.7)
missing	264 (39.1)	236 (36.6)	200 (29.5)	188 (30.0)	888 (33.9)	139 (38.7)	120 (33.8)	107 (30.9)	98 (29.1)	464 (33.2)	1352 (33.6)

[a]*Mean(SE)*

Table 5-2. Natural direct, natural indirect, and total marginal effects estimated by causal mediation analyses. All effect estimates are on the odds ratio scale unless otherwise specified.

	Normal Weight Pre-Pregnancy BMI							Overweight Pre-Pregnancy BMI						
Model	**N**	**NDE**	**p**	**NIE**	**p**	**MTE**	**p**	**N**	**NDE**	**p**	**NIE**	**p**	**MTE**	**p**
Adjusted for GA of Measurement	1565	2.67 (1.46; 4.88)	0.001	1.18 (0.94; 1.48)	0.153	3.14 (1.77; 5.59)	<0.001	1167	2.26 (1.32; 3.88)	0.003	1.06 (0.92; 1.21)	0.409	2.40 (1.42; 4.06)	0.001
Adjusted for GA of Measurement and All Covariates	1565	2.34 (1.25; 4.39)	0.008	1.21 (0.95; 1.54)	0.124	2.83 (1.55; 5.17)	0.001	1167	1.99 (1.10; 3.61)	0.024	1.06 (0.90; 1.24)	0.501	2.10 (1.17; 3.75)	0.012
Sensitivity Analysis, No Exposure-Mediator Interaction	1565	2.37 (1.27; 4.42)	0.007	1.19 (0.97; 1.46)	0.103	2.81 (1.54; 5.12)	0.001	1167	1.99 (1.11; 3.57)	0.022	1.06 (0.95; 1.17)	0.306	2.10 (1.17; 3.75)	0.012
Sensitivity Analysis, Alternate Exposure Version 1	2270	1.91 (1.12; 3.26)	0.018	1.24 (1.01; 1.52)	0.043	2.36 (1.41; 3.94)	0.001	1743	1.53 (0.94; 2.49)	0.089	1.01 (0.89; 1.14)	0.898	1.54 (0.95; 2.49)	0.079
Sensitivity Analysis, Alternate Exposure Version 2	1532	2.38 (1.19; 4.75)	0.014	1.22 (0.94; 1.59)	0.140	2.90 (1.49; 5.61)	0.002	1135	2.17 (1.17; 4.04)	0.014	1.08 (0.91; 1.29)	0.386	2.35 (1.28; 4.31)	0.006
Sensitivity Analysis, Includes Only Test-Confirmed GDM	1503	2.77 (1.42; 5.40)	0.003	1.22 (0.96; 1.57)	0.110	3.39 (1.78; 6.45)	<0.001	1127	2.09 (1.11; 3.93)	0.022	1.10 (0.91; 1.32)	0.326	2.30 (1.25; 4.23)	0.008
Sensitivity Analysis, Includes only Nulliparous Pregnancies	780	2.47 (0.99; 6.19)	0.053	1.24 (0.89; 1.73)	0.200	3.07 (1.25; 7.52)	0.014	522	1.95 (0.93; 4.12)	0.078	1.09 (0.87; 1.36)	0.462	2.12 (1.02; 4.42)	0.044
Sensitivity Analysis, Includes only Common Years	1450	2.45 (1.30; 4.63)	0.006	1.22 (0.94; 1.58)	0.128	2.99 (1.63; 5.49)	0.000	110	2.01 (1.10; 3.67)	0.022	1.09 (0.91; 1.30)	0.365	2.19 (1.21; 3.93)	0.009
Sensitivity Analysis, Outcome is Blood Glucose from Glucose Challenge Test[a]	1527	6.98 (3.89; 10.08)	0.000	1.59 (0.18; 3.00)	0.027	8.57 (5.81; 11.34)	0.000	1134	7.19 (3.21; 11.16)	0.000	0.11 (-1.16; 1.37)	0.869	7.29 (3.49; 11.10)	0.000

[a]*Effect estimates per 1-unit increase in blood glucose (mg/dL) for glucose challenge test*

5.8. Figures

Figure 5-1. Sample selection for twin pregnancies.

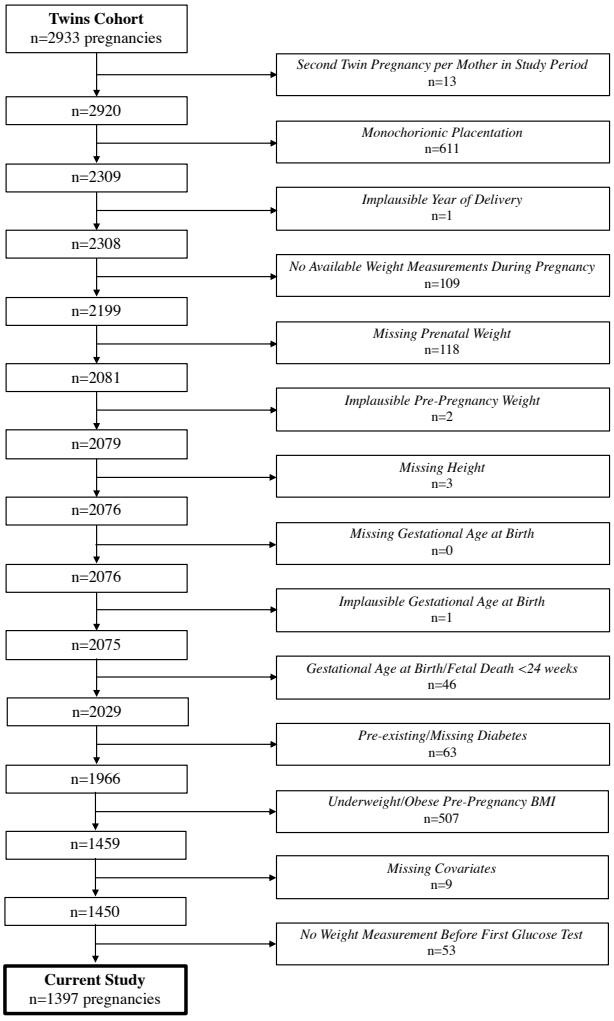


Figure 5-2. Sample selection for singleton pregnancies.

