COMPARATIVE EFFECTIVENESS AND SAFETY OF METFORMIN VERSUS INSULIN FOR GESTATIONAL DIABETES TREATMENT: MATERNAL, NEONATAL, AND CHILD HEALTH OUTCOMES

Suzanne Noelle Landi

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Epidemiology in the Gillings School of Global Public Health.

Chapel Hill
2018

Approved by:

Michele Jonsson Funk

Kim Boggess

Stephanie Engel

Sarah Radke

Til Stürmer

ABSTRACT

Suzanne Noelle Landi: Comparative Effectiveness and Safety of Metformin Versus Insulin
for Gestational Diabetes Treatment: Maternal, Neonatal, and Child Health Outcomes
(Under the direction of Michele Jonsson Funk)

Background: Metformin is an emerging option for treating gestational diabetes
(GDM). The evidence for the comparative safety and effectiveness of metformin versus
insulin (standard of care) is inconclusive.

Objective: To measure the comparative safety and effectiveness of metformin versus
insulin for pharmacological management of GDM.

Methods: We conducted a cohort study using New Zealand healthcare data (2005-
2015). We identified pregnant women treated with metformin or insulin for GDM and
excluded women with evidence of preexisting diabetes. We assessed birth outcomes using
maternity care and hospitalization data. We examined child growth and psychosocial
development as measured by parent- and teacher-reported Strengths and Difficulties
Questionnaires (SDQ), recorded during pre-school health assessments at age 4. We
adjusted for measured covariates using inverse probability of treatment weights and used
multiple imputation for missing covariates. For dichotomous outcomes, we estimated risk
differences (RD) per 100 and risk ratios (RR) with 95% confidence intervals (CI) and used
linear regression for continuous measures.

Results: We identified 7,268 pregnancies treated initially with metformin (n=3,818) or
insulin (n=3,450). Treatment initiation varied by ethnicity, socioeconomic status, region and
calendar year, though both groups were similar with respect to age, body mass index, and
timing of diagnosis and treatment initiation. After adjustment, metformin was associated with

a reduced risk of large-for-gestational-age (RD= -3.7, 95% CI -5.5, -1.8), and neonatal

hypoglycemia (RD= -5.0, 95% CI -6.9, -3.2), compared to insulin. After adjustment, children

of metformin-treated mothers were not more likely to be ≥85th percentile for weight

(RR=1.04, 95% CI 0.93 1.16) or weight-for-height (RR=0.92, 95% CI 0.83, 1.02) than

children of insulin-treated mothers. Children of metformin-treated mothers were not at

significantly increased risk of having a concerning parent-reported SDQ score (RR=1.13,

95% CI 0.88, 1.46) compared to insulin, although the upper confidence limit does not

exclude a modest increase in risk.

Conclusions: Metformin appears to be an effective treatment for women with GDM

and may reduce the risk of some adverse neonatal outcomes when compared to insulin.

Growth assessments of children of metformin-treated mothers were similar to those of

children of insulin-treated mothers. These results will help inform future GDM treatment

guidelines.

To my mother, Deborah, to my grandmother, Elizabeth and to my grandfather, Robert for their lifelong encouragement of my education.

To my husband, Aaron and son, Graham for their love and support.

Department of Epidemiology including Nancy Colvin, Carmen Woody, Valerie Hudock, Jennifer Moore, and Spencer Gee. I would like to extend a special thank you to Virginia Pate for her patience and her willingness to develop my programming skills. Thank you to all the friends I have made along the way, most notably those in my cohort and those I met through Michele's advisee group, with special thanks to Mitchell Conover for his wisdom, help, and support.

I would like to thank my friends and family. I am grateful for my family's lifelong support of my academic pursuits, especially my mother, Deborah. I am thankful to my friends who supported me during this journey, especially my dearest friend Jennay Ghowrwal and my caring Durham neighbors.

Finally, I could not have made it to the end of this journey without the love and support of my husband, Aaron. I am humbled to receive his encouragement, loyalty, and patience. I thank Aaron for always seeing the best in me. I would also like to thank Aaron for being an amazing, giving father to our son, Graham. And I thank Graham for always being my bright spot in any darkness. I love you both and hope I have made you proud.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACOG   American Congress of Obstetricians and Gynecologists

ADA   American Diabetes Association

B4SC   B4 School Check Screening Programme

BMI   Body mass index

CI   Confidence interval

GDM   Gestational diabetes mellitus

ICD-9-CM   International Classification of Diseases, Ninth Revision, Clinical Modification

ICD-10-AM   International Classification of Diseases, Tenth Revision, Australian Modification

IPTW   Inverse probability of treatment weights

IQR   Interquartile range

LGA   Large for gestational age

LMC   Lead Maternity Carer

LMP   Last menstrual period

MAT   Maternity Collection

MiG   Metformin in Gestational Diabetes Trial

mmol/L   millimoles per liter

NHI   National Health Identifier

NICU   Neonatal intensive care unit

NMDS   National Minimum Dataset

NZ   New Zealand

NZDep13   New Zealand Deprivation Index, 2013

OHA   Oral hypoglycemic agent

OR   Odds ratio

PCOS   Polycystic ovary syndrome

PHARMAC   Pharmaceutical Management Agency

PHARMS   Pharmaceutical Claims Data Mart

PRAMS   Pregnancy Risk Assessment Monitoring System

RCT   Randomized controlled trial

RD   Risk difference

RR   Risk ratio

SD   Standard deviation

SDQ   Strengths and Difficulties Questionnaire

SDQ-P   Strengths and Difficulties Questionnaire, Parent-Reported

SDQ-T   Strengths and Difficulties Questionnaire, Teacher-Reported

UK   United Kingdom

US   United States

WHO   World Health Organization

CHAPTER 1 – REVIEW OF THE LITERATURE

1.1 Background and Significance

Over the course of a normal human pregnancy, a series of metabolic changes occur that promote adipose tissue growth in early gestation, accompanied by insulin resistance later on. In early pregnancy, insulin sensitivity is relatively unchanged or slightly increased, and decreases over time into mid and late pregnancy. As this occurs, insulin secretion from the pancreas increases to compensate (1). Insulin is produced by the body to help deliver glucose to the cells, and production increases during pregnancy to help transport glucose from the bloodstream to the developing fetus. During the second and third trimesters of pregnancy, placental hormone production may inhibit the action of insulin. When a woman's body is unable to produce enough insulin during this time, glucose is not adequately diverted from the bloodstream and causes her blood glucose levels to increase more than normal (2-4).When this rise in blood glucose levels occurs for the first time during pregnancy, this condition is defined as gestational diabetes mellitus and is distinct from diabetes mellitus (type 2 diabetes).

Worldwide, the incidence of GDM is increasing (2). Estimated prevalence in the United States has varied between 1 and 14 percent, depending on the diagnostic criteria being used (5). The change in prevalence of GDM from 1989-1990 to 2003-2004 in the United States was estimated to be a 122% increase overall (1.9% to 4.2%) (6). A recent examination comparing the Pregnancy Risk Assessment Monitoring System (PRAMS) questionnaire to birth certificate records from 2007 to 2010 revealed that the GDM prevalence may vary between 4.6% and 9.2%, though the 21 states and cities represented in PRAMS may not be representative of the entire United States. In particular, states from

the Southeast, where obesity prevalence and chronic disease burden tends to be higher, were not well-represented in this study (5). In England and Wales, the average prevalence of any diabetes in pregnancy is 3.5%, varying across region and ethnicity, with approximately 87.5% of cases attributed to GDM (7). In Australia, GDM incidence in recorded hospitalizations increased by 21% from 2000-01 to 2009-10 (3.6% to 4.4%) (8). Prevalence of GDM also increased in New Zealand by approximately 14% between 2001 and 2012, and current prevalence varies widely by ethnicity: from 3% for New Zealand Europeans to 5-10% for Māori persons (9, 10). Higher prevalences can be found in India, where estimates approach 16.5% in a given study (11). The International Diabetes Foundation estimates that 16.9% of all pregnancies in women aged 20 to 49 years worldwide are affected by hyperglycemia, including GDM or diabetes prior to pregnancy (2). In Southeast Asia, the estimated hyperglycemia prevalence in 2015 was 24.2%, and an estimated 88% of hyperglycemia cases in pregnancy are found in low and middle income countries (2). Less than 10% of these cases are due to preexisting diabetes (2), thus GDM impacts a large number of pregnancies worldwide.

1.2 Treatment for Gestational Diabetes

While treatment for GDM greatly improves pregnancy outcomes (12-14), the question of which treatment option is best for mother and child in terms of both health status and ease of utilization remains unanswered. After diagnosis, patients typically receive diet and exercise counseling to control their condition; however, some women with GDM require treatment beyond lifestyle modification. The standard of pharmacologic treatment for GDM has been injectable insulin; however, practice has shifted towards exploring the use of oral hypoglycemic agents (OHAs) for this purpose. Insulin therapy requires subcutaneous injections, careful self-monitoring of blood glucose levels, and proper storage (15, 16). This increases difficulty in adhering to a treatment regimen, increasing the risk of adverse

outcomes stemming from substandard treatment (17). Provision of OHAs, which require less monitoring and do not require injections, has emerged as an alternative treatment regimen.

One of the most important considerations for prescribing to pregnant women is whether or not a given medication has the potential to cross the placenta from mother to the fetus. Insulin has been shown to minimally cross the placenta (16), which indicates that there will be minimal risk to the health and development of the fetus. Certain formulations of insulin, such as insulin aspart, do not appear to cross the placenta at all (18). In general, placental transport of insulin is likely obstructed by its large molecular weight. Metformin is a biguianide, which acts as an insulin sensitizer mainly by reducing hepatic glucose production through suppression of gluconeogenesis and increasing peripheral glucose uptake. Metformin does not enhance insulin secretion, reducing the risk of hypoglycemia associated with treatment. Because of its small molecular size and low protein binding, metformin does cross the placenta, though the extent varies (18, 19). A study of women using metformin for polycystic ovary syndrome (PCOS) demonstrated that fetal metformin concentrations are comparable to maternal levels (20). First trimester exposure to metformin for PCOS has not been associated with an increased in observed birth defects or early pregnancy loss (21), however, prior research may be limited by small sample sizes.

Despite the potential for placental transport, metformin may have some advantages over glyburide when deciding on an OHA treatment. Metformin has been associated with less maternal weight gain than glyburide or insulin, which has important consequences for the pregnancy, delivery, and infant health (22, 23). Additionally, two RCTs which directly compared metformin to glyburide have observed fewer cases of macrosomia and large for gestational age in neonates exposed to metformin in utero versus glyburide; however, failure rates of metformin were also higher in these studies (24, 25). Failure rates in a recent trial of south Indian women were similar among women assigned metformin or glyburide, however, glyburide dosage differed from the trials by Moore et al and Silva et al (maximum

dose of 15 mg/day versus 20 mg/day) (26). Because of the ongoing concerns about placental transport of metformin and lack of information on long-term outcomes in children, there may be some hesitation to recommend metformin for first line GDM treatment over glyburide.

1.3 Long-term Effects of Oral Hypoglycemic Agent Use During Pregnancy

In addition to the concern surrounding placental transport of metformin, there may also be long-term implications of how well GDM is controlled during pregnancy that contribute to child health as well. While GDM alone has been associated with adverse child outcomes, particularly childhood obesity, these associations are inconsistent and often attenuated after controlling for maternal body mass index (BMI) (27-29). No association between maternal hyperglycemia and offspring obesity at 2-3 and 5-7 years has been demonstrated in a subset of women participated in the Hyperglycemia and Adverse Pregnancy Outcomes study (30, 31). However, a study of 263 mother-child pairs found a positive association between increasing maternal glucose concentration and childhood BMI at age 3, independent of maternal prepregnancy BMI (32). Nonetheless, long-term metabolic outcomes related to medication management of GDM remain a high priority concern (33).

Based in part on the results of the two-year follow-up to the MiG trial, there is speculation that metformin may have beneficial effects on the development of fat mass for children exposed in utero, due to a reduction in insulin resistance (34). Treatment with metformin for GDM has been associated with less weight gain during pregnancy than treatment with insulin, and metformin has been shown to contribute to weight loss in women of reproductive age (35). Whether or not these effects contribute to childhood obesity after in utero exposure is unknown. The comparative effectiveness of pharmacologic treatment options for GDM may therefore have implications for childhood growth, but there have been few opportunities to explore this in existing small, short randomized trials.

Additionally, both maternal prepregnancy BMI and GDM have been associated with adverse developmental outcomes. In a study of 1,311 mother-child pairs in the U.S., children of obese mothers had increased odds of problems with emotion, conduct, and hyperactivity (36). While the authors state that they observed differences in effects by GDM status, they do not present stratified results because the interaction term was not statistically significant. Both pre-pregnancy diabetes and GDM were consistently associated with lower IQ scores in offspring at age 8 in the Avon Longitudinal Study of Parents and Children (37), however, small numbers of diabetes cases (n=59) contributed to wide confidence intervals of estimates. There is also conflicting information about how neonatal hypoglycemia, a common outcome of GDM, influences long-term neurodevelopment (38-40). Further research is needed to understand if pharmacologic treatment type for GDM influences child development, independent of maternal prepregnancy BMI.

1.4 Evidence Summary for Comparative Effectiveness of Gestational Diabetes Treatments

To date, there have been a number of randomized controlled trials comparing maternal and neonatal outcomes of metformin to insulin (Appendix 1). The most influential trial is the Metformin in Gestational Diabetes (MiG) trial, performed in 10 New Zealand and Australian urban obstetric hospitals (41). Researchers enrolled 751 women aged 18 to 45 years and diagnosed with GDM between 20 to 33 weeks of gestation. Women were excluded if they had a prepregnancy diagnosis of diabetes, contraindication to metformin, a known fetal anomaly, gestational hypertension, preeclampsia, fetal growth restriction, or ruptured membranes. Participating women were block randomized (block size 4) and stratified according to study site and gestational age at time of randomization to receive either a starting dose of 500mg of metformin once or twice daily with food, or insulin according to usual practice of the study site. A total of 731 women remained in the study for analysis (metformin N=363, insulin N=370). There were no significant differences between women assigned treatment with metformin or insulin for infant respiratory distress,

5

phototherapy, birth trauma, 5-minute Apgar score, or neonatal intensive care unit admission. Risk of neonatal hypoglycemia (defined as any blood glucose level <28.8 mg/dL) was reduced for women assigned to metformin (RR=0.41, 95% CI 0.21, 0.78), while risk of spontaneous preterm birth was higher among women assigned to metformin (RR=1.77, 95% CI 0.95, 3.28). Average gestational age at birth was slightly reduced for the metformin group as well (38.3 [SD=1.4] weeks versus 38.5 [SD=1.3] weeks, p=0.02). Women assigned to metformin experienced less weight gain from enrollment to the 37th week of gestation and more weight loss from enrollment to the postpartum visit, while weight gain from early pregnancy to enrollment was the same for both study groups, prior to randomization (41). Women who required supplemental insulin (N=168) were also compared to women who remained on metformin alone (N=195). Those requiring supplemental insulin were more likely to have a higher BMI, Polynesian (Pacific Peoples or Māori), and a previous GDM diagnosis. No differences were observed between the two groups with respect to other neonatal outcomes (41).

Following publication of the primary results of the MiG trial, additional results from this trial have been made available. The percent increase of maternal plasma triglycerides from randomization to 36 weeks of gestation was greater for women assigned to metformin (21.9%) versus insulin (9.7%), but there were no differences observed in cord plasma triglycerides, neonatal metabolic markers, or anthropometric measurements (42). A follow-up study of offspring born to women enrolled in the MiG trial (N=323) found that children who had been exposed to metformin in utero did not differ from children exposed to insulin with regard to weight, height, or abdominal fat at two years of age. However, children exposed to metformin had increased upper-arm circumference, subscapular skinfold thickness, and biceps skinfold thickness (43). These differences suggest that treatment with metformin may be associated with the development of less visceral fat in offspring as compared with insulin, though longer follow-up is needed to determine the true implications

of these observations. No differences in systolic or diastolic blood pressure were observed between children born to mothers in either treatment group (44). A separate follow-up of children born to MiG trial participants at age 2 found no meaningful differences in Bayley Scales of Infant Development Scores by treatment status, though they did observe differences by country (New Zealand versus Australia) (45).

Additional evidence from smaller randomized controlled trials generally echoes the conclusions of the MiG trial, that metformin is safe and effective for GDM treatment. Ijäs et al. conducted a randomized controlled trial among 97 eligible women with GDM identified in primary care in Finland using risk factor-based screening (46). Forty-seven of the participating women were randomized to receive metformin (750mg once daily for week one, twice daily for week two, and thrice daily for week three), and 50 received insulin according to hospital guidelines. Mean birthweight of neonates did not differ significantly between the two treatment groups (metformin: 3712g ± 432, insulin: 3558g ± 593), while mode of delivery outcomes did differ by treatment: women receiving metformin were less likely to undergo spontaneous vaginal delivery and more likely to need a caesarean section. There were more cases of neonatal hypoglycemia in the insulin group (n=7 [14.0%] versus n=4 [8.5%]) but this outcome was rare overall. In comparing women assigned to metformin who required supplemental insulin versus those who did not, the mean birthweight was higher among women requiring supplemental insulin, but baseline BMI was also significantly higher, and estimated relative effect measures appear to be crude rather than adjusted for this likely confounder (46). A follow-up study of 18-month-old offspring born to these women revealed slight differences in weight for those exposed to metformin in utero (mean=12.051kg, SD=1.87) versus insulin (mean=11.318kg, SD=1.45). No differences were observed in social, emotional, or language development, which has been observed to be impacted in offspring born to women with GDM (47).

Other randomized controlled trials have concluded that metformin is similar to insulin, though these have largely been limited by small sample sizes (N=200 or less), limited follow-up time, or methodological limitations (see Appendix, Table 1, page 38) (22, 48-51). Meta-analyses of these trials have similarly stated that metformin and insulin appear to be equally efficacious, and in some cases, may be a better treatment option than insulin based on existing evidence (52, 53). However, there are still limited data on the long-term outcomes stemming from treatment decisions. Additionally, many of the maternal and neonatal outcomes studied are rare, with some studies observing fewer than 10 cases of adverse neonatal outcomes in each study arm. The lack of observed differences between metformin- and insulin-treated pregnancies may reflect a lack of adequate power to detect rare outcomes and as a result, poor precision. Outcome definitions, including neonatal hypoglycemia, also varied among RCTs.