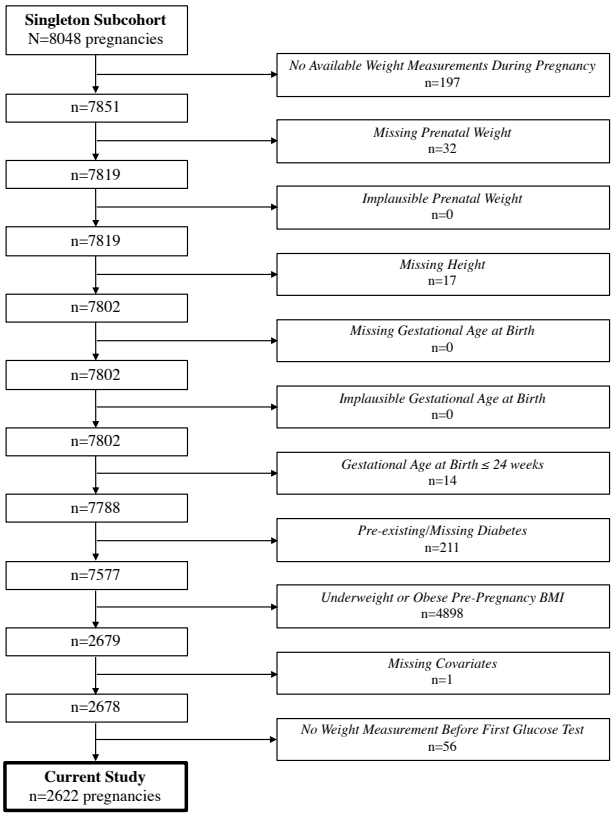


Figure 5-3. Causal diagram for the relationship between plurality, gestational weight gain, and gestational diabetes. Green node/arrows indicate measured confounders, yellow nodes/arrows indicate unmeasured confounders, the presence of which violate identifiability conditions in the presence of exposure-mediator interaction only, and red nodes/arrows indicate unmeasured confounders, the presence of which violate identifiability conditions regardless of exposure-mediator interaction.

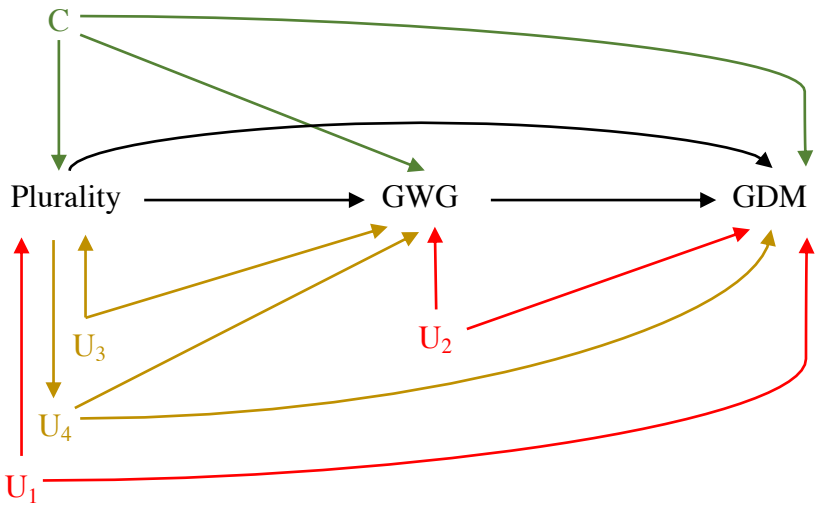


Figure 5-4. Distribution of mid-pregnancy GWG by plurality (twin/singleton) and GDM (yes/no) among women with normal weight pre-pregnancy BMI.

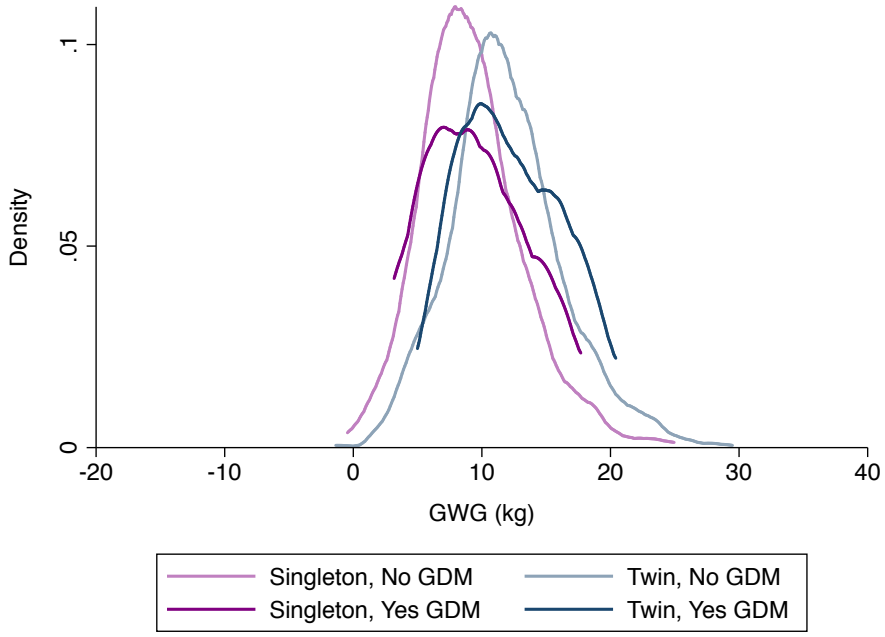


Figure 5.5. Distribution of mid-pregnancy GWG by plurality (twin/singleton) and GDM (yes/no) among women with overweight pre-pregnancy BMI.

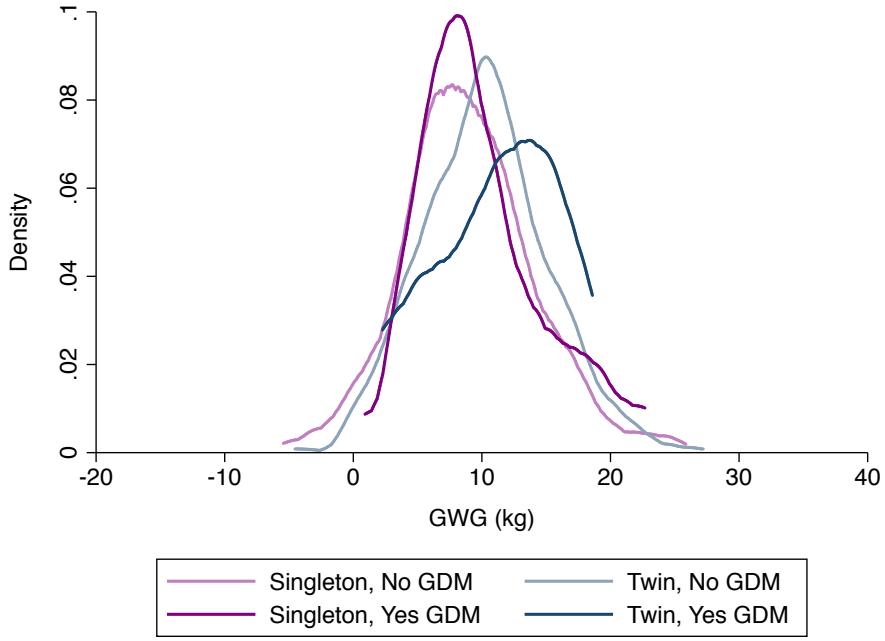


Figure 5-6. Controlled direct effect by value at which gestational weight gain is set for primary model among normal weight pre-pregnancy BMI.

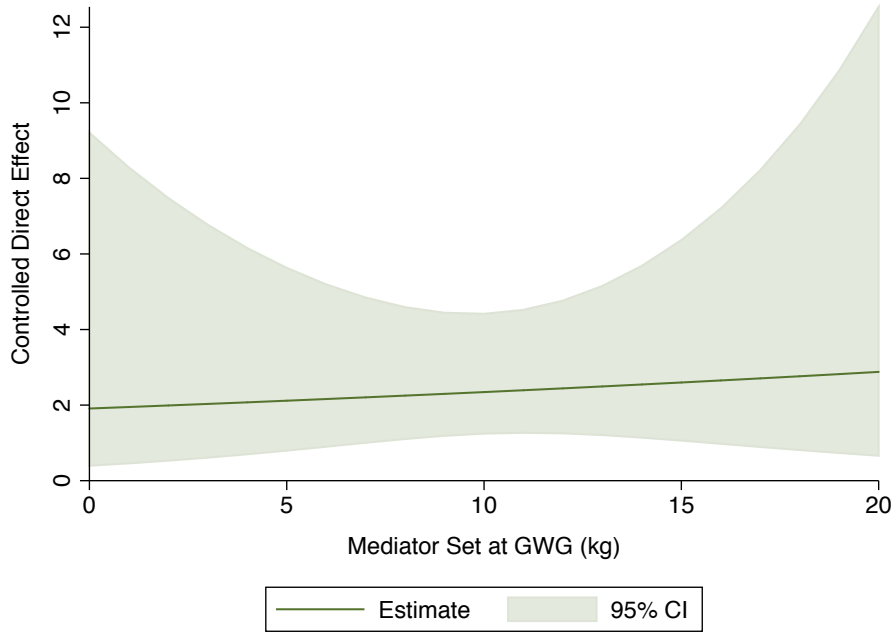
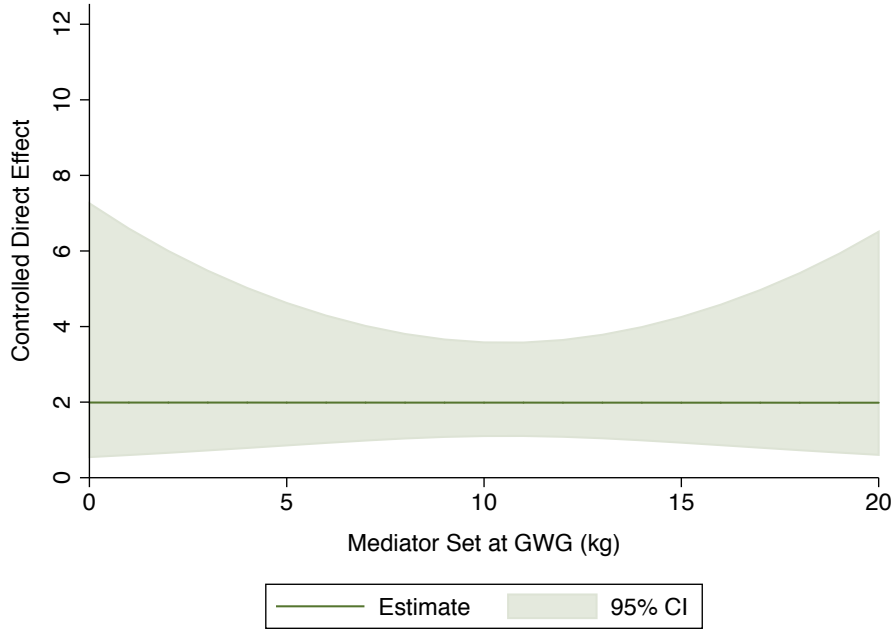


Figure 5-7. Controlled direct effect by value at which gestational weight gain is set for primary model among overweight pre-pregnancy BMI.



5.9. Appendix

Algorithm for identifying GDM in twins and singletons:

1. Glucose tolerance test
    a. Confirmed positive if at least two blood glucose values met/exceeded thresholds OR if test was marked "abnormal"; glucose tolerance test values take precedence over abnormal indicator.
    b. Confirmed negative if at least three blood glucose values were below thresholds OR if test was marked "normal"; glucose tolerance test values take precedence over normal indicator.
    c. Inconclusive if available blood glucose values were not confirmed positive/confirmed negative.
    d. Missing if no glucose tolerance test information was available.
    e. If there were two glucose tolerance tests with conflicting results, the first between 20-32 weeks' GA was taken; if the first was inconclusive, the second was taken. If order of two conflicting glucose tolerance tests was uncertain due to missing GA, glucose tolerance test was considered inconclusive.
2. Glucose screening test
    a. Confirmed positive if blood glucose ≥200 mg/dL.
    b. Confirmed negative if blood glucose <135 mg/dL OR if test was marked normal; glucose screening value takes precedence over normal indicator.
    c. Inconclusive if blood glucose ≥135 mg/dL and blood glucose <200 mg/dL OR if test was marked "abnormal" with no associated blood glucose value; this is because "abnormal" could refer to a value that exceeds 135 mg/dL but does not meet the 200 mg/dL threshold for automatic diagnosis in this population.

d. Missing if no glucose screening test information.

e. If there were two glucose screening tests with conflicting results, the first between 20-32 weeks' GA was taken; if the first was inconclusive, the second was taken.