There are few existing observational studies comparing the effectiveness of metformin and insulin for GDM treatment. The largest observational study to date was performed in Auckland, NZ and enrolled 1,269 women with GDM (54). Three hundred and seventy-one women received dietary counseling, 399 received insulin, and 465 received metformin (249 metformin alone, 216 metformin and insulin). Although this population was similar to that of the MiG trial, overall preterm birth was observed less frequently in women receiving metformin (12.5%) versus insulin (19.2%), though incidence of spontaneous preterm birth did not differ by treatment. In contrast, overall preterm birth was more common among metformin-treated women (12.1%) than insulin-treated women (7.6%) in the MiG trial, which was driven largely by spontaneous preterm birth. NICU admissions, hypoglycemia (defined as glucose <2.3 mmol/l), and large for gestational age were all more frequently observed among the women receiving insulin versus metformin or dietary counseling. In a comparison of women receiving metformin alone versus metformin and insulin, infants exposed to supplemental insulin were more likely to be large for gestational

age, admitted to the NICU, and experience respiratory distress, while differences in neonatal hypoglycemia were not significant. This study did not estimate relative effects or absolute differences, and observed outcomes were possibly driven by differences in baseline severity (54).

Tertti et al. performed a case control study among women treated with metformin, insulin, or neither, matched on pre-pregnancy BMI and age in Finland (55). Many of their estimated effect measures are not reported because they were not statistically significant. Neonates with hypoglycemia were less likely to have been exposed to metformin than insulin in utero (crude odds ratio [OR]=0.4, 95% CI 0.2, 0.9). However, women who received insulin treatment had higher blood glucose levels at the time of screening, which does not appear to be controlled for in their analysis. In another case-control study of 200 women treated for GDM in the United Kingdom, half had been treated with metformin and half with insulin (56). Among the 127 women originally identified as being treated with metformin, 13 required supplemental insulin, 5 stopped taking metformin because of side effects, and 9 stopped treatment because they were uncomfortable with using a drug not licensed for use in pregnancy at the time (2007 through 2008). Women treated with metformin and insulin were similar with regard to baseline characteristics such as age, ethnicity, BMI, and gestational age at entry; however, insulin-treated women had higher median glucose values and were more likely to have had a previous diagnosis of GDM. Women who had received insulin experienced greater weight gain and more preterm births (n=10 for insulin, n=0 for metformin). Adverse neonatal outcomes were rare, however, jaundice, hypoglycemia, special care baby unit admissions, and birthweight ≥ 4kg were all more common in neonates born to women treated with insulin versus metformin (56).

Existing observational studies of the comparative effectiveness of glyburide and insulin have utilized health care data in the United States, generating large sample sizes but suffering from certain limitations (57, 58). While point estimates from these studies largely

fell within the confidence intervals of estimates from previous research with respect to neonatal macrosomia, hypoglycemia, NICU admission, and jaundice, these estimates tended to be much more precise due to a large sample size. For example, the point estimate and 95% confidence interval for the risk ratio estimating the incidence of macrosomia in women treated with glyburide versus insulin were both above the null in Camelo Castillo et al. and Cheng et al. However, in Langer et al. and other studies (59, 60), confidence intervals crossed the null while point estimates remained above the null, indicating increased risk of macrosomia associated with glyburide, but not statistical significance. This degree of precision is a key benefit to large observational studies; nonetheless, Camelo Castillo et al. also had a large proportion of missing data on laboratory values and may have had residual confounding by obesity due to the use of ICD-9-CM diagnosis codes rather than actual anthropometric measurements. The lack of information on glucose values and degree of glycemic control has been cited as a major limitation of these studies and highlights the typical epidemiological trade-off of sample size versus granularity of data (61). To date, there are no studies of equivalent or better quality evaluating the comparative effectiveness of metformin.

1.5 International Treatment Guidelines for Gestational Diabetes

Despite the limitations in prior research, including small sample sizes and limited follow-up time, clinical guidelines for the treatment of GDM worldwide have begun to recommend consideration of metformin or glyburide for first-line treatment. In the US, the American College of Obstetricians and Gynecologists (ACOG) previously stated that insulin and oral medications (either glyburide or metformin) are equally efficacious and either can be used as first-line therapy, though insulin is currently the only treatment approved by the US Food and Drug Administration for treatment of GDM (62). However, in recent guidelines, the ACOG stance has changed based on evidence from recent meta-analyses, and now does not recommend either OHA as a first-line treatment (63). The choice of OHA may

depend on the health care system where the woman receives treatment; for example, the University of North Carolina Chapel Hill School of Medicine directs obstetricians to manage GDM patients with glyburide first (64), whereas GroupHealth Cooperative guidelines recommend metformin before glyburide in women who are unwilling to take insulin (65). The Canadian Diabetes Association notes that women who are nonadherent to or refuse insulin can be offered glyburide or metformin, but that use of oral agents in pregnancy is considered off-label and this decision should be discussed with the patient (66). In the United Kingdom, the National Institute for Health Care and Excellence recommends offering metformin first, followed by the addition of insulin if blood glucose targets are not met, or replacement with glyburide if metformin is not tolerated (7). Neither metformin nor glyburide have been approved for use during pregnancy in Australia at this time and are not explicitly recommended by the Royal Australian College of General Practitioners (7). There are no existing statements from the Royal Australian and New Zealand College of Obstetricians and Gynaecologists on pharmacologic treatment options for GDM. Preliminary data suggest that treatment with metformin has been more prevalent than treatment with glyburide in NZ. While treatment with metformin is endorsed by the NZ Ministry of Health (10), the NZ Medicines and Medical Devices Safety Authority states that the safety of metformin use during pregnancy is not yet established (67). While these guidelines recognize the benefits of providing OHA medications to women with GDM, they note that more information on the long-term consequences of this treatment decision is needed.

CHAPTER 2 – SPECIFIC AIMS

2.1 Specific Aims

GDM is a frequently observed pregnancy complication characterized by glucose intolerance, and incidence is expected to increase with rises in maternal obesity (2, 68). Women with elevated blood glucose are at increased risk for delivery of large-for-gestational age infants, and there is a strong monotonic association between maternal glucose control and infant birthweight (69, 70). Up to one-third of women with GDM require pharmacologic treatment to achieve adequate glucose control (15, 71). While glucose control greatly improves pregnancy outcomes (12, 14, 72), the question of whether OHAs are as safe and effective as injectable insulin for the mother and her infant remains unanswered. Injectable insulin has previously been the standard of care for GDM; however, glyburide and metformin (both OHAs) have become increasingly used for this indication (15). Recently, the safety and effectiveness of OHAs has been identified as a priority research topic (33). Limited randomized controlled trials and observational studies have compared the treatments with regard to immediate maternal and infant outcomes; however, the effectiveness of OHAs for this purpose has not been definitively established. While prior research comparing metformin to insulin has concluded that metformin is safe for use during pregnancy (22, 41, 46, 48-52, 55), these studies had limited analytic power to examine less frequent outcomes or heterogeneity of treatment effects, and short follow-up times prevented evaluation of long-term effects. While U.S. providers have primarily turned to glyburide as the OHA of choice for GDM, but emerging research from large observational studies (n≈9,000) suggested that glyburide may be associated with more adverse neonatal outcomes than insulin (57). The studies of metformin have not been adequately powered to evaluate

differences in these outcomes, despite international recommendations to prescribe metformin as an alternative to insulin for GDM (7, 10, 66).

Therefore, we propose to examine the comparative effectiveness of metformin versus insulin for the treatment of GDM in a longitudinal observational study using routinely collected health care data from New Zealand (2005 to 2015). These extant data sources are readily linked, and will provide data on pharmaceutical dispensing, primary maternity care services, laboratory data, and postpartum outcomes for mothers and their infants for pregnancies requiring pharmacotherapy for GDM. We will also examine child growth and development outcomes assessed at age 4 from a pre-school screening program introduced in 2009. Linking these data sources will allow us to report precise measures of any differences between the treatment groups with robust control of key confounders (glucose control and pre-pregnancy obesity).

2.1.1 Specific Aim 1

Aim: Measure the comparative effectiveness of treating GDM with metformin versus insulin on perinatal and neonatal outcomes in pregnant women and newborns. To accomplish this, we will estimate risk ratios, risk differences and 95% confidence intervals to compare the incidence of adverse maternal outcomes (e.g., preeclampsia, caesarean section delivery, induction) and adverse neonatal outcomes (e.g., preterm birth, large for gestational age, macrosomia, and neonatal hypoglycemia) in pregnancies treated with metformin versus those treated with insulin.

2.1.2 Specific Aim 2

Aim: Estimate the association of treating GDM with metformin versus insulin with adverse child health outcomes. Using data obtained from the B4 School Check health and development visit offered at age 4 in children born between 2005 and 2012, we will obtain measures of height, weight, and the results of a screening questionnaire to assess

psychosocial development of children born to mothers treated with metformin or insulin for GDM.

2.2 Hypotheses

Based on results from published randomized controlled trials and meta-analyses, we hypothesize that metformin will be a non-inferior treatment option to insulin with respect to infant birthweight, an indicator of how well treatment is able to control a mother's hyperglycemia. We expect the treatment groups to be similar but differ based on markers of disease severity like maternal BMI and timing of treatment initiation. We hypothesize that metformin will be associated with a reduced risk of certain neonatal events, based on meta-analysis results. We further expect to observe no meaningful differences in child growth or psychosocial development when comparing children born to mothers treated with metformin versus insulin for GDM.

2.3 Rationale

To date, there are no large studies comparing the effectiveness of metformin versus insulin. Although randomized controlled trials have long been considered the gold standard of clinical research, observational studies have many advantages for determination of real-world comparative effectives of two therapies. Observational research is comparatively affordable and practical, as treatments are not provided and assigned by the study, but rather are observed in regular clinical practice, allowing for consideration of real-life treatment determinants. Additionally, the use of secondary data that is routinely collected leads to greater efficiency in research. The availability of longitudinal health care data allows us to examine long-term outcomes of treatment decisions without waiting years for these outcomes to materialize. Furthermore, observational research using extant data is less subject to volunteer bias and loss to follow-up, which are often major threats to validity in prospective studies. These advantages are particularly relevant to the comparative effectiveness of metformin versus insulin, as the prevailing concern about metformin

14

treatment for GDM is the extent to which it impacts the long-term health of the offspring due to placental transport. Although the MiG trial, the largest trial to date, has planned to examine outcomes in children born to women participating in the study at age 5, there has already been significant loss to follow-up (44%) observed at the 2-year follow-up (43). A new prospective experimental study would need an even larger initial sample size, which quickly becomes prohibitively expensive and difficult to recruit.

As the prevalence of GDM increases and more women require pharmacologic therapy to achieve glucose control, it is important to study the comparative benefits and risks of different treatment options. Optimization of medical management of GDM could greatly improve pregnancy outcomes for a growing population of women. By using existing health care data, we are able to follow a large population of mothers and their children longer than in most randomized controlled trials with more comprehensive follow-up. When completed, this study will represent the largest study to date to compare metformin to insulin for GDM treatment. Our findings have the potential to inform clinical practice and provide needed data on the long-term effects of in utero exposure to metformin.

CHAPTER 3 – METHODS

3.1 Data Sources

We conducted a retrospective observational cohort study of pregnant women who delivered between 2005 and 2015, and their offspring. This study was conducted in the NZ population, due to the availability of extant secondary health care data and knowledge that metformin has been recommended to treat GDM in clinical practice. The NZ Ministry of Health maintains 20 District Health Boards for the purpose of funding, planning, and providing health services. From this, approximately 14 individual collections of health information exist, initially collected for payment, funding, or monitoring, though recently there has been more focus on research. Because health services in NZ are funded by the government, eligible persons may receive free inpatient and outpatient public hospital services, including pregnancy services, and subsidies on prescription items (73). Data collections are linkable through an encrypted National Health Identifier (NHI). This also includes linkages between mother and child through unique identifiers, allowing us to directly examine infant and child outcomes of medications used during pregnancy. Utilizing previously collected data allowed us to conduct analyses at relatively low-cost while being confident that these data are of high quality and valid.

These longitudinal health data are a rich source of information on prescription dispensing, maternal health status, and infant and child health outcomes. The linked data resources in NZ data for an entire population, rather than restricting to insured individuals, a single clinical center, or study volunteers. Additionally, existing NHI number linkages between mother and child simplified long-term follow-up and improved rates of loss-to-follow-up and missing data.

Multiple data collections were linked with mother and child NHI numbers, including the National Health Index (NHI), Database, National Maternity Collection (MAT), Pharmaceutical Claims Data Mart (PHARMS), National Minimum Data Set (NMDS), Mortality Dataset (MORT), Laboratory Claims Collection, and B4 School Checks (B4SC) (73, 74). Brief descriptions of each collection are available in Table 1. Mother-infant pairs were previously linked and available in MAT via unique pregnancy identifiers that link to National Health Identifiers (NHIs) for mothers and infants. All data were requested with encrypted NHI numbers to protect the identity of individuals in the study. Linkage between data sources was accomplished using the encrypted identifiers.

3.2 Study Population

New Zealand is a Pacific Island nation with a population of approximately 4.5 million people. The majority ethnicity is NZ European (68%), followed by the indigenous Māori population (15%), Asian (9%). and Pacific Peoples (7%) (74). The proportion of pregnancies experiencing GDM has increased in NZ, with an annual increase of nearly 14% between 2001 and 2012 (76). Prevalence of GDM varies widely by ethnicity, reported as 3% for New Zealand Europeans to 5-10% among Māori, 4-8% percent among Pacific peoples, and 4-8% among Asian Indians (10). It is likely that the rates of GDM among Māori and Pacific women reflect disparities in screening during the time these data were collected and are in fact, higher than currently measured.

Pregnancies were identified using the MAT which provides information on primary maternity services from the first midwife encounter through the postpartum period for pregnancies with a duration or 20 weeks or later, or resulting in an infant weighing at least 400 grams at delivery. This includes information from midwife encounters and hospitalizations. Because stillbirths are not reliably reported in MAT, we linked available mortality data from years 2005 to 2013 to more accurately classify infant birth outcomes. We estimated the start of pregnancy using gestational age as reported in MAT. Gestational age

in MAT is reported as either the duration of completed weeks between a woman's self-reported last menstrual period (LMP) date and her delivery date or derived from a clinical assessment during pregnancy or at birth. For those missing gestational age but having a recorded date of LMP, we estimated gestational age by calculating the number of completed weeks from the LMP to the delivery date. If both LMP and gestational age were missing, we removed these deliveries from our cohort.

Pharmaceutical data from the PHARMS are reliably available from 2005 onwards. Thus, in order to ensure that we observed the full 6-month lookback history of prescription use for each pregnancy, we restricted our cohort to pregnancies with an estimated start date of June 30, 2005 or later. We additionally excluded pregnancies if the mother was younger than 15 years old at delivery, or older than 45 years old. Finally, we excluded pregnancies with missing information on gestational age.

3.3 Exposure Assessment

We limited our cohort to pregnancies complicated by GDM. In NZ, according to national guidelines, women are first screened for gestational diabetes with the "post-polycose" glucose tolerance test (10). If the value of this test is ≥7.8 mmol/L they will undergo the oral glucose tolerance test and diagnosed with GDM if the fasting test result is ≥5.5mmol/L or the 2-hour test result is ≥9.0mmol/L. Following diagnosis of GDM, New Zealand women are provided with a blood glucose meter for home-based glucose testing and are given a prescription for testing strips (personal communication, Lesley MacLennan, Counties Manukau District Health Board). Because these prescriptions are subsidized by the NZ government, claims are filed for payment to the NZ Pharmaceutical Management Agency (PHARMAC) and recorded in the PHARMS collection. We considered a claim for prescription blood glucose test strips to be evidence of a GDM diagnosis in each pregnancy if the first prescription was filled after the start of the GDM screening period, 24 weeks gestation. If a prescription for test strips was filled anytime during the lookback period (6

months prior to pregnancy through 24 weeks gestation), we considered this to be evidence of prior type 2 diabetes and excluded these pregnancies. We used a shorter lookback period (3 months prior to pregnancy) for women with closely spaced pregnancies. This was done to avoid excluding pregnancies immediately following a pregnancy complicated by GDM and beginning less than 6 months after the prior delivery. We also examined claims for antidiabetic medications (insulin, metformin, glibenclamide, glipizide, glicazide, pioglitazone, and acarbose) and excluded women who had a prescription claim for any antidiabetic prior to 24 weeks gestation.

The cohort was further restricted to women who were treated pharmacologically with metformin, insulin, or both. For women who received both metformin and insulin during their pregnancy for the first time after filling a test strip prescription, we defined "combination therapy" as a dispensing for both medications within 7 days of the first claim. We treated combination therapy co-initiators as a separate group, as the need for both medications may have indicated an additional degree of GDM severity. For our primary analyses, we restricted our study population to pregnancies initially treated with metformin or insulin monotherapy. A woman who later added another therapy (>7 days after the first claim) remained in the cohort and analyzed as part of the treatment group with which she started. Exposure classification was based on first prescription dispensed for primary analyses, using a first-treatment-carried-forward approach for effectiveness.

3.4 Outcome Assessments

3.4.1 Outcome assessment for Aim 1

A summary of all targeted outcomes is presented in Table 3.2. In Specific Aim 1, we will assess the comparative effectiveness of treating GDM with metformin versus insulin on perinatal outcomes in pregnant women and newborns (Figure 3.1). Perinatal outcomes will primarily be obtained from the MAT collection, with enhancement from NMDS where information is not available in MAT.

3.4.1.1 Maternal outcomes

The risk of the following maternal outcomes will be compared between women initiating metformin treatment versus insulin:

Caesarean section delivery

Caesarean births, or c-sections, are defined as surgical deliveries of an infant through incisions made in the mother's abdomen and uterus. There are a number of risk factors for a c-section, including problems with the placenta, failure of labor to progress, multiple gestation pregnancy, a large fetus, or maternal medication conditions such as diabetes or high blood pressure (75). C-section delivery was measured using variables from the MAT collection, which distinguish between elective caesarean deliveries (performed as a planned procedure before or after the onset of labor, when the decision was made prior to commencement of labor) versus emergency caesarean (indicating that it was performed as an emergency procedure before or after the onset of labor). Both elective and emergency caesarean deliveries were of interest. Emergency caesarean deliveries may occur during birth because of problems with labor. Elective caesareans may be indicated, however, for women with anticipated problems, including a large-for-gestational age neonate. In NZ, c-section deliveries are commonly performed surgical obstetric interventions, with approximately one-quarter (24.3%) of women giving birth via caesarean section over the past decade (23.0-25.3%). Among these, approximately half were emergency caesarean section deliveries and others were elective. From 2004 to 2013, elective caesarean section deliveries increased from 8.8% of live births to 12.0% (76). Because both outcomes are plausible downstream consequences of GDM, we considered both types of c-section deliveries in our analysis. In MAT, these variables indicating the procedure was performed were derived from ICD-10-AM delivery procedure codes available in the NMDS.

Induction of labor

Labor induction refers to the use of medications or other methods to cause contractions to start and induce labor. This is generally limited to situations where there is a problem with the pregnancy, or when the baby is overdue (77). This was defined using dichotomous variable indicting whether or not a woman received a medical or surgical induction of labor during a delivery event, available from MAT and based on ICD-10-AM obstetric procedure codes.

Severe perineal trauma

Severe perineal trauma is defined as third- or fourth-degree perineal tears during childbirth. Third-degree tears are defined as partial or complete disruption of anal sphincter muscles. A fourth-degree tear involves the disruption of anal sphincter muscles with a breach of the rectal mucosa (78). Fecal incontinence may result from perineal tears in 29 to 57% of affected patients. High birth weight is considered a risk factor for perineal tears (79). Severe perineal trauma was defined by a flag third or fourth degree perineal tears or surgical repair of a tear taking place during a delivery event, derived from the MAT data collection. These variables were based on selected ICD-10-AM diagnosis codes during the delivery event.

Preeclampsia

Preeclampsia is defined by high blood pressure (greater than 140/90) occurring after 20 weeks or pregnancy, after experiencing normal blood pressure in the beginning of pregnancy. This condition is also accompanied by increased swelling and protein in the urine (80). While an increased risk of preeclampsia has been observed among women with GDM (81, 82), the exact relationship is unclear, as preeclampsia may have begun to develop before GDM is recognized. More severe GDM has been associated with increased risk of preeclampsia (83), suggesting that glucose control is an important determinant of preeclampsia incidence. Previous comparisons of metformin and insulin have not shown

significant differences in preeclampsia incidence between treatment groups, though the risk of preeclampsia is often lower in women treated with metformin. To examine this outcome, we identified diagnoses of preeclampsia in NMDS hospitalization data.

3.4.1.2 Infant outcomes

The incidence of the following neonatal outcomes will be compared between women initiating metformin treatment versus insulin:

Macrosomia/Birthweight

Macrosomia is defined as a birthweight of >4,000g or >4,500g or more, regardless of gestational age. We used 4,000g as the primary cut-off in analyses, as this is the definition cited in the NZ clinical guidelines for GDM screening and treatment (10). Birthweight of the infant at time of birth in grams was also available as a continuous measure, usually measured to the nearest 5g and obtained within one hour of birth. Thus, we were able to evaluate macrosomia as a dichotomous variable according to NZ definitions, as well as measure birth weight continuously. Fetal macrosomia is a common outcome of GDM and may lead to other adverse outcomes such as birth injuries and hypoglycemia.

Large for gestational age

Large for gestational age (LGA) is defined as a birthweight greater than the 90th percentile for age. An infant that is LGA is larger than expected for his/her age and gender and is a common outcome of gestational diabetes (2, 84). Current age-, sex-, and ethnicity-specific birthweight percentiles were not available for our study population. To obtain this information, we calculated these birthweight percentiles for infants born between 2001 and 2015 in NZ using available data in the MAT data collection. We categorized infants as LGA if their birthweight exceeded the 90th percentile, based on their gestational age, sex, and ethnicity.

Neonatal hypoglycemia

Neonatal hypoglycemia (low blood sugar) occurs when the infant does not have enough glucose in their blood that is needed for energy. This is a major concern in pregnancies complicated by GDM. Because glucose is primarily needed for the infant's brain, neonatal hypoglycemia may have long-term implications for neurodevelopment. Transient hypoglycemia, which occurs when a single initial measurement of glucose concentration is low, followed by a second measurement above a defined cutoff, has recently been linked to lower test achievement scores in childhood (38), though these findings have yet to be replicated. In addition, there are no uniform management guidelines for infants with low glucose concentrations with no overt symptoms (85, 86). Neonatal hypoglycemia was defined as an outcome using clinical codes (ICD-10-AM) available through the NMDS.

Respiratory distress

In Vignoles et al., women with either pregestational diabetes or GDM were more likely to deliver a neonate that suffered from severe respiratory failure (87). While metformin has generally not been associated with respiratory distress, glyburide was associated with an increased risk over insulin in Camelo Castillo et al.(57). The association of metformin and respiratory distress may have been difficult to detect in small studies because it has been a rare outcome (3.8% in the MiG trial) (41). To determine neonatal respiratory distress in our study, we identified diagnoses from clinical codes in NMDS.

Shoulder dystocia

Shoulder dystocia occurs during a vaginal delivery when, following delivery of the fetal head, additional obstetrics maneuvers are required to continue delivery. This can lead to birth injuries like brachial plexus injuries and fractures of the clavicle or humerus. Additionally, these complications are associated with postpartum hemorrhage and fourth-degree tears (88). GDM is associated with an increased risk of shoulder dystocia (89).

Macrosomia also increases the risk of shoulder dystocia, as well as maternal obesity or excessive weight gain (90). A preliminary examination of NZ hospital admissions from 2006 to 2013 for infants with relevant ICD-10 codes for conditions related to pregnancy and childbirth (O66.X) revealed that the incidence of shoulder dystocia is low in the general pregnant population (approximately 0.3%). While it's estimated that the incidence of shoulder dystocia is closer to 0.6-3% (91), this estimate pertains to vaginal deliveries only; our denominator for the NZ hospital admissions would include c-section deliveries as well. We used ICD-10-AM diagnosis codes for hospital admissions recorded in the NMDS to identify shoulder dystocia cases.

Preterm birth

Preterm birth, defined as delivery prior to 37 completed weeks of gestation, is a pregnancy outcome with a complicated etiology. There are many known risk factors for preterm birth in singleton pregnancies, including GDM, sexually transmitted infections, maternal age, smoking, stress, being either underweight or obese before pregnancy, ethnicity, and certain environmental exposures (92). The association of treatment with metformin for GDM and preterm birth is unclear. In the MiG trial, researchers found an increased risk of preterm birth overall associated with metformin versus insulin (RR=1.60, 95% CI 1.02, 2.52) (41). However, in Goh et al., an observational study conducted in a similar setting, there were more preterm births observed in women who received insulin treatment as compared with women receiving metformin (54). The MAT data collection contains information on gestational age, which is derived from either the first day of the woman's last menstrual period, a clinical assessment during pregnancy, or examination of the infant after birth. We used a dichotomous variable for preterm birth (less than 37 weeks versus 37 weeks or more), derived from gestational age.

Stillbirth

Stillbirth is a recognized adverse outcome of pregnancies complicated by preexisting diabetes (93, 94); however, the relationship with GDM is less clear. Although an increased risk of stillbirth has not been universally observed among women with GDM, studies may have been too small to detect such a rare outcome, or methodologically flawed (95). According to a recent study by Hutcheon et al., studies of GDM and stillbirth may be impacted by immortal time bias by beginning follow-up at 20 weeks awhile requiring pregnancies to survive to 24-28 weeks in order to receive a GDM diagnosis (96). In an analysis restricting to births occurring after 28 completed weeks of gestation, Hutcheon et al observed increased odds of stillbirth associated with GDM (OR=1.25, 95% CI 1.11-1.41). We defined stillbirth based on NZ mortality data.

Neonatal intensive care unit admissions

Neonatal intensive care unit (NICU) admissions may be indicated from newborns suffering from conditions closely related to GDM, including macrosomia, respiratory distress, or hypoglycemia. In a study of pregnancies complicated by GDM or type 2 diabetes in the National Women's Health database in Auckland, NZ from 1999 to 2000, 29% of NICU admissions were a result of GDM complications (97). The incidence of NICU admissions ranged from 12.7% in the metformin group to 18.7% in the insulin group in the Goh et al. observational study occurring in Auckland, NZ. (54) The NMDS contains detailed information on hospitalization events, including a Health Specialty Code to broadly categorize the event type (e.g. cardiology, oncology, rheumatology). Included are two codes that indicate if an infant's event took place in a Level II or Level III NICU (P42 and P43). While we were able to use these codes to determine if an infant was admitted to the NICU, we were not able to accurately assess the number of hours the infant remained under special care. The amount of time an infant spends in the NICU may provide information about the severity of the condition for which they are under care. To estimate this without exact measurements on

length of stay, we examined the proportion of infants who were discharged from the hospital after their mothers were discharged. This would suggest that the infant needed care for longer than a standard hospitalization, and likely wouldn't include the proportion of infants who are in the NICU for less severe conditions.

3.4.2 Outcome assessment for Aim 2

For Specific Aim 2, we assessed the comparative safety and effectiveness of treating GDM with metformin versus insulin on long-term childhood outcomes at age 4 (Figure 1). As noted above, the long-term effects of metformin treatment have not been extensively studied. Metformin crosses the placenta while insulin does not, causing concerns that that treatment with metformin may impact fetal development in ways that insulin would not and result in adverse childhood outcomes. To explore this, we leveraged regularly collected data on children in NZ at age 4 through a pre-school readiness assessment called the B4 School Check (B4SC). The B4SC program began in 2009 and offers a comprehensive health and development screen for 4-year-old children prior to school entry. From 2009 to 2012, participation varied from 52.5 percent to 75.6 percent in 2012; overall, participation rates during those 4 years was 66.3 percent. In 2015-16, participation was nearly 100% (98). We incorporated data from all available years (2010 through mid-2017), following infants born to women treated for GDM between 2006 and mid-2012.

3.4.2.1 Child outcomes

Growth outcomes

As part of the B4SC assessment, providers measure a child's height and weight. Providers are instructed to take all measurements on a hard surface. Weight is measured using a Seca 862 electronic floor scale or Tanita WB 100 S MA floor scale (or Seca 770 or Tanita HD-351 weighing scale). Height is measured using a Leicester Height Measure portable stadiometer or a Seca 214 portable stadiometer. Providers are further instructed to

take two readings of each measurement (height, weight, height, weight) and if the measurements do not vary by greater than 0.5kg for weight or 0.5cm for height, the reported measurement is the average of the two, rounded to the nearest 0.1kg or cm. If the measurements do differ by more than 0.5kg or 0.5cm, providers are instructed to take a third reading and average the two closest measurements (99).

We used height and weight values from the B4SC to calculate child BMI. These values were also used to estimate z-scores and percentile-based outcomes for weight, height, weight-for-height, and BMI. We used the 2006 World Health Organization (WHO) reference standards to calculate sex- and age-specific z-scores and percentiles (100). We used the 85th percentile as a cutpoint for defining overweight and obesity in children and applied this cutpoint for creating dichotomous outcomes for each growth measurement. We also examined the proportion of children with weight, height, weight-for-height and BMI z-scores ≥1 or ≥2. Additionally, the B4SC recommends referral for children with weight above the 97th percentile and BMI of 21 or over. We therefore used the 97th percentile as a cutpoint for extreme growth measurements and examined the proportion of children with a BMI measurement of 21 or over.

Behavioral and emotional development outcomes

The B4 School Check offers screening to help identify children that may need help with learning and development before entering school using the Strengths and Difficulties Questionnaire (SDQ) (101, 102). The SDQ is designed to assess the child's strengths in five subscales pertaining to different emotions or behaviors: prosocial behavior, hyperactivity, emotional symptoms, conduct, and peer problems (Appendix, Table 4). In New Zealand, the SDQ is offered in two different formats, depending on whether a parent or a teacher is administering the questionnaire. Parents or teachers respond, "not true", "somewhat true", or "certainly true" to each statement pertaining to the child's behavior (e.g., "Considerate of other people's feelings", "Often loses temper") in addition to other questions about emotional

difficulties. This instrument has been previously shown to be valid and reliable (101, 102).

Responses to these statements can be organized into the four subscales of hyperactivity,

emotional symptoms, conduct, and peer problems. Scores from these subscales are

summed to result in a Difficulties Score which can range from 0 to 40, with higher scores

indicating concern and possible need for referral.(99, 101, 103) The cut-off for a

"concerning" score is typically 17 for the parent-completed SDQ and 16 for the teacher-

completed SDQ; however, a review of the SDQ in the NZ population found that a lower cut-

off is more appropriate (14 for SDQ-P, 11 for SDQ-T).(104, 105) An additional prosocial

behavior scale is not included in the total Difficulties Score, and is scored so that an

absence of prosocial behavior (e.g. helping, sharing) receives a lower score.

3.5 Covariates and Treatment Effect Heterogeneity

We identified potential confounders of the association between GDM treatment and

our primary effectiveness outcome, birthweight, by reviewing literature and consulting

clinical experts. We assessed covariates in the baseline period (6 months prior to pregnancy

through 23 completed weeks of gestation). These covariates included maternal age, BMI,

timing of diagnosis and treatment, geographic region, ethnicity, socioeconomic status,

parity, and smoking status. Most maternal characteristics were recorded in MAT. For

demographic characteristics (age, ethnicity, region, socioeconomic status), we used the

MAT data collection as our primary source and augmented missing MAT data with available

information in the NHI. Noting that prevalence of GDM and availability of treatment may vary

across regions of NZ, we used the territorial local authority of the mother's residential

address is recorded at the time of delivery in the MAT data collection in our analysis.

Additionally, GDM prevalence varies across ethnic groups in NZ. The mother's self-identified

ethnic group is categorized as prioritized ethnicity, ethnicity code 2, and ethnicity code 3. We

used prioritized ethnicities in our analyses and reported the proportion of women who

reported a given ethnicity in any of the three categories. To measure socioeconomic status,

we used the NZ Deprivation Index (NZDep13). The NZDep13 combines census data related to income, home ownership, employment, qualifications, family structure, housing, and access to transportation or communications into a score for measuring relative socioeconomic deprivation. These scores are measured on a meshblock level, which is the smallest geographical area defined by Statistics New Zealand, and generally comprises 60 to 110 individuals. These scores are grouped into deciles, with a score of 1 indicating an area with the least deprived scores (106, 107).

The key health characteristics that we measured were maternal BMI, timing of diagnosis and treatment, smoking status, and parity. Maternal BMI was measured at the time of the first Lead Maternity Care (e.g., midwife, obstetrician) visit for a given pregnancy in the MAT data collection. This means that women may not have BMI measured until later in their pregnancy, if they did not seek prenatal care earlier. For our study, we used BMI measurements that were recorded prior to the initiation of GDM treatment and set other measurements to missing if they were taken after the prescription fill date. The timing of GDM recognition and treatment may have an impact on pregnancy outcomes and the initial treatment assignment (i.e., later recognition of worse glucose control may require management with insulin). We determined the timing of diagnosis and treatment based on dates of prescription claims for diagnostic test strips and antidiabetic treatments. We also characterized women as having a previous diagnosis of GDM if they met our diagnostic criteria in a previous pregnancy, and likewise characterized them as being treated for GDM if they met the eligibility criteria of metformin or insulin treatment for GDM more than once. Maternal smoking is an important risk factor for neonatal outcomes like preterm birth and stillbirth. Although there is no clear link between smoking and GDM, it could be an indicator of other lifestyle choices that we were unable to measure, such as nutrition. We will explore this as a possible confounder based on reported smoking status at the first encounter with the Lead Maternity Carer, as stored in the MAT data collection. Parity, defined by the

number of pregnancies lasting until a viable gestational age, was similarly measured in the MAT data collection and women were categorized as having 0, 1, 2, or 3 or more prior pregnancies.

Comparative effectiveness and safety of metformin versus insulin may vary by observable factors. We explored treatment effect heterogeneity by maternal ethnicity and infant sex. To achieve this, we presented stratified effect estimates of risk ratio and risk difference from stratified models for Aim 1 and presented stratified risk ratio estimates for the interaction with sex for Aim 2.

3.6 Statistical Analysis

For Specific Aim 1, we created a uniform period of follow-up between the initial anti-diabetic prescription and 92 days postpartum. While duration of pregnancy after initial treatment varied, we maintained uniform follow-up on the time scale of initial prescription to end of the perinatal period. We similarly maintained a uniform follow-up period between initial prescription and assessment at the B4SC. For this reason, we were able to calculate risks in our cohort. We used log-binomial regression to estimate risk ratios and 95% confidence intervals and linear risk models to estimate risk differences and 95% confidence intervals. For continuous outcomes (i.e., birthweight, z-scores) we used linear regression to estimate to model the relationship between metformin treatment as compared with insulin treatment for GDM and these outcomes.

Multiple imputation using chained equations was used to estimate missing covariate information (108). BMI was measured at the first prenatal care encounter in NZ, which means some women may have had their BMI recorded late in pregnancy. We reassigned extreme BMI values (<14 and >72) to missing. We then used multiple imputation to estimate BMI for pregnancies with missing data, including reassigned extreme values. We similarly imputed values for maternal ethnicity, smoking status (yes/no) and parity (0, 1, 2, or 3 or more prior births) at the time of the first LMC encounter.

To control for identified covariates, we estimated propensity scores for each pregnancy, calculated as the estimated probability of treatment with metformin based on measured covariates using logistic regression. We then used these values to estimate inverse probability of treatment weights (IPTW), calculated as the inverse of the probability of receiving the treatment actually received. Because women could contribute multiple pregnancies to our study, we used generalized estimating equations to account for within-woman correlations. We specified an independent correlation structure in IPT-weighted models, as is recommended (109). We assessed balance of covariates pre- and post-weighting by calculating standardized mean differences for individual covariates for comparison.

3.7 Software and Approval

All analyses were conducted using SAS 9.3 (SAS Institute, Cary, NC). This study was approved by the NZ Health and Disability Ethics Committee (Reference 16/NTA/66) and by the University of North Carolina at Chapel Hill Institutional Review Board (16-1121).