3. International Classification of Disease (ICD-9) code for physician diagnosis of GDM

    a. Yes or No.

6. Discussion

Goals of my dissertation were twofold: Primarily, I aimed to conduct a thorough investigation of maternal weight gain during pregnancy and its relationship with perinatal outcomes in twin pregnancies, a population that is severely understudied[58]. Secondarily, I aimed to compare the role of GWG in twins and singletons by both analyzing data from singletons within the same study population, and by drawing on what is already known about GWG in singletons from previous research. Underpinning both goals was my aim to address new questions and common methodological pitfalls of GWG research using novel approaches.

In Manuscript 1, my objective was to evaluate several methods for estimating GWG between measurements in both twins and singletons. I compared methods commonly used in GWG research, such as linear interpolation using all measurements, as well as individual and pooled methods that varied on characteristics of flexibility and parametricity. I found that individual (linear interpolation using most proximal measurements) and pooled methods (models with restricted cubic splines for GA and random effects parameters for intercepts and slopes) that were highly flexible and less parametric performed substantially better than commonly used methods in both twin and singleton pregnancies. My findings were consistent with previous research[27,68], although I found that methods incorporating restricted cubic splines exhibited superior accuracy and precision across GA, while this method was found to be less precise in other studies[68]. I used this method to estimate GWG between measurements, which required estimates of GWG per day of GA across gestational duration.

In Manuscript 2, I used a flexible extension of the Cox proportional hazards model to examine potential non-linear and time-dependent relationships between time-varying GWG z-score and GA at birth. Among twins, I found that the u-shaped relationship between GWG and preterm birth, which has been previously described in singletons[1,38], may be due to two different phenomena, namely increased relative hazard of early spontaneous delivery among pregnancies with low GWG and increased relative

hazard of late preterm delivery without labour among pregnancies with high GWG. Findings regarding spontaneous onset of labour were similar to those in normal weight singletons, but evidence for a u-shaped relationship among induced deliveries was also observed. Thus, there appear to be some similarities and differences between this relationship in twin and singleton pregnancies.

In Manuscript 3, I combined twin and singleton pregnancies in the same model to investigate whether any additional risk of gestational diabetes observed in twin versus singleton pregnancies is caused by increased GWG. I found that crude relationships between GWG and GDM were similar by plurality, and that increased GWG played a relatively small role in conferring additional risk of GDM to twin pregnancies.

Strengths of my dissertation include the analysis of detailed information, including serial weight gain measurements, abstracted from medical charts for two large cohorts of twin and singleton pregnancies within the same study population. My dissertation contained many methodological considerations, particularly around the timing of GWG and the inherent link between GWG and gestational duration. Both the size of the parent studies and the detail of information analyzed for these cohorts, as well as the careful methodological considerations in my dissertation, are unique characteristics in regard to studies of GWG and perinatal outcomes. Moreover, I focused on twin pregnancies, which is a high-risk but neglected population in this field of research[58].

Weaknesses of my dissertation pertain specifically to the identifiability conditions for causal inference. Specifically, GWG is an ill-defined intervention in that it may be modifiable by nutrition or exercise-based interventions, or unmodifiable by genetic profile; changes in GWG caused by these and other factors may not have equal impacts on maternal and infant health outcomes. Causal inference from observational data is often difficult in the context of confounding that can and cannot be measured. However, informing GWG guidelines necessitates a dialogue between observational and intervention-based studies. Observational studies are necessary for proposing ranges of ideal GWG,

while intervention-based studies are required to determine whether pregnancies can meet these guidelines and/or whether this substantially alters maternal and child health. My dissertation, and particularly Manuscript 2, provided evidence to support the upper range of existing GWG guidelines in twin pregnancies, which are currently 17 to 25 kg in normal weight, 14 to 23 kg in overweight, and 11 to 19 kg in obese pre-pregnancy BMI strata[49].

6.1 Conclusion

In my doctoral dissertation, I partially filled the gap in methodologically rigorous research needed to refine provisional guidelines for GWG in twin pregnancies. I compared GWG and its relationship to perinatal outcomes in twin and singleton pregnancies and found both similarities and differences. I recommend that future research focus on the assumptions needed to identify causal effects, and in particular the assumption that changes in GWG produce equivalent effects on perinatal outcomes regardless of intervention, as well as investigate effect modification by pre-pregnancy BMI in more detail. For example, intervention-based studies that aim to modify GWG in accordance with current/future guidelines may additionally examine whether equivalent changes in GWG caused by different interventions yield similar effects on maternal and infant outcomes.

7. References

1. Bodnar LM, Hutcheon JA, Parisi SM, Pugh SJ, Abrams B. Comparison of Gestational Weight Gain z-Scores and Traditional Weight Gain Measures in Relation to Perinatal Outcomes. *Paediatr Perinat Epidemiol*. 2015;29(1):11-21. doi:10.1111/ppe.12168
2. Margerison Zilko CE, Rehkopf D, Abrams B. Association of maternal gestational weight gain with short- and long-term maternal and child health outcomes. *Am J Obstet Gynecol*. 2010;202(6):574.e1-574.e8. doi:10.1016/j.ajog.2009.12.007
3. Hutcheon JA, Bodnar LM, Joseph KS, Abrams B, Simhan HN, Platt RW. The bias in current measures of gestational weight gain. *Paediatr Perinat Epidemiol*. 2012;26(2):109-116. doi:10.1111/j.1365-3016.2011.01254.x
4. Martin JA, Hamilton BE, Osterman MJK, Driscoll AK, Drake P. Births: Final Data for 2016. *Natl Vital Stat Rep*. 2018;67(1):1-55.
5. Blondel B, Kogan MD, Alexander GR, et al. The Impact of the Increasing Number of Multiple Births on the Rates of Preterm Birth and Low Birthweight: An International Study. *Am J Public Health*. 2002;92(8):1323-1330. doi:10.2105/AJPH.92.8.1323
6. Kulkarni AD, Adashi EY, Jamieson DJ, Crawford SB, Sunderam S, Kissin DM. Affordability of Fertility Treatments and Multiple Births in the United States. *Paediatr Perinat Epidemiol*. August 2017. doi:10.1111/ppe.12383
7. Basso O, Nohr E, Christensen K, Olsen J. Risk of Twinning as a Function of Maternal Height and Body Mass Index. *JAMA J Am Med Assoc*. 2004;291(13):1564-1566. doi:10.1001/jama.291.13.1564-c
8. Reddy UM, Branum AM, Klebanoff MA. Relationship of Maternal Body Mass Index and Height to Twinning. *Obstet Gynecol*. 2005;105(3):593-597. doi:10.1097/01.AOG.0000153491.09525.dd

9. Chauhan SP, Scardo JA, Hayes E, Abuhamad AZ, Berghella V. Twins: prevalence, problems, and preterm births. *Am J Obstet Gynecol*. 2010;203(4):305-315. doi:10.1016/j.ajog.2010.04.031

10. Fisher SC, Kim SY, Sharma AJ, Rochat R, Morrow B. Is obesity still increasing among pregnant women? Prepregnancy obesity trends in 20 states, 2003–2009. *Prev Med (Baltim)*. 2013;56(6):372-378. doi:10.1016/j.ypmed.2013.02.015

11. Bortolus R. The epidemiology of multiple births. *Hum Reprod Update*. 1999;5(2):179-187. doi:10.1093/humupd/5.2.179

12. Beemsterboer SN, Homburg R, Gorter NA, Schats R, Hompes PGA, Lambalk CB. The paradox of declining fertility but increasing twinning rates with advancing maternal age. *Hum Reprod*. 2006;21(6):1531-1532. doi:10.1093/humrep/del009

13. Grantz KL, Kawakita T, Lu Y-L, Newman R, Berghella V, Caughey A. SMFM Special Statement: State of the science on multifetal gestations: unique considerations and importance. *Am J Obstet Gynecol*. 2019;221(2):B2-B12. doi:10.1016/j.ajog.2019.04.013

14. Witteveen T, Van Den Akker T, Zwart JJ, Bloemenkamp KW, Van Roosmalen J. Severe acute maternal morbidity in multiple pregnancies: a nationwide cohort study. *Am J Obstet Gynecol*. 2016;214(5):641.e1-641.e10. doi:10.1016/j.ajog.2015.11.003

15. Ananth C V., Chauhan SP. Epidemiology of Twinning in Developed Countries. *Semin Perinatol*. 2012;36(3):156-161. doi:10.1053/j.semperi.2012.02.001

16. Hardin J, Carmichael SL, Selvin S, Lammer EJ, Shaw GM. Increased prevalence of cardiovascular defects among 56,709 California twin pairs. *Am J Med Genet Part A*. 2009;149A(5):877-886. doi:10.1002/ajmg.a.32745

17. Goldenberg RL, Culhane JF, Iams JD, Romero R. Preterm Birth 1: Epidemiology and Causes of Preterm Birth. *Obstet Anesth Dig*. 2009;29(1):6-7. doi:10.1097/01.aoa.0000344666.82463.8d

18. Schwartz DB, Daoud Y, Zazula P, et al. Gestational diabetes mellitus: metabolic and blood

glucose parameters in singleton versus twin pregnancies. *Am J Obstet Gynecol*. 1999;181(4):912-914.

19. Rauh-Hain JA, Rana S, Tamez H, et al. Risk for developing gestational diabetes in women with twin pregnancies. *J Matern Fetal Neonatal Med*. 2009;22(4):293-299. doi:10.1080/14767050802663194
20. Buhling KJ, Henrich W, Starr E, et al. Risk for gestational diabetes and hypertension for women with twin pregnancy compared to singleton pregnancy. *Arch Gynecol Obstet*. 2003;269(1):33-36. doi:10.1007/s00404-003-0483-z
21. Henderson CE, Scarpelli S, LaRosa D, Divon MY. Assessing the risk of gestational diabetes in twin gestation. *J Natl Med Assoc*. 1995;87(10):757-758.
22. Engle WA. Morbidity and Mortality in Late Preterm and Early Term Newborns: A Continuum. *Clin Perinatol*. 2011;38(3):493-516. doi:10.1016/j.clp.2011.06.009
23. Parikh LI, Reddy UM, Männistö T, et al. Neonatal outcomes in early term birth. *Am J Obstet Gynecol*. 2014;211(3):265.e1-265.e11. doi:10.1016/j.ajog.2014.03.021
24. Cheong-See F, Schuit E, Arroyo-Manzano D, et al. Prospective risk of stillbirth and neonatal complications in twin pregnancies: systematic review and meta-analysis. *BMJ*. September 2016:i4353. doi:10.1136/bmj.i4353
25. Smith GCS. Life-table analysis of the risk of perinatal death at term and post term in singleton pregnancies. *Am J Obstet Gynecol*. 2001;184(3):489-496. doi:10.1067/mob.2001.109735
26. Muglu J, Rather H, Arroyo-Manzano D, et al. Risks of stillbirth and neonatal death with advancing gestation at term: A systematic review and meta-analysis of cohort studies of 15 million pregnancies. Smith GC, ed. *PLOS Med*. 2019;16(7):e1002838. doi:10.1371/journal.pmed.1002838
27. Mitchell EM, Hinkle SN, Schisterman EF. It's About Time: A Survival Approach to Gestational

Weight Gain and Preterm Delivery. *Epidemiology*. October 2015:1. doi:10.1097/EDE.0000000000000413