3.8 Tables and Figures

Figure 3.1 Study schematic
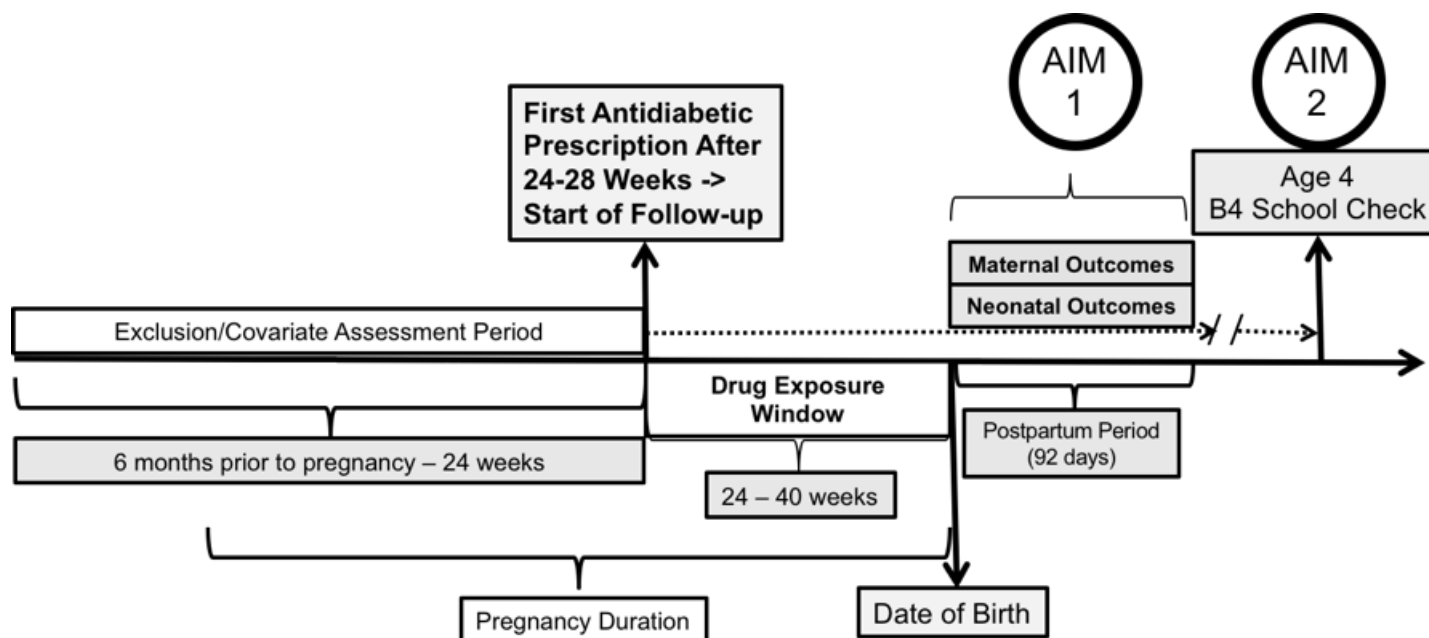
Table 3.1 Summary of New Zealand Data sources

Data Source	Description	Years Available
National Health Index Database (NHI)	Contains demographic information for those assigned a National Health Identifier	2004 - 2015
National Maternity Collection (MAT)	Contains data on primary maternity services and on inpatient and day-patient health event data during pregnancy, birth, and the postnatal period for women and their babies.	2004 - 2015
Pharmaceutical Dispensing Warehouse (Pharmhouse)	Contains information on all subsidized prescriptions dispensed from community pharmacies.	2003 - 2015
National Minimum Data Set (NMDS)	Contains records of all hospital discharges in New Zealand.	2003 - 2015
Mortality Dataset (MORT)	Classifies the underlying cause of death for all deaths registered in NZ, including all registered fetal deaths.	2004 - 2013
Laboratory Claims Data Warehouse	Contains claim and payment information for community laboratory tests.	2003 - 2015

Table 3.2 Outcome definitions

Outcome	Defined By	Data Source
Maternal Outcomes		
Elective c-section	ICD-10-AM delivery procedure code (O82.0, 1652000, 1652002)	NMDS, MAT
Emergency c-section	ICD-10-AM delivery procedure code (O82.1, 1652001, 1652003)	NMDS, MAT
Induction of labor	ICD-10-AM obstetric procedure code (9046500, 9046501, 9046502, 9046503, 9046504, 9046505)	NMDS, MAT
Severe perineal trauma	ICD-10-AM diagnosis codes (O70.2, O70.3, 1657300)	NMDS, MAT
Preeclampsia	ICD-10-AM diagnosis codes (O14.X)	NMDS
Infant Outcomes		
Birthweight	Birthweight in grams	MAT
Macrosomia (birthweight)	Birthweight in grams (>4,000g)	MAT
Large for gestational age	Birthweight in grams (> 90th percentile)	MAT
Neonatal hypoglycemia	ICD-10-AM diagnosis codes (P70, P70.0, P70.1, P70.3, P70.4)	NMDS
Respiratory distress	ICD-10-AM diagnosis codes (P22, P22.0, P22.1, P22.8, P22.9)	MAT/NMDS
Shoulder dystocia	ICD-10-AM diagnosis codes (O66, O66.0)	NMDS
Preterm birth	Delivery <37 completed weeks of gestation	MAT
Stillbirth	Fetal death after study entry (24 weeks)	Mortality Collection
NICU admission	Health Specialty codes (P42, P43)	NMDS
Child Outcomes		
Child growth (height, weight, body mass index)	Height and weight measurements	B4 School Check
Growth percentiles	Height and weight measurements, compared with 2006 World Health Organization reference standard	B4 School Check, WHO
Growth z-scores	Height and weight measurements, compared with 2006 World Health Organization reference standard	B4 School Check, WHO
SDQ[d]-Parent Difficulties Score	Summed scores for emotional problems, conduct problems, hyperactivity/inattention, and peer relationship problems parent-reported subscales	B4 School Check
SDQ-Teacher Difficulties Score	Summed scores for emotional problems, conduct problems, hyperactivity/inattention, and peer relationship problems teacher-reported subscales	B4 School Check
Prosocial behavior (parent-reported)	Subscale of SDQ, reported by parents	B4 School Check
Prosocial behavior (teacher-reported)	Subscale of SDQ, reported by parents	B4 School Check

CHAPTER 4 – MANUSCRIPT 1: COMPARATIVE EFFECTIVENESS OF METFORMIN VERSUS INSULIN FOR INITIAL TREATMENT OF GESTATIONAL DIABETES: MATERNAL AND NEONATAL OUTCOMES

4.1 Introduction

Gestational diabetes mellitus (GDM) is characterized by relative insulin deficiency that leads to maternal hyperglycemia (2). Worldwide, an estimated 1 in 7 births are affected by GDM (2). Identification and treatment of GDM improves perinatal outcomes (12, 14, 70). While many GDM women are able to achieve euglycemia with lifestyle interventions (diet and exercise), an estimated one-third of GDM women require pharmacologic treatment to maintain euglycemia (15, 71). When women fail to achieve euglycemia with lifestyle interventions, the optimal pharmacologic agent for GDM treatment is unclear.

Historically, injectable insulin has been the standard GDM treatment. Newer data suggest that oral hypoglycemic agents (OHAs) such as metformin may be an acceptable alternative to insulin to treat GDM. The evidence for the comparative effectiveness of OHAs versus insulin, however, is limited and inconclusive (41, 46, 48-52, 54-56, 72). The most recent Cochrane review could not recommend nor refute use of OHAs for GDM treatment based on available evidence (110).

To date, there have been no large observational studies evaluating the comparative effectiveness of metformin versus insulin. GDM treatment with metformin has been endorsed by the numerous professional and governmental organizations (7, 10, 62), yet guidelines caution that the safety of metformin use during pregnancy is not yet established and emphasize the need for more information on the effects of OHA for GDM treatment. To address this, the objective of this study was to estimate the comparative effectiveness of

metformin versus insulin in a large cohort study using routinely collected health care data from New Zealand.

4.2 Methods

4.2.1 Study population

The New Zealand Ministry of Health maintains 20 District Health Boards for the purpose of funding, planning, and providing health services. Because health services in New Zealand are funded by the government, eligible persons receive free inpatient and outpatient public hospital services, including pregnancy services and subsidies on prescription items. Nationally, approximately 14 repositories of health information exist and are linkable through a National Health Identifier (NHI), including the Pharmaceutical Claims Data Mart (PHARMS) and the National Minimum Dataset (NMDS) for hospitalizations (111). We identified a retrospective cohort of pregnancies with deliveries occurring in New Zealand between 2005 and 2015. Pregnancies were identified using the National Maternity Collection (MAT) which provides information on primary maternity services from the first midwife encounter through the postpartum period. This includes information from midwife encounters and hospitalizations. Mothers and infants are linked via a unique pregnancy key. Because stillbirths are not reliably reported in MAT, we linked available mortality data from years 2005 to 2013 to more accurately classify infant birth outcomes.

We estimated the start of pregnancy using gestational age as reported in MAT. Gestational age in MAT is reported as either the duration of completed weeks between a woman's self-reported last menstrual period (LMP) date and her delivery date, or derived from a clinical assessment during pregnancy or at birth. For those missing gestational age but having a recorded date of LMP, we estimated gestational age by calculating the number of completed weeks from the LMP to the delivery date. If both LMP and gestational age were missing, we removed these deliveries from our cohort.

We used the estimated start of pregnancy to define our baseline period for covariates. We examined the healthcare history for each pregnancy using data from hospitalizations, pharmaceutical claims and maternity care from 6 months prior to the estimated start of pregnancy through the start of the GDM screening and diagnosis window of time, 24-28 weeks gestation. To augment missing data on demographic characteristics, we linked the National Health Index (NHI) dataset to MAT. The NHI contains basic demographic information on NZ residents, including date of birth, residence, ethnicity, resident status, and socioeconomic status via the New Zealand Deprivation Index (NZDep13). The NZDep13 is an index of social deprivation that ranges from 1 to 10, where 1 represents the areas with the least deprived scores and 10 the areas with the most deprived scores (107). These data are updated on individuals at the time of a hospital encounter. Infants are registered with the NHI at birth.

In New Zealand, according to national guidelines, women are first screened for gestational diabetes with the "post-polycose" glucose tolerance test (10, 112). If the value of this test is ≥7.8 mmol/L they will undergo the oral glucose tolerance test and diagnosed with GDM if the fasting test result is ≥5.5mmol/L or the 2-hour test result is ≥9.0mmol/L. Following diagnosis of GDM, New Zealand women are provided with a blood glucose meter for home-based glucose testing and are given a prescription for testing strips (personal communication, Lesley MacLennan, Counties Manukau District Health Board). Because these prescriptions are subsidized by the NZ government, claims are filed for payment to the NZ Pharmaceutical Management Agency (PHARMAC) and recorded in the PHARMS collection. We considered a claim for prescription blood glucose test strips to be evidence of a GDM diagnosis in each pregnancy if the first prescription was filled after the start of the GDM screening period, 24 weeks gestation. If a prescription for test strips was filled anytime during the lookback period (6 months prior to pregnancy through 24 weeks gestation), we considered this to be evidence of prior type 2 diabetes and excluded these pregnancies. We

used a shorter lookback period (3 months prior to pregnancy) for women with closely spaced pregnancies. This was done to avoid excluding pregnancies immediately following a pregnancy complicated by GDM and beginning less than 6 months after the prior delivery. We also examined claims for antidiabetic medications (insulin, metformin, glibenclamide, glipizide, glicazide, pioglitazone, and acarbose) and excluded women who had a prescription claim for any antidiabetic prior to 24 weeks gestation.

Pharmaceutical data from the PHARMS are reliably available from 2005 onwards. Thus, in order to ensure that we observed the full 6-month lookback history of prescription use for each pregnancy, we restricted our cohort to pregnancies with an estimated start date of June 30, 2005 or later. We further restricted our cohort to pregnancies with a duration of 20 weeks or greater, and to infants weighing at least 400 grams. We additionally excluded pregnancies if the mother was younger than 15 years old at delivery, or older than 45 years old. Finally, we excluded pregnancies with missing information on gestational age. After identifying pregnancies with GDM, we identified prescriptions filled after or just before (one week or less) the first prescription for a blood glucose test strip. Our cohort was then narrowed down to women who initiated metformin, insulin, or both on a given date (Figure 1).

4.2.2 Exposure definition

For women who received both metformin and insulin during their pregnancy for the first time after filling a test strip prescription, we defined "combination therapy" as a dispensing for both medications within 7 days of the first claim. We treated combination therapy co-initiators as a separate group, as the need for both medications may have indicated an additional degree of GDM severity. For our primary analyses, we restricted our study population to pregnancies initially treated with metformin or insulin monotherapy. A woman who later added another therapy (>7 days after the first claim) remained in the cohort and analyzed as part of the treatment group with which she started.

4.2.3 Outcome definitions

Maternal outcomes (elective and emergency caesarean-sections, induction, preeclampsia and severe perineal trauma) were identified from the MAT dataset and derived from ICD-10-AM diagnosis codes from the NMDS collection of hospitalization events. For infants, outcomes of interest from the MAT dataset included birthweight and preterm birth. For delivery event outcomes not reported in MAT (neonatal hypoglycemia, shoulder dystocia, and respiratory distress) we used ICD-10-AM diagnosis codes recorded in NMDS during the delivery event in the hospital. We defined neonatal intensive care unit (NICU) admission by a hospitalization event in NMDS hospitalization event with a health specialty code for a Level II or Level III neonatal intensive care unit. A small proportion (<1%) of mothers and infants were missing records in hospitalization data, and were excluded from analyses comparing outcomes that were not available from MAT. Information on stillbirths were derived from the both the MAT and Mortality Data Collections and used as an outcome in our analysis. We estimated sex-, gestational age-, and ethnicity-specific birthweight percentiles from a sample of New Zealand infants born between 2001 and 2015 and recorded in MAT (N=850,506) to estimate risk of large-for-gestational-age.

4.2.4 Covariates

We identified potential confounders of the relationship between GDM treatment and infant birthweight from the relevant literature and based on clinical knowledge. We focused on infant birthweight as the outcome most directly-related to the effectiveness of these medications, and a contributing factor to other adverse outcomes such as macrosomia, severe perineal trauma or shoulder dystocia. Potential confounders included maternal age, maternal ethnicity, maternal socioeconomic status as determined by the NZ Deprivation Index, maternal BMI, maternal smoking status, prior history of GDM (treated and untreated), and timing of GDM diagnosis and treatment. We also examined claims for laboratory tests

from community laboratories to determine which screening and diagnostic tests were commonly received in our cohort.

4.2.5 Statistical analysis

We used multiple imputation using chained equations to estimate values of BMI, smoking status, and parity for pregnancies missing this information (108). BMI was measured at the first encounter with a Lead Maternity Carer in New Zealand, which means some women may have had their BMI recorded late in pregnancy. We reassigned extreme BMI values (<14 and >72) to missing. We then used multiple imputation to estimate BMI for pregnancies with missing data, including reassigned extreme values. We similarly imputed values for smoking status (yes/no), maternal ethnicity, and parity (0, 1, 2, or 3 or more prior births) at the time of the first LMC encounter.

To control for identified covariates, we estimated propensity scores for each pregnancy, calculated as the estimated probability of treatment with metformin versus insulin for each pregnancy conditional on measured covariates using logistic regression. We then used these values as inverse probability of treatment weights (IPTW), calculated as the inverse of probability of receiving the treatment actually received. Assuming no unmeasured confounders, estimated treatment effects from an IPTW analysis can be interpreted as the average treatment effect in the study population (113). We assessed balance of covariates pre- and post-weighting by calculating standardized mean differences for individual covariates for comparison and judged balance to be adequate if standardized mean differences post-weighting were less than 0.05. To measure the association between treatment with metformin versus insulin and the maternal and infant outcomes listed above, we estimated risk ratios and risk differences with 95% confidence intervals. Because some women contributed multiple pregnancies to our study cohort, we used generalized estimating equations for repeated measures to control for within-woman correlations. These models assumed an independent correlation structure, as is recommended with models

weighted using IPTW (109). We also included geographic region as a covariate in outcome models.

We conducted sensitivity analyses to estimate the comparative effectiveness of metformin versus insulin in key subgroups. We estimated unadjusted and weighted risk ratios and risk differences stratified by ethnicity. We conducted a sensitivity analysis restricting deliveries to those occurring after 2008, following the publication of the landmark MiG trial. Finally, we conducted an analysis of infant outcomes stratified by sex to explore a possible biological interaction between treatment and sex. This study was approved by the New Zealand Health and Disability Ethics Committee (Reference 16/NTA/66) and by the University of North Carolina at Chapel Hill Institutional Review Board (16-1121).

4.3 Results

We identified 679,125 deliveries occurring in New Zealand and recorded in MAT between 2005 and 2015 (Figure 4.1). Among these deliveries, 669,174 were singleton births. We were able to link mothers and infants in 662,290 pregnancies total. We eliminated pregnancies that were estimated to have begun earlier than June 30, 2005 (n=71,554), those missing gestational age (n=1,490), mothers aged less than 15 or over 45 years old (n=988), infants with a birthweight less than 400g (n=219), and infants born prior to 20 weeks gestation (n=105). From the 588,034 pregnancies meeting the inclusion criteria, we identified 22,060 (3.8%) with a claim for diabetic testing strips prior to delivery. We further excluded pregnancies with evidence of diabetic test strip fills prior to 24 weeks gestation (n=6,580), as this may be more indicative of overt diabetes than incident GDM. Of the 15,480 pregnancies assumed to have incident GDM based on diabetic test strip fill timing, 8,249 (53.3%) initiated a prescription for metformin or insulin after or shortly prior to obtaining test strips. We identified 981 pregnancies with metformin and insulin initiated concomitantly (separate fills within 7 days), 3,818 pregnancies with initiated metformin, and 3,450 pregnancies with initiated insulin.

Women who initiated metformin or insulin monotherapy were similar with respect to maternal characteristics including age, smoking status, and BMI (Table 4.1). On average, women were approximately 32 years old and had an average BMI of 29 prior to treatment for GDM. We observed substantial differences in other characteristics such as geographic region, ethnicity, and NZDep13 scores. The majority of metformin-treated pregnancies (62%) were from women in the Northern region. The prevalence of insulin treatment was more spread out geographically, with 34% of insulin-treated pregnancies in the Northern region and 24% on the South Island. We also observed trends in use by ethnicity. With regard to ethnicity, over half (57%) of insulin-treated pregnancies were among women of European ethnicity, while the majority (61%) of metformin-treated pregnancies were among women who identified as Asian or Pacific Peoples ethnicity. Additionally, pregnancies treated with metformin were more likely to be among women of lower socioeconomic status according to the NZDep13. Approximately 9% of insulin-treated pregnancies were among women assigned a score of 10 (indicating most deprivation) while 19% of metformin-treated pregnancies belonged to women with that score. We found evidence of diabetic screening or diagnostic tests (Hba1c, fructosamine, oral glucose tolerance test or serum glucose test) in over 90% of pregnancies. Only approximately 50% of pregnancies had a laboratory claim for the screening post-polycose test, though over 80% received the diagnostic oral glucose tolerance test. Roughly 60% of pregnancies were registered for prenatal care with a Lead Maternity Carer (e.g., midwife, obstetrician, or general practitioner) in the first trimester. Approximately 12% of women who initiated metformin required supplemental insulin, while just 1% of insulin initiators required supplemental treatment with metformin.