28. Reece EA, Leguizamón G, Wiznitzer A. Gestational diabetes: the need for a common ground. *Lancet*. 2009;373(9677):1789-1797. doi:10.1016/S0140-6736(09)60515-8
29. Hapo Study Cooperative Research Group. Hyperglycemia and Adverse Pregnancy Outcomes. *N Engl J Med*. 2008;358(19):1991-2002. doi:10.1056/NEJMoa0707943
30. Landon MB, Spong CY, Thom E, et al. A Multicenter, Randomized Trial of Treatment for Mild Gestational Diabetes. *N Engl J Med*. 2009;361(14):1339-1348. doi:10.1056/NEJMoa0902430
31. Hiersch L, Berger H, Okby R, et al. Gestational diabetes mellitus is associated with adverse outcomes in twin pregnancies. *Am J Obstet Gynecol*. 2019;220(1):102.e1-102.e8. doi:10.1016/j.ajog.2018.10.027
32. McGrath RT, Hocking SL, Scott ES, Seeho SK, Fulcher GR, Glastras SJ. Outcomes of twin pregnancies complicated by gestational diabetes: a meta-analysis of observational studies. *J Perinatol*. 2017;37(4):360-368. doi:10.1038/jp.2016.254
33. Vani K, Scifres C, Himes K. The Impact of Gestational Diabetes in Twin Pregnancies on Maternal and Neonatal Outcomes. *Am J Obstet Gynecol*. 2018;218(1):S590.
34. Caimari F, Ramos A, Pujol I, García-Patterson A, Adelantado JM, Corcoy R. Gestational diabetes mellitus in women with multiple pregnancies: is the metabolic abnormality milder? *J Matern Neonatal Med*. October 2015:1-5. doi:10.3109/14767058.2015.1090424
35. Gandhi M, Gandhi R, Mack LM, et al. Estimated energy requirements increase across pregnancy in healthy women with dichorionic twins. *Am J Clin Nutr*. 2018;108(4):775-783. doi:10.1093/ajcn/nqy184
36. Fox NS, Gerber RS, Saltzman DH, et al. Glycemic control in twin pregnancies with gestational diabetes: are we improving or worsening outcomes? *J Matern Neonatal Med*. 2016;29(7):1041-

1045. doi:10.3109/14767058.2015.1038517

37. Muktabhant B, Lawrie TA, Lumbiganon P, Laopaiboon M. Diet or exercise, or both, for preventing excessive weight gain in pregnancy. *Cochrane Database Syst Rev*. June 2015. doi:10.1002/14651858.CD007145.pub3

38. Santos S, Voerman E, Amiano P, et al. Impact of maternal body mass index and gestational weight gain on pregnancy complications: an individual participant data meta-analysis of European, North American, and Australian cohorts. *BJOG An Int J Obstet Gynaecol*. March 2019:1471-0528.15661. doi:10.1111/1471-0528.15661

39. Shepherd E, Gomersall JC, Tieu J, Han S, Crowther CA, Middleton P. Combined diet and exercise interventions for preventing gestational diabetes mellitus. *Cochrane Database Syst Rev*. November 2017. doi:10.1002/14651858.CD010443.pub3

40. Bennett CJ, Walker RE, Blumfield ML, et al. Interventions designed to reduce excessive gestational weight gain can reduce the incidence of gestational diabetes mellitus: A systematic review and meta-analysis of randomised controlled trials. *Diabetes Res Clin Pract*. 2018;141:69-79. doi:10.1016/j.diabres.2018.04.010

41. Du M, Ouyang Y, Nie X, Huang Y, Redding SR. Effects of physical exercise during pregnancy on maternal and infant outcomes in overweight and obese pregnant women: A meta-analysis. *Birth*. 2019;46(2):211-221. doi:10.1111/birt.12396

42. Oken E, Kleinman KP, Belfort MB, Hammitt JK, Gillman MW. Associations of Gestational Weight Gain With Short- and Longer-term Maternal and Child Health Outcomes. *Am J Epidemiol*. 2009;170(2):173-180. doi:10.1093/aje/kwp101

43. Bodnar LM, Pugh SJ, Lash TL, et al. Low Gestational Weight Gain and Risk of Adverse Perinatal Outcomes in Obese and Severely Obese Women. *Epidemiology*. 2016;27(6):894-902. doi:10.1097/EDE.0000000000000535

44. Bodnar LM, Siega-Riz AM, Simhan HN, Himes KP, Abrams B. Severe obesity, gestational weight gain, and adverse birth outcomes. *Am J Clin Nutr*. 2010;91(6):1642-1648. doi:10.3945/ajcn.2009.29008

45. Wen L, Liu X, Wang L, et al. Correlation between second trimester weight gain and perinatal outcomes in dichorionic twin pregnancies: The LoTiS cohort study. *Eur J Obstet Gynecol Reprod Biol*. 2019;233:64-69. doi:10.1016/j.ejogrb.2018.12.014

46. Lutsiv O, Hulman A, Woolcott C, et al. Examining the provisional guidelines for weight gain in twin pregnancies: a retrospective cohort study. *BMC Pregnancy Childbirth*. 2017;17(1):330. doi:10.1186/s12884-017-1530-2

47. Hutcheon JA, Platt RW, Abrams B, et al. Pregnancy Weight Gain by Gestational Age in Women with Uncomplicated Dichorionic Twin Pregnancies. *Paediatr Perinat Epidemiol*. 2018;32(2):172-180. doi:10.1111/ppe.12446

48. Hutcheon JA, Platt RW, Abrams B, Himes KP, Simhan HN, Bodnar LM. Pregnancy weight gain charts for obese and overweight women. *Obesity*. 2015;23(3):532-535. doi:10.1002/oby.21011

49. Institute of Medicine (US) and National Research Council (US) Committee to Reexamine IOM Pregnancy Weight Guidelines, Rasmussen K, Yaktine A, Editors. *Weight Gain During Pregnancy: Reexamining the Guidelines*. Vol 33. Washington, DC: The National Academies Press; 2009.