We observed time-based trends in treatment over the course of the study period. Very few pregnancies (n=29, 6%) were treated with metformin pre-2008, compared with 444 insulin-treated pregnancies (94%). This changed dramatically in 2008, when the number of metformin-treated pregnancies more than doubled between 2008 (33%) and 2009 (56%).

Our study cohort contained 634 women diagnosed with GDM in a prior pregnancy according to our study definition, while 405 of these women also received metformin or insulin treatment during the affected prior pregnancy. History of diagnosed and treated GDM was more common among the insulin treatment group. Both treatment groups had a median gestational week of diagnosis of week 30 (IQR 28, 32), and the median gestational week of treatment was week 32 (IQR 30, 34). After weighting, we observed no substantial differences in measured covariates between treatment groups (Appendix 3).

Before and after adjustment with IPTW compared to insulin, metformin use was associated with a reduced risk of elective c-section delivery and induction (Table 4.2). We did not observe differences in the risk of emergency c-section, preeclampsia or severe perineal trauma between treatment groups. Among infants, we observed similar birthweights and no meaningful difference in the risk of macrosomia associated with metformin compared with insulin use. The risk of large-for-gestational-age and neonatal hypoglycemia were significantly reduced in metformin-treated pregnancies compared with insulin. Shoulder dystocia was rare overall, yielding imprecise estimates, though the magnitude of the estimates suggests no substantial differences in risk associated with metformin versus insulin use. Stillbirths were extremely rare (n=8) and estimates of the association of treatment with stillbirth are not reported. The risk of NICU admission after birth was less common among metformin-exposed infant, with an attenuated effect after adjusting (RD= -15.4, 95% CI -17.2, -13.6 pre-weighting, RD= -5.9, 95% CI -7.5, -4.3 post-weighting). We further examined hospital discharge dates for mothers and infants and observed that pregnancies treated with metformin were less likely to result in an extended hospitalization for the infant, defined as an infant being discharged after the mother, compared with insulin (RD= -2.3, 95% CI -3.4, -1.2). We also observed a decrease in risk of preterm birth (defined as fewer than 37 completed weeks of gestation) associated with metformin use compared with insulin (RD= -1.7, 95% CI -3.0, -0.3).

In our cohort, we observed minor differences in effect estimates by ethnicity in our cohort (Figure 4.2, Appendix 4). The risk difference estimates for the association of metformin versus insulin treatment and elective c-section varied from 1.7 (95% CI -3.4, 6.7) among mothers who identified as Pacific Peoples to -5.7 (95% CI -10.3, -1.2) among Māori mothers. The effect estimates for neonatal outcomes were largely similar in magnitude and direction among infants born to mothers who identified as European, Māori, or Pacific Peoples. With respect to maternal and neonatal outcomes, among Asian women, metformin and insulin appeared to be equivalent. There was a reduced risk of neonatal hypoglycemia associated with metformin use compared with insulin among women who identified as European, Māori, or Pacific Peoples, but not among Asian women. In our sensitivity analysis examining outcomes among deliveries occurring after 2008, we observed risk ratios and risk differences similar to the overall cohort (Appendix 5).

Based on the results of our sex-specific analyses, metformin treatment for GDM was associated with a greater reduction in risk for neonatal hypoglycemia compared to insulin among boys (RD= -6.3, 95% CI -9.0, -3.7) than among girls (RD= -2.0, 95% CI -4.8, 0.8) (Appendix 6). Similarly, metformin appeared to reduce the risk of large-for-gestational-age to a greater extent among boys (RD= -4.8, 95% CI -7.3, -2.3) than among girls (RD= -2.0, 95% CI -4.8, 0.8) compared to insulin. For other neonatal outcomes, risk difference estimates were similar in direction and magnitude for both boys and girls.

4.4 Discussion

In a population-based cohort of NZ pregnancies treated pharmacologically for GDM, we did not observe increased risks of adverse maternal or neonatal outcomes associated with metformin use compared with insulin. Initial treatment with metformin was associated with reduced risks of elective c-section, labor induction, infant macrosomia, neonatal hypoglycemia, and preterm birth. Our estimates for rare adverse outcomes such as preeclampsia and shoulder dystocia do not suggest major differences in risk of these events

44

among metformin and insulin users. While rare, our confidence intervals for these associations were more precise than the combined estimates from a recent meta-analysis (52) (Appendix 7). In stratified models, we observed variation in elective c-section estimates by ethnicity, as well as variation in the degree to which metformin use affected neonatal hypoglycemia risk. We also observed variation by infant sex in the risk difference estimates for the association of metformin versus insulin treatment and two outcomes, neonatal hypoglycemia and large-for-gestational-age.

Our findings are consistent with results from the MiG trial and from a previous observational study in New Zealand (41, 54). The most recent meta-analysis of trials comparing metformin to insulin also calculated a summary estimate showing reduced risk of neonatal hypoglycemia associated with metformin (52). Similarly to an observational study conducted in the Auckland region, we observed a reduced incidence of NICU admission was associated with metformin versus insulin (54).

We observed potential differences by infant sex in how effective metformin is compared with insulin for preventing neonatal hypoglycemia and being large-for-gestational-age. Emerging research suggests that there may be clinically important differences in how men and women respond to diabetic therapies (114, 115).However, there is no information on how diabetic therapies used during pregnancy may differentially affect male and female fetuses. Our data suggest that further research on sex-specific effects of different GDM pharmacotherapy is needed.

Our study has a number of strengths. We studied over 7,000 treated pregnancies using routinely-collected healthcare data from the New Zealand Ministry of Health. This allowed us to compare metformin versus insulin in key subgroups, including stratifying models by ethnicity and infant sex. Additionally, the MAT data collection offered a unique opportunity to link information collected as part of maternity care with pharmaceutical dispensing, hospitalizations, laboratory claims and demographic data. We were able to

measure potential confounders such as maternal BMI, parity, smoking status, and socioeconomic status. We were also able to restrict our cohort to women with GDM diagnosed after 24 weeks of pregnancy. This represents an important improvement over other pharmacoepidemiological studies of treatment in pregnant women, which often must assume gestational age based on delivery diagnostic and procedure codes, and may restrict their populations to live births only (57, 58).

The use of secondary data presents some limitations. The MAT data collection only contains information on BMI recorded during the first encounter with a Lead Maternity Carer, which may have occurred in the second or third trimester. This means that the BMI measurement in MAT may not match a woman's BMI prior to pregnancy. However, approximately 60% of women in our cohort registered with a Lead Maternity Carer in the first trimester during which weight changes are typically minimal. We used multiple imputation using chained equations to overcome this limitation. Other limitations include missing information on weight gain during pregnancy and infant laboratory test results, limiting our ability to examine gestational weight gain as a confounder or confirm neonatal hypoglycemia as an outcome. As a strength, our treatment groups were already well-balanced by the active comparator new user design (116, 117). This is demonstrated by how stable our estimates were pre- and post-adjustment.

Overall, we observed no substantial differences in effectiveness between metformin and insulin as treatment for GDM based on a wide array of clinically relevant maternal and child outcomes. Metformin appears to be associated with reduced risks of neonatal hypoglycemia and NICU admission, which has been previously shown in randomized trials. We observed no clinically significant differences in average birthweight between the two treatment groups. Our data, as well of that of two recent meta-analyses show that metformin may have positive effects on GDM-treated women and infants (52, 53). While reassuring,

there is still inconclusive evidence for long-term infant safety following in utero exposure to metformin. Longitudinal research on child health outcomes from children exposed to metformin in utero is still needed.

4.5 Tables and Figures

Figure 4.1 Study cohort formation


```
Deliveries in New Zealand, 2005-2015
N= 679,125
                                          →  Excluded pregnancies:
                                             Twins or multiples (N=9,951)

Singleton births in New Zealand, 2005-2015
N=669,174
                                          →  Excluded pregnancies:
                                             Pregnancies without linked infant (N=6,884)

Linked mother-infant pairs
N=662,290
                                          →  Excluded pregnancies:
                                             6-month prior to pregnancy lookback period occurring
                                             before January 1, 2005 (N=71,554)
                                             Missing gestational age (N=1,490)
                                             Maternal age <15 or >45 (N=988)
                                             Birthweight <400g (N=219)
                                             Gestational age <20 weeks (N=105)

Eligible pregnancies meeting inclusion criteria
N=588,034
                                          →  Excluded pregnancies:
                                             Pregnancies with no evidence of diabetic testing strip
                                             fills (N=565,974)

Pregnancies with fill claim for diabetic testing strips
N=22,060
                                          →  Excluded pregnancies:
                                             Pregnancies with diabetic test strip fills prior to 24 weeks
                                             gestation (N=6,580)

Pregnancies with fill claim for diabetic testing strips after 24 weeks
N=15,480
                                          →  Excluded pregnancies:
                                             Pregnancies with diabetic test strip fills and no
                                             antidiabetic prescription fills (N=6,486)
                                             Pregnancies with antidiabetic prescription fills >7 days
                                             prior to test strip fill date (N=745)

Pregnancies with antidiabetic prescription fill after OR <=7 days prior to
diabetic testing strip fill date
N=8,249

    Insulin           Metformin        Combination
    N= 3,450          N= 3,818         Therapy
                                       N= 981
```

48

Table 4.1 Initial demographic and clinical characteristics of New Zealand pregnancies complicated by gestational diabetes and treated with metformin or insulin, 2006-2015

Characteristic	Metformin N=3,818		Insulin N=3,450	
	N	%	N	%
Maternal age (Mean, SD)	31.9	5.4	32.4	5.4
Maternal age (Median, IQR)	33	29, 36	32	28, 36
BMI (kg/m^2) (Mean, SD)[a]	29.4	7.6	29.2	8.4
BMI (kg/m^2) (Median, IQR)[a]	28	23, 33	28	24, 33
Gestational Week of Diagnosis (Mean, SD)	30.5	2.9	30.1	2.9
Gestational Week of Diagnosis (Median, IQR)	30	28, 32	30	28, 32
Gestational Week of Treatment Initiation (Mean, SD)	32.0	2.9	31.6	2.9
Gestational Week of Treatment Initiation (Median, IQR)	32	30, 34	32	30, 34
Maternal Age Categories				
15 - 20	74	1.9	64	1.9
21 - 25	407	10.7	330	9.6
26 - 30	1038	27.2	834	24.2
31 - 35	1256	32.9	1180	34.2
36 - 40	835	21.9	844	24.5
41 - 45	208	5.5	198	5.7
Geographic Region				
Northern	2382	62.4	1205	34.9
Midland	380	10.0	715	20.7
Central	540	14.1	690	20.0
South Island	516	13.5	840	24.4
Delivery Year				
2006	20	0.5	174	5.0
2007	9	0.2	270	7.8
2008	143	3.8	289	8.4
2009	337	8.8	266	7.7
2010	412	10.8	249	7.2
2011	540	14.1	332	9.6
2012	536	14.0	355	10.3
2013	673	17.6	466	13.5
2014	683	17.9	596	17.3
2015	465	12.2	453	13.1
Ethnicity (Prioritized)[b]				
European	1581	41.41	1953	56.61
Māori	555	14.54	585	16.96
Pacific Peoples	854	22.37	462	13.39
Asian	1461	38.27	1002	29.04
Other	150	3.93	112	3.25
Ethnicity (Any, Yes)[b,c]				
European	1085	28.4	1468	42.6
Māori	555	14.5	585	17.0
Pacific Peoples	781	20.5	393	11.4
Asian	1279	33.5	909	26.4
Other	117	3.1	93	2.7
NZ Resident (Yes)[d]	3453	90.7	3207	93.2
NZ Deprivation Index				

Characteristic	Metformin N=3,818		Insulin N=3,450	
1 (least deprived)	164	4.3	261	7.6
2	235	6.2	252	7.3
3	292	7.7	263	7.6
4	276	7.2	298	8.6
5	302	7.9	339	9.8
6	359	9.4	335	9.7
7	413	10.8	422	12.2
8	407	10.7	469	13.6
9	652	17.1	491	14.2
10 (most deprived)	718	18.8	320	9.3
Parity[e]				
0 (primipara)	1194	36.4	1214	38.8
1	1139	34.7	1069	34.2
2	476	14.5	475	15.2
3 or more	474	14.4	371	11.9
Trimester of Registration[f]				
1st	2252	64.8	2132	64.8
2nd	1009	29.0	974	29.6
3rd	214	6.2	183	5.6
Smoker at First Visit (Yes)[g]	268	7.9	267	9.3
BMI Categories[a]				
Underweight (<18.5)	37	1.2	55	2.0
Normal (18.5 – 24.9)	940	29.5	824	30.0
Overweight (25 – 29.9)	878	27.6	741	27.0
Obese (30 +)	1330	41.8	1127	41.0
Prior GDM (Yes)	315	8.3	319	9.3
Prior Treated GDM (Yes)	192	5.0	213	6.2

[a]BMI values are recorded at the first prenatal care encounter. These values reflect BMI recorded prior to metformin or insulin initiation. BMI values were missing for 1,336 women (18.4%).
[b] Maternal ethnicity was missing for 3 women (0.1%). The category of "Other" ethnicity includes Middle Eastern, Latin American/Hispanic, African, and other unspecified ethnicities.
[c]This includes any mention of a given ethnicity reported in first, second, or third ethnicity codes. As multiple ethnicities can be reported, proportions do not add up to 1.
[d]New Zealand resident status was missing for 21 women (0.3%).
[e]Parity information was missing for 856 women (11.8%).
[f]Trimester of maternity care registration was missing for 317 women (6.9%).
[g]Smoking status was missing for 814 women (13.8%).

Table 4.2 Unadjusted and IPTW-adjusted risk ratios (RR), risk differences (RD) per 100, and 95% CI for maternal and neonatal outcomes comparing pregnancies treated with metformin or insulin for gestational diabetes in New Zealand, 2006-2015

	Event Counts (%)		Unadjusted		IPTW Adjusted[a]	
Maternal Outcome	Metformin N=3,818	Insulin N=3,450	RR (95% CI)	RD (95% CI)	RR (95% CI)	RD (95% CI)
Elective c-section	720 (18.9)	808 (23.4)	0.81 (0.74, 0.88)	-4.5 (-6.4, -2.6)	0.89 (0.81, 0.98)	-2.3 (-4.3, -0.3)
Emergency c-section	640 (18.6)	650 (17.0)	0.92 (0.83, 1.01)	-1.5 (-3.3 0.2)	0.92 (0.82, 1.02)	-1.5 (-3.4, 0.5)
Induction	1965 (51.5)	1884 (54.6)	0.94 (0.90, 0.98)	-3.2 (-5.5, -0.8)	0.92 (0.88, 0.97)	-4.3 (-6.8, -1.8)
Preeclampsia[b]	139 (3.6)	116 (3.5)	1.08 (0.85, 1.38)	0.3 (-0.6, 1.1)	1.07 (0.83, 1.39)	0.2 (-0.7, 1.1)
Severe perineal trauma[c]	74 (1.9)	65 (1.9)	1.03 (0.75, 1.43)	0.1 (-0.6, 0.7)	0.89 (0.63, 1.27)	-0.3 (-0.9, 0.4)
Infant Outcome						
Birthweight (grams), Mean (SD)[d,e]	3367.6 (538.6)	3390.8 (557.0)	-20.69 (-46.09, 4.71)		0.58 (-27.06, 28.23)	
Macrosomia[d,f]	405 (10.7)	415 (12.2)	0.88 (0.77, 1.00)	-1.5 (-3.0, 0.0)	0.91 (0.80, 1.06)	-1.0 (-2.6, 0.6)
Large-for-gestational-age[d,g]	549 (14.5)	653 (19.1)	0.77 (0.69, 0.85)	-4.4 (-6.2, -2.7)	0.80 (0.71, 0.89)	-3.7 (-5.5, -1.8)
Neonatal hypoglycemia[h]	543 (14.3)	754 (22.0)	0.65 (0.59, 0.72)	-7.7 (-9.5, -5.9)	0.74 (0.66, 0.82)	-5.0 (-6.9, -3.2)
Shoulder dystocia[b]	37 (1,0)	42 (1.2)	0.80 (0.51, 1.24)	-0.2 (-0.7, 0.2)	0.88 (0.55, 1.41)	-0.1 (-0.7, 0.4)
Respiratory distress[h]	194 (5.1)	244 (7.1)	0.72 (0.60, 0.86)	-2.0 (-3.1, -0.9)	0.87 (0.71, 1.06)	-1.0 (-2.2, 0.2)
NICU admission[h,i]	458 (12.1)	962 (28.1)	0.44 (0.40, 0.49)	-15.4 (-17.2, -13.6)	0.63 (0.57, 0.70)	-5.9 (-7.5, -4.3)
Extended postnatal hospitalization[b,h,j]	168 (4.4)	260 (7.6)	0.58 (0.48, 0.70)	-3.2 (-4.3, -2.1)	0.69 (0.56, 0.84)	-2.3 (-3.4, -1.2)
Preterm birth[k]	269 (7.1)	311 (9.0)	0.78 (0.67, 0.92)	-1.9 (-3.2, -0.7)	0.79 (0.67, 0.93)	-1.7 (-3.0, -0.3)

[a]Inverse probability of treatment weights (IPTW) were estimated using a propensity score model containing the following covariates: maternal age, maternal ethnicity, maternal NZ deprivation decile score, parity, BMI prior to prescription initiation, smoking status, timing of GDM diagnosis and treatment, history of GDM and history of GDM treatment. The IPTW adjusted model uses imputed values for parity, BMI, and smoking status. These models are also adjusted for geographic region.
[b]Hospitalization data for women were missing for 39 delivery events (18 metformin, 21 insulin).
[c]Severe perineal trauma is defined as third- or fourth-degree tear or a tear requiring surgical repair.
[d]Birthweight values were missing for 63 infants (29 metformin, 34 insulin).
[e]Linear regression was used to compare mean birthweight in infants treated initially with metformin versus insulin.
[f]Macrosomia is defined as birthweight greater than 4,000g.
[g]Large-for-gestational-age is defined as a birthweight greater than the 90th percentile. These percentiles were calculated using a sample of New Zealand infants born between 2001 and 2015 with available data on birthweight and gestational age (n=850,506). These percentiles are additionally sex- and ethnicity-specific.
[h]Hospitalization data for infants were missing for 39 delivery events (18 metformin, 21 insulin).
[i]NICU admission is defined as a hospitalization event with a health specialty code for a Level II or Level III neonatal intensive care unit.