50. Gandhi M, Gandhi R, Mack LM, et al. Impact of changes in maternal body composition on birth weight and neonatal fat mass in dichorionic twin pregnancies. *Am J Clin Nutr*. 2018;108(4):716-721. doi:10.1093/ajcn/nqy180

51. Andersson ES, Silventoinen K, Tynelius P, Nohr EA, Sørensen TIA, Rasmussen F. Heritability of Gestational Weight Gain – A Swedish Register-Based Twin Study. *Twin Res Hum Genet*. 2015;18(4):410-418. doi:10.1017/thg.2015.38

52. Hernán MA, Robins J. *Causal Inference*. Boca Raton: Chapman & Hall/CRC, forthcoming.; 2015.

53. Hutcheon JA, Bodnar LM. A systematic approach for establishing the range of recommended weight gain in pregnancy. *Am J Clin Nutr*. 2014;100(2):701-707. doi:10.3945/ajcn.114.085258

54. Brunner S, Stecher L, Ziebarth S, et al. Excessive gestational weight gain prior to glucose screening and the risk of gestational diabetes: a meta-analysis. *Diabetologia*. 2015;58(10):2229-2237. doi:10.1007/s00125-015-3686-5

55. Hedderson MM, Gunderson EP, Ferrara A. Gestational Weight Gain and Risk of Gestational Diabetes Mellitus. *Obstet Gynecol*. 2010;115(3):597-604. doi:10.1097/AOG.0b013e3181cfce4f

56. Zhong C, Li X, Chen R, et al. Greater early and mid-pregnancy gestational weight gain are associated with increased risk of gestational diabetes mellitus: A prospective cohort study. *Clin Nutr ESPEN*. 2017;22:48-53. doi:10.1016/j.clnesp.2017.08.013

57. Gibson KS, Waters TP, Catalano PM. Maternal Weight Gain in Women Who Develop Gestational Diabetes Mellitus. *Obstet Gynecol*. 2012;119(3):560-565. doi:10.1097/AOG.0b013e31824758e0

58. Bodnar LM, Pugh SJ, Abrams B, Himes KP, Hutcheon JA. Gestational weight gain in twin pregnancies and maternal and child health: a systematic review. *J Perinatol*. 2014;34(4):252-263. doi:10.1038/jp.2013.177

59. Pettit KE, Lacoursiere DY, Schrimmer DB, Alblewi H, Moore TR, Ramos GA. The association of inadequate mid-pregnancy weight gain and preterm birth in twin pregnancies. *J Perinatol*. 2015;35(2):85-89. doi:10.1038/jp.2014.160

60. Pettit KE, Lacoursiere DY, Schrimmer DB, Alblewi H, Moore TR, Ramos GA. Maternal and neonatal outcomes in women with twin pregnancies with excessive gestational weight gain. *J Matern Neonatal Med*. September 2015:1-4. doi:10.3109/14767058.2015.1079613

61. Pécheux O, Garabedian C, Drumez E, et al. Maternal and neonatal outcomes according to gestational weight gain in twin pregnancies: Are the Institute of Medicine guidelines associated with better outcomes? *Eur J Obstet Gynecol Reprod Biol*. 2019;234:190-194. doi:10.1016/j.ejogrb.2019.01.010

62. Greenan C, Newman R, Wojciechowski B, Christensen C. Achievement of Body Mass Index Specific Weight Gain Recommendations: Impact on Preterm Birth in Twin Pregnancies. *Am J Perinatol*. 2017;34(13):1293-1301. doi:10.1055/s-0037-1603318

63. Kosinska-Kaczynska K, Szymusik I, Kaczynski B, Wielgos M. Observational study of associations between gestational weight gain and perinatal outcomes in dichorionic twin pregnancies. *Int J Gynecol Obstet*. 2017;138(1):94-99. doi:10.1002/ijgo.12171

64. Shamshirsaz AA, Haeri S, Ravangard SF, et al. Perinatal outcomes based on the institute of medicine guidelines for weight gain in twin pregnancies. *J Matern Neonatal Med*. 2014;27(6):552-556. doi:10.3109/14767058.2013.836177

65. Yee LM, Caughey AB, Grobman WA, Cheng YW. Gestational weight gain for women with twins: are the IOM guidelines appropriate? *Am J Obstet Gynecol*. 2017;216(1):S437.

66. Ram M, Berger H, Ray JG, et al. Gestational weight gain in twin pregnancies modeled as a function of gestational age and pre-pregnancy body mass index. *Am J Obstet Gynecol*. 2018;218(1):S482.

67. MacDonald SC, Bodnar LM, Himes KP, Hutcheon JA. Patterns of Gestational Weight Gain in Early Pregnancy and Risk of Gestational Diabetes Mellitus. *Epidemiology*. 2017;28(3):419-427. doi:10.1097/EDE.0000000000000629

68. Darling AM, Werler MM, Cantonwine DE, Fawzi WW, McElrath TF. Accuracy of a mixed effects model interpolation technique for the estimation of pregnancy weight values. *J Epidemiol Community Health*. 2019;73(8):786-792. doi:10.1136/jech-2018-211094

69. Bodnar LM, Himes KP, Abrams B, Parisi SM, Hutcheon JA. Early-pregnancy weight gain and the risk of preeclampsia: A case-cohort study. *Pregnancy Hypertens*. 2018;14:205-212. doi:10.1016/j.preghy.2018.10.005

70. Sebire NJ, Snijders RJ, Hughes K, Sepulveda W, Nicolaides KH. The hidden mortality of monochorionic twin pregnancies. *Br J Obstet Gynaecol*. 1997;104(10):1203-1207.

71. American College of Obstetricians and Gynecologists. ACOG Practice Bulletin No. 101: Ultrasonography in Pregnancy. *Obstet Gynecol*. 2009;113(2, Part 1):451-461. doi:10.1097/AOG.0b013e31819930b0

72. Oken E, Baccarelli AA, Gold DR, et al. Cohort Profile: Project Viva. *Int J Epidemiol*. 2015;44(1):37-48. doi:10.1093/ije/dyu008

73. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int J Epidemiol*. 1999;28(5):964-974. doi:10.1093/ije/28.5.964

74. Durrleman S, Simon R. Flexible regression models with cubic splines. *Stat Med*. 1989;8(5):551-561. doi:10.1002/sim.4780080504

75. Zeger S, Liang L. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73:13-22.

76. Yang S, Hutcheon JA. Identifying outliers and implausible values in growth trajectory data. *Ann Epidemiol*. 2016;26(1):77-80.e2. doi:10.1016/j.annepidem.2015.10.002

77. Shi J, Korsiak J, Roth DE. New approach for the identification of implausible values and outliers in longitudinal childhood anthropometric data. *Ann Epidemiol*. 2018;28(3):204-211.e3. doi:10.1016/j.annepidem.2018.01.007

78. StataCorp. Stata Statistical Software: Release 14. 2015.

79. Hutcheon JA, Stephansson O, Cnattingius S, Bodnar LM, Johansson K. Is the Association Between Pregnancy Weight Gain and Fetal Size Causal? *Epidemiology*. 2019;30(2):234-242.

doi:10.1097/EDE.0000000000000959

80. Hutcheon JA, Stephansson O, Cnattingius S, Bodnar LM, Wikström A-K, Johansson K. Pregnancy Weight Gain Before Diagnosis and Risk of Preeclampsia. *Hypertension*. 2018;72(2):433-441. doi:10.1161/HYPERTENSIONAHA.118.10999

81. Headen I, Cohen AK, Mujahid M, Abrams B. The accuracy of self-reported pregnancy-related weight: a systematic review. *Obes Rev*. 2017;18(3):350-369. doi:10.1111/obr.12486

82. Kalish RB, Thaler HT, Chasen ST, et al. First- and second-trimester ultrasound assessment of gestational age. *Am J Obstet Gynecol*. 2004;191(3):975-978. doi:10.1016/j.ajog.2004.06.053

83. Dietz PM, Callaghan WM, Cogswell ME, Morrow B, Ferre C, Schieve LA. Combined Effects of Prepregnancy Body Mass Index and Weight Gain During Pregnancy on the Risk of Preterm Delivery. *Epidemiology*. 2006;17(2):170-177. doi:10.1097/01.ede.0000198470.26932.9a

84. Nohr EA, Bech BH, Vaeth M, Rasmussen KM, Henriksen TB, Olsen J. Obesity, gestational weight gain and preterm birth: a study within the Danish National Birth Cohort. *Paediatr Perinat Epidemiol*. 2007;21(1):5-14. doi:10.1111/j.1365-3016.2007.00762.x

85. Han Z, Lutsiv O, Mulla S, Rosen A, Beyene J, McDonald SD. Low gestational weight gain and the risk of preterm birth and low birthweight: a systematic review and meta-analyses. *Acta Obstet Gynecol Scand*. 2011;90(9):935-954. doi:10.1111/j.1600-0412.2011.01185.x

86. Darling AM, Werler MM, Cantonwine DE, Fawzi WW, McElrath TF. Timing and Amount of Gestational Weight Gain in Association with Adverse Birth Outcomes. *Epidemiology*. 2019;30(5):695-705. doi:10.1097/EDE.0000000000001055

87. Abrahamowicz M, MacKenzie TA. Joint estimation of time-dependent and non-linear effects of continuous covariates on survival. *Stat Med*. 2007;26(2):392-408. doi:10.1002/sim.2519

88. Hutcheon JA, Platt RW, Abrams B, Himes KP, Simhan HN, Bodnar LM. A weight-gain-for-gestational-age z score chart for the assessment of maternal weight gain in pregnancy. *Am J Clin*

Nutr. 2013;97(5):1062-1067. doi:10.3945/ajcn.112.051706

89. Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr*. 1974;19(6):716-723. doi:10.1109/TAC.1974.1100705

90. Quantin C, Abrahamowicz M, Moreau T, et al. Variation Over Time of the Effects of Prognostic Factors in a Population-based Study of Colon Cancer: Comparison of Statistical Models. *Am J Epidemiol*. 1999;150(11):1188-1200. doi:10.1093/oxfordjournals.aje.a009945

91. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2018.

92. Ooi S, Wong VW. Twin Pregnancy With Gestational Diabetes Mellitus: A Double Whammy? *Diabetes Care*. 2018;41(2):e15-e16. doi:10.2337/dc17-2227

93. Weissman A, Drugan A. Glucose tolerance in singleton, twin and triplet pregnancies. *J Perinat Med*. 2016;44(8). doi:10.1515/jpm-2016-0186

94. Brown J, Alwan NA, West J, et al. Lifestyle interventions for the treatment of women with gestational diabetes. *Cochrane Database Syst Rev*. May 2017. doi:10.1002/14651858.CD011970.pub2

95. Carpenter M, Coustan D. Criteria for screening tests for gestational diabetes. *Am J Obstet Gynecol*. 1982;144(7):768-773.

96. Metzger BE, Coustan DR. Summary and recommendations of the Fourth International Workshop-Conference on Gestational Diabetes Mellitus. The Organizing Committee. *Diabetes Care*. 1998;21 Suppl 2:B161-7.

97. VanderWeele TJ. *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press; 2015.

98. Emsley R, Liu H. PARAMED: Stata module to perform causal mediation analysis using parametric regression models. 2013.

99. VanderWeele TJ, Knol MJ. A Tutorial on Interaction. *Epidemiol Method*. 2014;3(1). doi:10.1515/em-2013-0005

100. Kim SY, England L, Wilson HG, Bish C, Satten GA, Dietz P. Percentage of gestational diabetes mellitus attributable to overweight and obesity. *Am J Public Health*. 2010;100(6):1047-1052. doi:10.2105/AJPH.2009.172890

101. Lynch A. Assisted reproductive interventions and multiple birth. *Obstet Gynecol*. 2001;97(2):195-200. doi:10.1016/S0029-7844(00)01145-5

102. Qin J, Wang H, Sheng X, Liang D, Tan H, Xia J. Pregnancy-related complications and adverse pregnancy outcomes in multiple pregnancies resulting from assisted reproductive technology: a meta-analysis of cohort studies. *Fertil Steril*. 2015;103(6):1492-1508.e7. doi:10.1016/j.fertnstert.2015.03.018

103. Qin J, Liu X, Sheng X, Wang H, Gao S. Assisted reproductive technology and the risk of pregnancy-related complications and adverse pregnancy outcomes in singleton pregnancies: a meta-analysis of cohort studies. *Fertil Steril*. 2016;105(1):73-85.e6. doi:10.1016/j.fertnstert.2015.09.007