[j]Extended postnatal hospitalization was defined as an infant being discharged from the hospital after a mother was discharged following delivery.
[k]Preterm birth is defined as gestational age <37 weeks.

Figure 4.2 Ethnicity-stratified risk difference estimates and 95% confidence intervals for the association between metformin versus insulin initiation and adverse maternal and neonatal outcomes

	RD (95% CI)
Elective c-section	
European	-4.7 (-8.2, -1.1)
Maori	-5.7 (-10.3, -1.2)
Pacific Peoples	1.7 (-3.4, 6.7)
Asian	-1.3 (-4.8, 2.2)
Emergency c-section	
European	0.5 (-2.6, 3.6)
Maori	-1.7 (-6.2, 2.9)
Pacific Peoples	-3.4 (-8.8, 1.9)
Asian	-1.8 (-5.2, 1.6)
Induction	
European	-4.2 (-8.3, 0.0)
Maori	0.2 (-5.7, 6.1)
Pacific Peoples	-8.3 (-14.8, -1.8)
Asian	-4.5 (-9.0, 0.1)
Large-for-gestational-age	
European	-4.4 (-7.4, -1.4)
Maori	-6.0 (-11.3, -0.7)
Pacific Peoples	-5.7 (-11.5, 0.0)
Asian	-1.0 (-3.8, 1.8)
Macrosomia	
European	0.1 (-2.5, 2.7)
Maori	-3.7 (-8.1, 0.6)
Pacific Peoples	-3.1 (-8.6, 2.5)
Asian	-1.2 (-2.8, 0.4)
Neonatal hypoglycemia	
European	-7.0 (-10.0, -3.9)
Maori	-8.2 (-12.9, -3.4)
Pacific Peoples	-7.8 (-13.1, -2.5)
Asian	0.2 (-2.8, 3.2)
NICU admission	
European	-6.7 (-9.5, -3.9)
Maori	-15.1 (-20.4, -9.8)
Pacific Peoples	-4.3 (-8.1, -0.5)
Asian	-4.1 (-6.6, -1.5)

-25   -12.5   0   12.5   25

CHAPTER 5 – MANUSCRIPT 2

5.1 Introduction

Gestational diabetes mellitus (GDM) is a common pregnancy complication, characterized by relative insulin deficiency and in most women, impaired insulin action, leading to hyperglycemia (2). An estimated 1 in 7 births are affected by GDM worldwide (2). While many women are able to control GDM through lifestyle changes, an estimated one-third of women with GDM require pharmacologic treatment to manage their condition (15, 118).

Oral hypoglycemic agents like metformin have emerged as a promising option for managing GDM (110). In randomized trials comparing maternal and infant outcomes at delivery, metformin has been shown to be an effective alternative to insulin, the historical standard treatment for GDM (22, 41, 46, 50). In the absence of long-term data on the safety of metformin use during pregnancy, concerns remain regarding the potential development effects that may only be evident later in childhood (7, 10, 62, 63, 71, 118).Unlike insulin, metformin has been shown to cross the placenta (16, 18, 19). This has caused concerns that metformin may impact fetal development in ways that insulin would not.

To date, researchers have examined the comparative safety of metformin versus insulin treatment during pregnancy on the offspring in children up to age 4 (43-45, 47, 119). Although these studies have suggested no clinical differences in growth or neurodevelopment among the offspring of women treated with metformin versus insulin for GDM, they were limited by small initial sample sizes and loss-to-follow-up from the original, randomized study population. Thus, patients and providers remain concerned about the long-term effects of metformin (120). To address the noted lack of long-term safety

information in children of women treated with metformin or insulin during pregnancy, we conducted a retrospective cohort study in a New Zealand (NZ) population of treated women linked with their child's growth and developmental data at age 4, taken from a pre-school readiness screening program provided as routine well-child care throughout NZ.

5.2 Methods

5.2.1 Study population

Health services in NZ are funded by the government, meaning eligible persons receive free inpatient and outpatient public hospital services, including pregnancy services and subsidies on prescription items. Nationally, approximately 14 repositories of health information exist and are linkable through a National Health Identifier (NHI) (73). We identified a cohort of pregnancies with deliveries occurring in NZ between 2005 and 2012. Pregnancies were identified using the National Maternity Collection (MAT) which provides information on primary maternity services from the first prenatal encounter through the postpartum period. Mothers and their children were linked via a unique pregnancy key. Gestational age in MAT was reported as either the duration of completed weeks between a woman's self-reported last menstrual period (LMP) date and her delivery date, or derived from a clinical assessment during pregnancy or at birth. For those missing gestational age but having a recorded LMP date, we estimated gestational age by calculating the number of completed weeks from the LMP date to the delivery date. We estimated the start of pregnancy using gestational age as reported in MAT.

We examined the healthcare history for each pregnancy using data from hospitalizations, pharmaceutical claims and maternity care from 6 months prior to the estimated start of pregnancy through the start of the GDM screening and diagnosis window of time, 24-28 weeks gestation. Additional demographic characteristics were obtained through linkage to the National Health Index, which contains basic demographic information on NZ residents, including date of birth, residence, ethnicity, resident status, and

socioeconomic status via the NZ Deprivation Index (NZDep13). The NZDep13 is an index of social deprivation that ranges from 1 to 10, where 1 represents the areas with the least deprived scores and 10 the areas with the most deprived scores (107). These data are updated on individuals at the time of a hospital encounter. Infants are registered with the NHI at birth.

5.2.2 Exposure definition

In NZ, women diagnosed with GDM are first screened with the "post-polycose" glucose tolerance test (10, 112). According to national guidelines, if the value of this test is ≥7.8 mmol/L, patients undergo the oral glucose tolerance test and are diagnosed with GDM if the fasting test result is ≥5.5mmol/L or the 2-hour test result is ≥9.0mmol/L. Following diagnosis of GDM, NZ women are provided with a blood glucose meter for home-based glucose testing and are given a prescription for testing strips (personal communication, Lesley MacLennan, Counties Manukau District Health Board). Because these prescriptions are subsidized by the NZ government, claims are filed for payment to the NZ Pharmaceutical Management Agency (PHARMAC) and recorded in the Pharmaceutical Claims Data Mart (PHARMS). We considered a claim for prescription blood glucose test strips to be evidence of a GDM diagnosis in each pregnancy if the first test strip claim appeared after the start of the GDM screening period, 24 weeks gestation. If a prescription for test strips was filled anytime between 6 months prior to pregnancy through 24 weeks gestation, we considered this to be evidence of prior type 1 or 2 diabetes and excluded these pregnancies. We used a shorter window (3 months prior to pregnancy) for women with closely spaced pregnancies. This was done to avoid excluding pregnancies immediately following a pregnancy complicated by GDM and beginning less than 6 months after the prior delivery. We also examined claims for antidiabetic medications (metformin, insulin, glibenclamide, glipizide, glicazide, pioglitazone, and acarbose) and excluded women who had a prescription claim for any antidiabetic prior to 24 weeks gestation.

Pharmaceutical data from the PHARMS are reliably available from 2005 onwards. Thus, in order to ensure that we observed the full 6-month lookback history of prescription use for each pregnancy, we restricted our cohort to pregnancies with an estimated start date of June 30, 2005 or later (Figure 5.1). We further restricted our cohort to pregnancies with a duration of 20 weeks or greater, and to infants weighing at least 400 grams. We excluded pregnancies if the mother was younger than 15 years old or older than 45 years old at delivery. Finally, we excluded pregnancies with missing information on gestational age. After identifying pregnancies with GDM, we identified prescriptions filled after or just before (one week or less) the first claim for a blood glucose test strip. We excluded women who initiated any prescription other than insulin or metformin. For women who received both metformin and insulin during their pregnancy for the first time after filling a test strip prescription, we defined "combination therapy" as a dispensing for both medications within 7 days of the first claim. For our analyses, we restricted our study population to pregnancies initially treated with metformin or insulin monotherapy.

5.2.3 Outcome definitions

The B4 School Check (B4SC) is a universal health and development screening program offered to all children in NZ at age 4 as part of routine well-child care, prior to school entry at age (121). This includes growth surveillance involving height and weight measurement and identification of behavioral problems using the Strengths and Difficulties Questionnaire (SDQ) for parents and teachers (SDQ-P and SDQ-T) (101, 102). We linked children with available results to the mothers we identified as initiating metformin or insulin treatment for GDM.

5.2.3.1 Growth outcomes

We used height and weight values from the B4SC to calculate child BMI. These values were also used to estimate z-scores and percentile-based outcomes for weight,

height, weight-for-height, and BMI. We used the 2006 World Health Organization (WHO)

reference standards to calculate z-scores and percentiles (100). We used the 85th

percentile as a cutpoint for defining overweight and obesity in children and applied this

cutpoint for creating dichotomous outcomes for each growth measurement. We also

examined the proportion of children with weight, height, weight-for-height and BMI z-scores

≥1 or ≥2. Additionally, the B4SC recommends referral for children with weight above the

97th percentile and BMI of 21 or over (99). We therefore used the 97th percentile as a

cutpoint for extreme growth measurements and examined the proportion of children with a

BMI measurement of 21 or over. Details on how practitioners are instructed to take weight

and height measurements during the B4SC can be found in the program's handbook (99).

5.2.3.2 Strengths and Difficulties Questionnaire

The B4 School Check offers screening to help identify children that may need help

with learning and development before entering school using the SDQ (102). The SDQ is

designed to assess the child's strengths in five subscales pertaining to different emotions or

behaviors: prosocial behavior, hyperactivity, emotional symptoms, conduct, and peer

problems (101, 102). In NZ, the SDQ was offered in two different formats, depending on

whether a parent or a teacher is administering the questionnaire. Parents and teachers

respond "not true", "somewhat true", or "certainly true" to each statement pertaining to the

child's behavior (e.g., "Considerate of other people's feelings", "Often loses temper").

Responses to these statements can be organized into the four subscales of hyperactivity,

emotional symptoms, conduct, and peer problems. Scores from these subscales are

summed to result in a Difficulties Score which can range from 0 to 40, with higher scores

indicating concern and possible need for referral (99, 105). The cut-off for a "concerning"

score is typically 17 for the parent-completed SDQ and 16 for the teacher-completed SDQ;

however, a review of the SDQ in the NZ population found that a lower cut-off is more

appropriate (14 for SDQ-P, 11 for SDQ-T) (99, 105). An additional prosocial behavior scale

is not included in the total Difficulties Score and is scored so that an absence of prosocial behavior (e.g. helping, sharing) receives a lower score.

5.2.4 Maternal and infant characteristics

We identified potential confounders of the relationship between GDM treatment and child growth and development from the relevant literature and with clinical insight. In our study, we included maternal age, maternal ethnicity, maternal socioeconomic status as determined by the NZ Deprivation Index (NZDep13), maternal BMI, maternal smoking status, prior history of GDM (treated and untreated), and timing of GDM diagnosis and treatment as potential confounders. We used maternity care and hospitalization data to examine children's history of infant outcomes that have been associated with child obesity or neurodevelopmental difficulties, such as large-for-gestational-age and neonatal hypoglycemia (39, 40, 122). However, because these outcomes occurred after the mother initiated treatment with metformin or insulin, we limited this examination to a descriptive analysis, rather than including infant health outcomes as covariates.

Multiple imputation using chained equations was used to estimate missing covariate information (108). BMI was measured at the first prenatal care encounter in NZ, which means some women may have had their BMI recorded late in pregnancy. We reassigned extreme BMI values (<14 and >72) to missing. We then used multiple imputation to estimate BMI for pregnancies with missing data, including reassigned extreme values. We similarly imputed values for maternal ethnicity, smoking status (yes/no) and parity (0, 1, 2, or 3 or more prior births) at the time of the first LMC encounter.

5.2.5 Statistical analysis

To control for identified covariates, we estimated propensity scores for each pregnancy, calculated as the estimated probability of treatment with metformin based on measured covariates using logistic regression. We then used these values to estimate

inverse probability of treatment weights (IPTW), calculated as the inverse of the probability of receiving the treatment actually received (113). We assessed balance of covariates pre- and post-weighting by calculating standardized mean differences for individual covariates for comparison.

For continuous outcomes (e.g. weight z-scores), we used linear regression to compare outcomes between children of metformin- and insulin-treated pregnancies. Log-binomial regression was used to estimate risk ratios and 95% confidence intervals for binary outcomes (e.g. weight above the 97th percentile for age). Because some women contributed multiple pregnancies to our study cohort, we used generalized estimating equations for repeated measures to control for within-woman correlations. We also included geographic region as a covariate in outcome models.

We conducted a sensitivity analysis to estimate the association of initiating metformin versus insulin with child growth and development stratified by child sex to explore a potential interaction observed in other clinical settings (114, 115). This study was approved by the NZ Health and Disability Ethics Committee (Reference 16/NTA/66) and by the University of North Carolina at Chapel Hill Institutional Review Board (16-1121).

5.3 Results

We identified 501,672 deliveries occurring in New Zealand and recorded in MAT between 2005 and 2012 (Figure 5.1). We excluded deliveries of multiple gestations (n=7,403) and further excluded pregnancies that were estimated to begin before June 30, 2005 (n=42,758), pregnancies missing gestational age (n=1,204), mothers aged less than 15 or greater than 45 years old (n=749), infants with a birthweight less than 400g (n=141), and infants born prior to 20 weeks gestation (n=79). Among the pregnancies meeting the inclusion criteria, we identified 13,023 (3.1%) with a claim for diabetic testing strips prior to delivery. We excluded 4,070 deliveries for having a diabetic test strip fill prior to 24 weeks of gestation, suggesting overt rather than gestational diabetes. Of the 8,953 pregnancies

assumed to have incident GDM based on diabetic test strip fill timing, 4,318 (48.2%) initiated

a prescription for metformin or insulin after or shortly prior to obtaining test strips. We

identified 386 pregnancies with metformin and insulin initiated concomitantly (separate fills

within 7 days). We included all pregnancies resulting in liveborn children that could be linked

to the B4SC dataset. The final cohort included 1,996 pregnancies with initiated metformin,

and 1,932 pregnancies with initiated insulin.

Overall, women treated with either metformin or insulin were similar with respect to

age, BMI, and timing of GDM diagnosis and treatment initiation (Table 5.1). We observed

strong trends by region and delivery year, however. Over half (66.7%) of metformin use was

prescribed to women residing in the Northern region, which includes Auckland. Metformin

use was infrequent (n=29) prior to 2008. We also observed differences in treatment initiation

by ethnicity and NZ Deprivation Index Deciles, a measure of the relative socioeconomic

deprivation of an area. After inverse probability of treatment weighting, treatment groups

were well-balanced on measured covariates (Appendix 8).

The children of women treated with metformin were less likely to have experienced

neonatal hypoglycemia (14%) at birth compared with the offspring of women treated with

insulin (23%). Children from both treatment groups were similar with respect to average

gestational age at birth (38.2 weeks for metformin-treated pregnancies, 38.0 for insulin-

treated pregnancies) and average birthweight (3362g for metformin-treated pregnancies,

3437g for insulin-treated pregnancies).

When comparing growth outcomes (weight, height, weight-for-height and BMI)

between children of women treated with metformin or insulin, we did not observe significant

differences in adjusted models (Table 5.2). The average weight z-score for both groups was

0.7 and we observed no difference in weight z-scores in our adjusted linear regression

model (β= -0.028, 95% CI -0.125, 0.069). We compared the proportion of children with

weights, heights, weight-for-length and BMI values at or above the 85th percentile and found

no significant differences in risk of these outcomes (Table 5.3). Children of metformin-treated mothers were not significantly more likely to be ≥85th percentile for weight (adjusted risk ratio [aRR]=1.04, 95% CI 0.93 1.16) or weight-for-height (aRR=0.92, 95% CI 0.83, 1.02) than children of insulin-treated mothers. Overall, we consistently observed adjusted risk ratios below the null for weight, weight-for-length, and BMI outcomes, though the confidence intervals crossed the null. For height, we observed risk ratios to suggest that children of metformin-treated pregnancies may be taller on average than those of insulin-treated pregnancies, but after adjustment, these values were neither clinically nor statistically significant.

Children born to metformin- and insulin-treated pregnancies had similar average SDQ-P and SDQ-T scores (Table 5.2). Less than half of the children in our study had a completed SDQ-T (n=1,681, 42.8%), while approximately 80% had a completed SDQ-P. SDQ-P and SDQ-T Prosocial Behaviour Scores were, on average, similar in both groups. We also observed similar proportions of concerning SDQ-P and SDQ-T Difficulties Scores between treatment groups. After weighting, children born to metformin-treated women appeared to be at an increased risk of having a concerning SDQ-P score (aRR=1.13, 95% CI 0.88, 1.46). However, we did not observe meaningful differences in average SDQ-P scores in our linear regression models after adjustment (β= -0.015, 95% CI -0.147, 0.118). In sex-specific models (Table 5.4), we observed differences in adjusted models for growth outcomes between boys and girls. The risk of having an extreme weight-for-length z-score (≥2) was lower among metformin-exposed boys, but higher among metformin-exposed girls compared with those exposed to insulin (aRR for boys=0.77, 95% CI 0.59, 1.00; aRR for girls=1.08, 95% CI 0.78, 1.49) (Appendix 10). However, the corresponding p-value for interaction (p=0.3653) indicated that these differences were not statistically significant (Table 5.4). Having a concerning Difficulties Score on the SDQ-P or SDQ-T was overall more common among boys than girls. We observed no association between metformin

versus insulin treatment and concerning SDQ-P scores for girls (aRR=1.00, 95% CI 0.63, 1.58), while the risk of a concerning SDQ-P score for metformin-exposed boys was higher than for insulin-exposed boys (aRR=1.21, 95% CI 0.89, 1.64) (Table 5.4).

5.4 Discussion

In a population-based cohort of children born to women treated with metformin or insulin for GDM in NZ, we did not observe meaningful differences in child growth or developmental outcomes at age 4. Children whose mothers were treated initially with metformin versus insulin to manage GDM were similar with respect to weight, height, and BMI and we did not observe statistically or clinically significant differences between them after adjusting for measured covariates. We observed variation in model estimates by child sex, but there was not a statistically significant interaction between treatment and sex. As expected, the average height, weight, and BMI values for the population of children born to women treated pharmacologically for GDM were higher than average values for the entire NZ population (98). Studies have shown that maternal hyperglycemia is associated with obesity in children (32), though these results have been inconsistent and may be partially explained by maternal prepregnancy BMI (30, 31).

Our findings are consistent with prior studies focused on child health outcomes at younger ages. A follow-up study of offspring born to women enrolled in the MiG trial (N=323) found that children who had been exposed to metformin in utero did not differ from children exposed to insulin with regard to weight, height, or abdominal fat at two year of age (43). The results of this study suggested that treatment with metformin may be associated with the development of less visceral fat in offspring as compared with insulin, which we were unable to ascertain in this observational study. At 18-months post-randomization, Ijäs et al reported increased risks of having a height or weight at or above the 95th percentile for metformin-exposed children, but the sample was small (n=93.) Thus, the estimates were

imprecise (e.g., weight ≥ 95th percentile RR=5.33, 95% CI 0.65, 43.91) (47), making it

impossible to rule out a large range of possible effect sizes.

Previous follow-up studies have also found no differences between children born to

mothers treated with metformin versus insulin with respect to neurodevelopmental

outcomes. A subsequent follow-up of children born to MiG trial participants at age 2 found

no meaningful differences in Bayley Scales of Infant Development Scores by treatment

status, though they did observe differences by country (New Zealand versus Australia) (45).

Similarly, Terrti et al did not find differences in average Bayley-III scores between children at

age 2 born to mothers participating in a trial comparing metformin and insulin treatment for

GDM in Finland (119). Ijäs et al also concluded that metformin was not associated with

developmental delay compared with insulin when comparing developmental milestones like

standing, walking, and speech at 18 months (47).

This study had some limitations, primarily stemming from missing data. Outcome

measurements were missing for approximately 20% of children. We were also missing

important covariate data for mothers, including BMI, smoking status, and parity. Using

multiple imputation, we were able to augment the covariate data, but did not impute

outcomes. Despite the high proportion of missing data, our study was able to examine

outcomes in over 3,000 children, resulting in the largest study to date evaluating long-term

outcomes in children of metformin- and insulin-treated mothers. This results in more precise

estimates that reduce the uncertainty surrounding metformin's impact on child development

when used to treat GDM.

It is possible that there may be residual confounding related to disease severity due

to the observational nature of this study. However, the mothers included in this study were

similar with respect to important indicators of GDM severity, such as BMI and timing of

diagnosis. Additionally, the use of a comparative effective design helps to reduce residual

confounding by including exclusively women with the indication of treatment of GDM in our cohort and is a major methodological strength (116, 117).

This study is the first to examine long-term outcomes beyond early childhood in children born to women treated with metformin versus insulin for GDM. We observed no meaningful differences in growth or behavioral and emotional development between treatment groups. This may provide clinicians and patients with important information about the comparative safety of metformin versus insulin when making treatment decisions during pregnancy.

5.5 Tables and Figures

Figure 5.1 Study cohort formation

Figure 5.1 Study cohort formation

Flow	Excluded
Deliveries in New Zealand, 2005-2012 N= 501,672	Excluded pregnancies: Twins or multiples (N=7,403)
Singleton births in New Zealand, 2005-2012 N=494,269	Excluded pregnancies: Pregnancies without linked infant (N=6,884)
Linked mother-infant pairs N=488,944	Excluded pregnancies: 6-month prior to pregnancy lookback period occurring before January 1, 2005 (N=42,758) Missing gestational age (N=1,204) Maternal age <15 or >45 (N=749) Birthweight <400g (N=141) Gestational age <20 weeks (N=79)
Eligible pregnancies meeting inclusion criteria N=415,344	Excluded pregnancies: Pregnancies with no evidence of diabetic testing strip fills (N=402,321)
Pregnancies with fill claim for diabetic testing strips N=13,023	Excluded pregnancies: Pregnancies with diabetic test strip fills prior to 24 weeks gestation (N=4,070)
Pregnancies with fill claim for diabetic testing strips after 24 weeks N=8,953	Excluded pregnancies: Pregnancies with diabetic test strip fills and no antidiabetic prescription fills (N=4,180) Pregnancies with antidiabetic prescription fills >7 days prior to test strip fill date (N=455)
Pregnancies with antidiabetic prescription fill after OR <=7 days prior to diabetic testing strip fill date N=4,318	
Pregnancies with metformin, insulin, or combination treatment N=4,318	Excluded pregnancies: Combination therapy (N=386) Stillbirths (N=4)
Linked to B4 School Check Database, 2009 - 2017 N=3,928	

| Insulin N= 1,932 | Metformin N= 1,996 |

Table 5.1 Demographic and clinical characteristics of New Zealand mothers treated pharmacologically for gestational diabetes with metformin or insulin, 2005-2012

	Metformin N=1,996		Insulin N=1,932	
	N	%	N	%
Age (Mean, SD)	32.07	5.5	32.44	5.6
BMI (kg/m^2) (Mean, SD)	29.58	7.7	29.54	9.4
Gestational week of diagnosis (Mean, SD)	30.6	2.9	30.25	2.8
Gestational week of treatment (Mean, SD)	31.96	3.0	31.76	2.9
Age Categories				
15 - 20	37	1.9	42	2.2
21 - 25	209	10.5	195	10.1
26 - 30	536	26.9	441	22.8
31 - 35	633	31.7	638	33.0
36 - 40	463	23.2	499	25.8
41 - 45	118	5.9	117	6.1
Geographic Region				
Northern	1331	66.7	548	28.4
Midland	139	7.0	456	23.6
Central	274	13.7	418	21.6
South Island	252	12.6	510	26.4
Maternal Ethnicity (Prioritised)[a]				
European	576	28.9	867	44.9
Maori	290	14.5	360	18.7
Pacific Peoples	435	21.8	243	12.6
Asian	639	32.0	408	21.1
Other	55	2.8	52	2.7
Maternal Ethnicity (Any, Yes)[a,b]				
European	843	42.2	1143	59.2
Maori	290	14.5	360	18.6
Pacific Peoples	474	23.8	274	14.2
Asian	736	36.9	460	23.8
Other	75	3.8	62	3.2
NZ Deprivation Index Deciles				
1 (least deprived)	89	4.5	129	6.7
2	111	5.6	132	6.8
3	148	7.4	133	6.9
4	136	6.8	166	8.6
5	154	7.7	168	8.7
6	195	9.8	188	9.7
7	224	11.2	239	12.4
8	207	10.4	271	14.0
9	362	18.1	293	15.2
10 (most deprived)	370	18.5	213	11.0
Delivery Year				
2006	20	1.0	174	9.0
2007	9	0.5	270	14.0
2008	143	7.2	287	14.9
2009	336	16.8	266	13.8
2010	412	20.6	249	12.9
2011	540	27.1	332	17.2

	Metformin N=1,996		Insulin N=1,932	
	N	%	N	%
2012	536	26.9	354	18.3
NZ Resident (Yes)[c]	1797	90.5	1794	93.2
Parity[d]				
0 (primipara)	627	38.0	645	38.3
1	536	32.5	545	32.3
2	256	15.5	273	16.2
3 or more	231	14.0	222	13.2
Trimester of Maternity Care Registration[e]				
1st	1056	58.8	980	54.0
2nd	583	32.5	710	39.1
3rd	157	8.7	125	6.9
Smoker at First Visit (Yes)[f]	114	6.6	151	10.9
BMI (kg/m^2) Categories[g]				
Underweight (<18.5)	16	0.8	25	1.3
Normal (18.5 - 24.9)	442	22.1	376	19.5
Overweight (25 - 29.9)	431	21.6	340	17.6
Obese (30 +)	640	32.1	553	28.6
Prior GDM (Yes)	112	5.6	135	7.0
Prior Treated GDM (Yes)	56	2.8	84	4.4

[a]Maternal ethnicity was missing for 3 women (0.1%).

[b] This includes any mention of a given ethnicity reported in first, second, or third ethnicity codes. As multiple ethnicities can be reported, proportions do not add up to 1. The category of "Other" ethnicity includes Middle Eastern, Latin American/Hispanic, African, and other unspecified ethnicities.

[c]New Zealand resident status was missing for 18 women (0.5%).

[d]Parity information was missing for 593 women (15.1%).

[e]Trimester of maternity care registration was missing for 317 women (8.1%).

[f]Smoking status was missing for 814 women (20.7%).

[g]BMI values are recorded at the first prenatal care encounter. These values reflect BMI recorded prior to metformin or insulin initiation. BMI values were missing for 1,105 women (28.1%).

Table 5.2 Unadjusted and adjusted estimates from linear regression models comparing child health outcomes between children exposed in utero to metformin versus insulin for treatment of gestational diabetes, 2005-2012

	Metformin N=1,996 Mean (SD)	Insulin N=1,932 Mean (SD)	Unadjusted β Estimate (95% CI)	Adjusted[a] β Estimate (95% CI)
Child Growth Outcomes				
Weight z-score[b,c]	0.7 (1.1)	0.7 (1.2)	0.012 (-0.074, 0.097)	-0.028 (-0.125, 0.069)
Height z-score[c,d]	0.4 (1.1)	0.3 (1.1)	0.119 (0.044, 0.194)	0.052 (-0.034, 0.138)
Weight-for-Length[e] z-score [c]	0.7 (1.3)	0.9 (1.3)	-0.107 (-0.199, -0.016)	-0.095 (-0.199, 0.010)
BMI z-score[c,e]	0.8 (1.4)	0.9 (1.3)	-0.097 (-0.192, -0.002)	-0.095 (-0.202, 0.012)
Strengths and Difficulties Questionnaire				
SDQ-P Difficulties Score [f,g]	6.7 (4.7)	6.5 (4.9)	0.202 (-0.136 0.540)	0.119 (-0.271, 0.509)
SDQ-T Difficulties Score [f,h]	4.0 (4.4)	3.7 (4.5)	0.347 (-0.080, 0.774)	0.252 (-0.235, 0.740)
SDQ-P Prosocial Behaviour Score [g]	8.5 (1.7)	8.5 (1.7)	-0.027 (-0.146, 0.092)	-0.015 (-0.147, 0.118)
SDQ-T Prosocial Behaviour Score [h]	8.0 (2.3)	8.2 (2.3)	-0.252 (-0.472, -0.032)	-0.184 (-0.426, 0.058)

[a]Inverse probability of treatment weights (IPTW) were estimated using a propensity score model containing the following covariates: maternal age, maternal ethnicity, maternal NZ deprivation decile score, parity, BMI prior to prescription initiation, smoking status, timing of GDM diagnosis and treatment, history of GDM and history of GDM treatment. The IPTW adjusted model uses imputed values for parity, BMI, and smoking status. These models are also adjusted for geographic region.
[b]Weight was avaliable for 3,156 children (80.4%). Weight measurements were missing for 21.5% of insulin-exposed children and 17.9% of metformin exposed children.
[c]Z-scores and percentiles were calculated based on a WHO 2006 reference standard.
[d]Height was available for 3,154 children (80.3%). Height measurements were missing for 21.5% of insulin-exposed children and 17.9% of metformin exposed children.
[e]Weight for length and BMI measurements were calculated for children with non-missing weight and height (N=3,154, 80.3%). Height was rounded to the nearest 0.5 for determining weight-for-length percentiles.
[f]SDQ Difficulties Scores were calculated by summing scores from four scales: emotional problems, conduct problems, hyperactivity/inattention, and peer relationship problems.
[g]SDQ-P values were available for 3,129 children (79.7%). Scores could not be calculated for 22.3% of children exposed to insulin, and 18.4% of children exposed to metformin.
[g]SDQ-T values were available for 1,681children (42.8%). Scores could not be calculated for 47.1% of children exposed to insulin, and 38.7% of children exposed to metformin.

Table 5.3 Unadjusted and adjusted risk ratio estimates comparing child health outcomes between children exposed in utero to metformin versus insulin for treatment of gestational diabetes, 2005-2012

	Metformin N=1,996 N (%)	Insulin N=1,932 N (%)	Unadjusted Risk Ratio (95% CI)	Adjusted[a] Risk Ratio (95% CI)
Child Growth Outcomes				
Weight ≥ 85th Percentile	600 (36.6)	520 (34.3)	1.07 (0.98, 1.18)	1.04 (0.93, 1.16)
Weight ≥ 97th Percentile	235 (14.4)	202 (13.3)	1.08 (0.90, 1.28)	0.93 (0.76, 1.13)
Weight z score ≥1 SD	616 (37.6)	549 (36.2)	1.04 (0.95, 1.14)	1.01 (0.91, 1.12)
Weight z score ≥2 SD	207 (12.6)	172 (11.4)	1.13 (0.94, 1.37)	0.94 (0.76, 1.18)
Height ≥ 85th Percentile	441 (26.9)	341 (22.5)	1.19 (1.05, 1.34)	1.06 (0.92, 1.22)
Height ≥ 97th Percentile	136 (8.3)	95 (6.3)	1.32 (1.03, 1.70)	1.08 (0.81, 1.43)
Height z score ≥1 SD	450 (27.5)	355 (23.4)	1.17 (1.03, 1.31)	1.05 (0.91, 1.20)
Height z score ≥2 SD	116 (7.1)	76 (5.0)	1.41 (1.07, 1.87)	1.13 (0.82, 1.54)
Weight-for-Length ≥85th Percentile	600 (36.6)	612 (40.4)	0.91 (0.83, 0.99)	0.92 (0.83, 1.02)
Weight-for-Length ≥ 97th Percentile	252 (15.4)	244 (16.1)	0.96 (0.82, 1.13)	0.91 (0.75, 1.09)
Weight-for-Length z score ≥1 SD	599 (36.8)	614 (40.6)	0.91 (0.83, 0.99)	0.91 (0.83, 1.01)
Weight-for-Length z score ≥2 SD	219 (13.4)	210 (13.9)	0.97 (0.82, 1.16)	0.89 (0.73, 1.09)
BMI ≥ 85th Percentile	605 (36.9)	618 (40.8)	0.91 (0.83, 0.99)	0.92 (0.83, 1.01)
BMI ≥ 97th Percentile	267 (16.3)	256 (16.9)	1.01 (0.86, 1.18)	0.92 (0.77, 1.11)
BMI z score ≥1 SD	622 (38.0)	641 (42.3)	0.90 (0.83, 0.98)	0.91 (0.83, 1.00)
BMI z score ≥2 SD	233 (14.2)	220 (14.5)	0.99 (0.83, 1.17)	0.89 (0.73, 1.08)
BMI ≥ 21	59 (3.6)	51 (3.4)	1.07 (0.74, 1.55)	0.81 (0.53, 1.22)
Strengths and Difficulties Questionnaire				
Abnormal SDQ-P Difficulties Score (≥14)[f]	152 (9.3)	137 (9.1)	1.03 (0.83, 1.28)	1.13 (0.88, 1.46)
Abnormal SDQ-T Difficulties Score (≥11)[g]	73 (9.5)	74 (8.1)	1.16 (0.85, 1.58)	1.21 (0.85, 1.71)

[a]Inverse probability of treatment weights (IPTW) were estimated using a propensity score model containing the following covariates: maternal age, maternal ethnicity, maternal NZ deprivation decile score, parity, BMI prior to prescription initiation, smoking status, timing of GDM diagnosis and treatment, history of GDM and history of GDM treatment. The IPTW adjusted model uses imputed values for parity, BMI, and smoking status. These models are also adjusted for geographic region.

[b]Weight was avaliable for 3,156 children (80.4%). Weight measurements were missing for 21.5% of insulin-exposed children and 17.9% of metformin exposed children.

[c]Z-scores and percentiles were calculated based on a WHO 2006 reference standard for children aged 0-5. One child was measured at 62 months and excluded from these calculations.

[d]Height was available for 3,154 children (80.3%). Height measurements were missing for 21.5% of insulin-exposed children and 17.9% of metformin exposed children.

[e] Weight for length and BMI measurements were calculated for children with non-missing weight and height (N=3,154, 80.3%). Height was rounded to the nearest 0.5 for determining weight-for-length percentiles.

[f]SDQ-P Difficulties Scores were calculated by summing scores from four scales: emotional problems, conduct problems, hyperactivity/inattention, and peer relationship problems. SDQ-P values were available for 3,129 children (79.7%). The SDQ-P Difficulties Score could not be calculated for 22.3% of children exposed to insulin, and 18.4% of children exposed to metformin.

[g]SDQ-T Difficulties Scores were calculated by summing scores from four scales: emotional problems, conduct problems, hyperactivity/inattention, and peer relationship problems. SDQ-T values were available for 1,681children (42.8%). The SDQ-T Difficulties Score could not be calculated for 47.1% of children exposed to insulin, and 38.7% of children exposed to metformin.

Table 5.4 Sex-specific adjusted risk ratio estimates comparing child growth and development outcomes between children born to women treated with metformin versus insulin for gestational diabetes

	Boys			Girls			
	Metformin N=1082 N(%)	Insulin N=1044 N(%)	Adjusted[a] Risk Ratio (95% CI)	Metformin N=914 N(%)	Insulin N=888 N(%)	Adjusted[a] Risk Ratio (95% CI)	p-value for interaction
Child Growth Outcomes							
Weight ≥ 85th Percentile[b,c]	354 (39.8)	311 (38.2)	1.01 (0.88, 1.16)	246 (32.9)	209 (29.8)	1.05 (0.89, 1.25)	0.6736
Weight ≥ 97th Percentile	149 (16.7)	120 (14.7)	0.94 (0.72, 1.22)	86 (11.5)	82 (11.7)	0.88 (0.64, 1.21)	0.4213
Weight z score ≥1 SD	361 (40.6)	329 (40.4)	0.98 (0.86, 1.11)	255 (34.1)	220 (31.4)	1.03 (0.87, 1.22)	0.5084
Weight z score ≥2 SD	132 (14.8)	107 (13.1)	0.90 (0.68, 1.19)	75 (10.0)	65 (9.3)	0.98 (0.69, 1.39)	0.8699
Height ≥ 85th Percentile[c,d]	257 (28.9)	203 (24.9)	1.00 (0.84, 1.21)	184 (24.6)	138 (19.7)	1.12 (0.89, 1.39)	0.6115
Height ≥ 97th Percentile	84 (9.4)	57 (7.0)	1.05 (0.72, 1.52)	52 (7.0)	38 (5.4)	1.12 (0.72, 1.76)	0.6090
Height z score ≥1 SD	260 (29.3)	211 (25.9)	0.98 (0.82, 1.17)	190 (25.4)	144 (20.5)	1.12 (0.90, 1.38)	0.4728
Height z score ≥2 SD	73 (8.2)	47 (5.8)	1.11 (0.74, 1.67)	43 (5.8)	29 (4.1)	1.11 (0.66, 1.87)	0.7242
Weight-for-Length ≥85th Percentile[c,e]	339 (38.1)	345 (42.4)	0.91 (0.79, 1.03)	261 (34.9)	267 (38.0)	0.92 (0.78, 1.07)	0.7004
Weight-for-Length ≥ 97th Percentile	149 (16.7)	146 (17.9)	0.83 (0.65, 1.05)	103 (13.8)	98 (14.)	1.01 (0.75, 1.35)	0.6511
Weight-for-Length z score ≥1 SD	338 (38.2)	347 (42.7)	0.89 (0.78, 1.01)	261 (35.1)	267 (38.1)	0.92 (0.79, 1.08)	0.6189
Weight-for-Length z score ≥2 SD	128 (14.4)	129 (15.9)	0.77 (0.59, 1.00)	91 (12.2)	81 (11.5)	1.08 (0.78, 1.49)	0.3653
BMI ≥ 85th Percentile[c,e]	344 (38.7)	352 (43.2)	0.89 (0.79, 1.02)	261 (34.9)	266 (38.0)	0.92 (0.79, 1.08)	0.5900
BMI ≥ 97th Percentile	164 (18.4)	162 (19.9)	0.85 (0.67, 1.06)	103 (13.8)	94 (13.4)	1.02 (0.76, 1.38)	0.6126
BMI z score ≥1 SD	356 (40.0)	367 (45.1)	0.89 (0.79, 1.01)	266 (35.6)	274 (39.1)	0.91 (0.78, 1.07)	0.6145
BMI z score ≥2 SD	141 (15.8)	142 (17.4)	0.78 (0.61, 0.99)	92 (12.3)	78 (11.1)	1.08 (0.78, 1.50)	0.3507
BMI ≥ 21	33 (3.7)	26 (3.2)	0.81 (0.44, 1.46)	26 (3.5)	25 (3.6)	0.73 (0.40, 1.34)	0.5772
Strengths and Difficulties Questionnaire							
Abnormal SDQ-P Difficulties Score (≥14)[f,g]	103 (11.6)	86 (10.7)	1.20 (0.89, 1.63)	49 (6.6)	51 (7.3)	0.98 (0.61, 1.56)	0.5563

	Boys			Girls			
	Metformin N=1082 N(%)	Insulin N=1044 N(%)	Adjusted[a] Risk Ratio (95% CI)	Metformin N=914 N(%)	Insulin N=888 N(%)	Adjusted[a] Risk Ratio (95% CI)	p-value for interaction
Abnormal SDQ-T Difficulties Score (≥11)[f,h]	53 (5.9)	51 (6.2)	1.24 (0.82, 1.88)	20 (5.6)	23 (5.3)	1.24 (0.68, 2.25)	0.9692

[a]Inverse probability of treatment weights (IPTW) were estimated using sex-specific propensity score models containing the following covariates: maternal age, maternal ethnicity, maternal NZ deprivation decile score, parity, BMI prior to prescription initiation, smoking status, timing of GDM diagnosis and treatment, history of GDM and history of GDM treatment. The IPTW adjusted model uses imputed values for parity, BMI, and smoking status. These models are also adjusted for geographic region.

[b]Weight was avaliable for 3,156 children (80.4%). Weight measurements were missing for 21.5% of insulin-exposed children and 17.9% of metformin exposed children. Weight was missing for 421 boys (24.7%) and 351 girls (24.2%).

[c]Z-scores and percentiles were calculated based on a WHO 2006 reference standard.

[d]Height was available for 3,154 children (80.3%). Height measurements were missing for 21.5% of insulin-exposed children and 17.9% of metformin exposed children. Height was missing for 422 boys (24.7%) and 352 girls (24.2%).

[e] Weight for length and BMI measurements were calculated for children with non-missing weight and height (N=3,154, 80.3%). Height was rounded to the nearest 0.5 for determining weight-for-length percentiles.

[f]SDQ Difficulties Scores were calculated by summing scores from four scales: emotional problems, conduct problems, hyperactivity/inattention, and peer relationship problems.

[g]SDQ-P values were available for 3,129 children (79.7%) (1,692 boys, 1,437 girls). The SDQ-P Difficulties Score could not be calculated for 22.3% of children exposed to insulin, and 18.4% of children exposed to metformin.

[h]SDQ-T values were available for 1,681children (42.8%) (888 boys, 793 girls). The SDQ-T Difficulties Score could not be calculated for 47.1% of children exposed to insulin, and 38.7% of children exposed to metformin.

CHAPTER 6 – DISCUSSION

Globally, the prevalence of GDM is rising as maternal obesity and insulin resistance becomes more widespread. Women diagnosed with GDM are at greater risk of adverse pregnancy outcomes, including high blood pressure and high birthweight babies. This contributes to an increased need for birth interventions like labor induction and c-section delivery. Infants born to women with GDM are also at increased risk of low blood glucose at birth (neonatal hypoglycemia) and birth injuries (e.g., shoulder dystocia). These complications can be severe, particularly in resource-poor settings without access to advanced medical technology.

Managing GDM has been shown to be effective for mitigating these adverse outcomes, and therefore critical to improving maternal and infant health outcomes. Randomized controlled trials have shown that active management of GDM with diet or pharmacologic therapy reduces the risk of adverse maternal and neonatal outcomes compared with usual prenatal care without treatment. Identifying treatment strategies that can be implemented and accepted in multiple resource settings is therefore essential to addressing the adverse health outcomes of unmanaged GDM. Women may find insulin prohibitively expensive or challenging to obtain and store, and these barriers may be amplified in resource-poor areas. In a recent focus group study, 16 US women requiring pharmacologic management of GDM were asked about their experiences with diagnosis and treatment, including the options they considered and their concerns related to these options (120). Although some women reported feelings of relief after starting insulin, most struggled to manage their blood sugar and even reported episodes of hypoglycemia. Many women participating in the focus groups reported that taking insulin

was inconvenient, cumbersome, and physically and emotionally uncomfortable. While women were interested in alternate treatment options like oral hypoglycemic agents (OHAs), these were not offered, and questions remained about how these medications work and impact their child's health in the future. Providing high-quality evidence for the safety and effectiveness of these treatment options is therefore incredibly valuable for guiding women through treatments that work to manage their condition and fit their lifestyles.

Prenatal care providers have been increasingly interested in using OHAs as first-line treatment for GDM. Results from randomized controlled trials suggest that OHAs glyburide and metformin are both options that would be as effective as insulin in maintaining glycemic control for women diagnosed with GDM. However, these studies were not powered to evaluate modest but clinically meaningful differences in adverse maternal and neonatal outcomes. Such differences may only become evident in studies with large sample sizes or meta-analyses. In the case of glyburide, results from a large observational study and recent meta-analyses led the ACOG to change its recommendation on using glyburide as a first-line treatment in the US (63), while metformin remains an option for women who cannot safely administer insulin. Estimates from meta-analyses suggest that metformin may actually be more effective than insulin in preventing short-term complications, but even these estimates were limited in precision due to small sample sizes from randomized trials. To address the uncertainty regarding the comparative safety and effectiveness of metformin versus insulin, we conducted a large, population-based cohort of New Zealand pregnancies treated pharmacologically for GDM.

6.1 Summary of Findings

We did not observe increased risks of adverse maternal or neonatal outcomes associated with metformin use compared with insulin. Compared to insulin, metformin use was associated with reduced risks of elective c-section, labor induction, infant macrosomia, neonatal hypoglycemia, and preterm birth. Our estimates for rare adverse outcomes such as

preeclampsia and shoulder dystocia do not suggest major differences in risk of these events among metformin and insulin users. While rare, our confidence intervals for these associations were more precise than even the combined estimates from a recent meta-analysis.10 In stratified models, we observed variation in elective c-section estimates by ethnicity, as well as variation in the degree to which metformin use affected neonatal hypoglycemia risk. We also observed variation by infant sex in the risk difference estimates for the association of metformin versus insulin treatment and two outcomes, neonatal hypoglycemia and large-for-gestational-age. This suggests that while metformin may overall be an appropriate alternative to insulin for first-time GDM treatment, there may be individuals for whom insulin is more effective.

We were also able to leverage results from a universal pre-school assessment provided as part of regular well-child care to children in New Zealand to study health outcomes in children born to mothers who had been treated with metformin or insulin for GDM. In this study, we did not observe meaningful differences in child growth or developmental outcomes at age 4. Children whose mothers were treated initially with metformin versus insulin to manage GDM were similar with respect to weight, height, and BMI and we did not observe statistically or clinically significant differences between them after adjusting for measured covariates. We observed variation in model estimates by child sex, but there was not a statistically significant interaction between treatment and sex. Our findings are consistent with prior studies focused on child health outcomes at younger ages.

6.2 Strengths and Limitations

Our cohort study was conducted using population-level healthcare data and represents GDM treatment and outcomes throughout New Zealand. By using a maternity care dataset linked with other clinical and demographic data, we were able to incorporate information not usually found in administrative databases for pregnancy research, including information on gestational age and birthweight. Our study focused on maternal and infant

outcomes was approximately ten times the size of the largest randomized controlled trial to date. This allowed us to calculate estimates of association with greater precision, and to study outcomes that previously studies did not have the power to examine. Additionally, we were able to investigate associations within key subgroups, including estimating ethnicity- and infant sex-specific measures of association.

We did not have access to women's laboratory values from either outpatient or home-based glucose testing. This means that we could not control for baseline disease severity as measured by blood glucose levels. However, other indicators of disease severity like maternal BMI and timing of diagnosis and treatment were similar between the two groups. Additionally, we were also unable to restrict our cohort to women who were diagnosed with GDM based on results from oral glucose tolerance tests. Instead, we learned from discussions with NZ clinicians that women diagnosed with GDM are provided free blood glucose monitors and prescriptions for glucose test strips to identify newly diagnosed patients in prescription claims. While we were able to exclude women who had filled a prescription for blood glucose test strips or antidiabetics prior to 24 weeks gestation, it is possible that some women may have had undetected type 2 diabetes prior to pregnancy or GDM screening.

Our study had missing data on BMI, smoking status, parity and ethnicity for some women. We employed an advanced epidemiologic method for addressing missing data, multiple imputation using chained equations, for women with missing covariate information. However, because BMI is only reported at the first prenatal care encounter in the Maternity Collection, we did not have information on maternal weight gain during pregnancy, which is a likely confounder. We were also missing some granularity for certain outcomes. We used ICD-10 codes to define neonatal hypoglycemia and could not confirm this diagnosis based on laboratory values or determine severity. Our estimates are consistent with previous studies, including a recent meta-analysis, suggesting that an ICD-10 code-based definition

may be a suitable proxy. Likewise, we were unable to determine how many hours an infant stayed in the neonatal intensive care unit, and therefore could not distinguish between short- and long-term stays. However, we were able to observe if the infant remained in the hospital after the mother had already been discharged, indicating potentially serious complications that required extra medical attention.

Since this study was conducted in New Zealand, its results and conclusions may not be generalizeable to other populations. GDM disproportionately affects Black and Hispanic women in the US, but these ethnic groups are not well-represented in New Zealand. If there are ethnic differences in the effectiveness of metformin compared with insulin, our results may not be reproducible in a US population. Additionally, the New Zealand healthcare system is publicly-funded, and cost may be a less significant barrier to diagnosis and treatment of GDM than it is in the US. Women diagnosed with GDM in the US may be more likely to have undetected or untreated type 2 diabetes than in New Zealand and may also have a difficult time affording insulin treatment. Future research should examine how cost and healthcare history factor in to GDM treatment decisions. Because insulin may be a more expensive treatment in certain settings, including the US, metformin may be the most feasible option for women with GDM.

6.3 Public Health Implications

There is substantial uncertainty surrounding the options for pharmacologic treatment of GDM. Based on the results of our study, initiating metformin as pharmacologic therapy appears to be at least as effective as insulin for reducing serious adverse events at birth. In some cases metformin may be a better first line option for treating GDM than insulin, which is consistent with results from a recent meta-analysis (52). We observed a reduced risk of large-for-gestational-age, neonatal hypoglycemia, and NICU admissions associated with metformin versus insulin initiation. However, approximately 12% of women initiation metformin required supplemental treatment with insulin. This study provides new evidence

to support the consideration of metformin as a first-line pharmacotherapy for GDM. Careful blood glucose monitoring will be necessary to determine if supplemental treatment with insulin is indicated.

In our analyses stratified by ethnicity, we did not see substantial variation in the effectiveness of metformin compared with insulin for most outcomes. However, metformin was associated with a meaningfully reduced risk of neonatal hypoglycemia compared with insulin for every ethnicity except Asian women, for whom there was no difference in risk. We observed similarly attenuated estimates for Asian women when comparing risks of large-for-gestational-age as well. This may suggest that there are subgroups in whom metformin is clearly a better option for preventing serious adverse neonatal outcomes compared with insulin.

One major reason that metformin has not been universally endorsed as a first-line treatment for GDM is the concern that it may adversely impact fetal development in ways that insulin would not. Metformin has been shown to cross the placenta due to its small molecular size, while insulin does not. Additionally, metformin and insulin may differ in how effective each medication is for managing GDM and avoiding adverse infant outcomes like macrosomia or neonatal hypoglycemia. This can have important downstream consequences for child health. Without long-term health data in children born to women treated with metformin for GDM, patients and providers have not felt confident in metformin's safety for use during pregnancy. Our investigation of the comparative effectiveness and safety of metformin versus insulin therefore provides desired information to help guide treatment decisions for women requiring pharmacologic therapy to treat GDM.

6.4 Conclusions

This study presents new evidence for the comparative effectiveness and safety of initial metformin pharmacotherapy for GDM compared with insulin. Our examination of long-term growth and development outcomes in preschool-aged children is a novel contribution

79

to the literature and can be considered by physicians and patients when making treatment decisions for GDM. Future research should focus on indicators of treatment failure in women initiating metformin to refine pharmacotherapy recommendations. Additionally, more research is needed to examine the effectiveness of metformin compared to insulin in subgroups with elevated risk for GDM that this study could not include, specifically Black and Hispanic women in the US.

APPENDIX 1. SUMMARY OF PREVIOUS STUDIES

Reference	Study Design	Total and Stratified N	Study Population	Results	Limitations
Maternal and Infant Outcomes					
Moore et al. 2007	RCT	Metformin Group N=32 Insulin Group N=31 Total N=63	-Women receiving prenatal care at the University of Mississippi Medical Center in Jackson, MS -Diagnosis of GDM with 2 or more abnormal values during 3-h glucose tolerance test (cut-off s1-h 190 mg/dL, 2-h 165 mg/dL, 3-h 145 mg/dL)	More caesarean deliveries occurred in insulin group (n=10) versus metformin group (7). Mean birthweight between treatment groups was similar (insulin: 3500.2 ± 700.5, metformin: 3451.8 ± 727.5). There were no cases of neonatal hypoglycemia in the metformin group, and 2 cases for the insulin group.	-Small sample size (<100) -Treatment groups not well-balanced by pre-pregnancy weight, gravidity, or ethnicity. Metformin group patients had higher BMI than insulin patients.
Rowan et al. 2008	RCT	Metformin Group N=363 Insulin Group N=370 Total N=731 (751 enrolled)	-Women aged 18 to 45 years between 20-33 weeks gestation -10 New Zealand and Australia urban hospitals -Diagnosis of GDM according to Australasian Diabetes in Pregnancy Society -Had more than one fasting glucose measurement above 5.4 mmol/L or more than one 2-hour measurement above 6.7 mmol/L.	Primary neonatal composite outcome: RR=0.99 (95% CI 0.80-1.23) Neonatal hypoglycemia: RR 0.81 (95% CI 0.59-1.12) Respiratory distress: RR 0.76 (95% CI 0.37-1.59) Birth trauma: RR 0.96 (95% CI 0.49-1.87) Preterm birth (all): RR 1.60 (95% CI 1.02-2.52) Apgar score <7: RR 3.06 (95% CI 0.32-29.26) NICU admission: RR 0.89 (95% CI 0.66-1.19)	-Women receiving metformin were more likely to have ≥3 prior pregnancy terminations or miscarriages and less likely to be of Indian or Polynesian ethnicity. -Treatment was open-label and not blinded. -Primary composite outcome comprises events of varying clinical significance. The sample size is large for an RCT but too small overall to detect small but meaningful differences in rare outcomes.

Reference	Study Design	Total and Stratified N	Study Population	Results	Limitations
Ijäs et al. 2010	RCT	Metformin Group N=47 Insulin Group N=50 Total N=97 (100 enrolled)	-Women with singleton pregnancies between 12-34 weeks gestation (mean age 32) -Attending maternity welfare clinics at two hospitals in Kajaani, Finland -Diagnosed with GDM based on risk factor-based screening with 2-hour OGTT	LGA: RR 0.90 (95% CI 0.24-2.98) Neonatal hypoglycemia: RR 0.70 (95% CI 0.23-1.89) Hyperbilirubinemia: RR 0.80 (95% CI 0.43-1.39) NICU admission: RR 0.70 (95% 0.29-1.60) Caesarean section: RR 1.90 (95% 0.99-3.31) Labor induction: RR 1.00 (95% CI 0.67-1.45)	-Small sample size (<100) -Blinding was not possible for women due to route of administration, but unclear if providers were blinded to treatment as well.
Niromanesh et al. 2012	RCT	Metformin Group N=80 Insulin Group N=80 Total N=160 (172 enrolled)	-Women with singleton pregnancies aged 18 to 40 years -Between 20 and 34 weeks gestation -Receiving prenatal care at one of three hospitals in Tehran, Iran -Diagnosed with GDM using Carpenter and Coustan guidelines	Birthweight >90th percentile: RR 0.5 (95% CI 0.3-0.9) Neonatal hypoglycemia: RR 1.5 (95% CI 0.3-8.7) Respiratory distress: RR 2.5 (95% CI 0.5-12.5) Hyperbilirubinemia: RR 2.0 (95% CI 0.8-4.7) Preterm birth (all): RR 2.2 (95% CI 0.7-7.0) NICU admission: RR 2.5 (95% CI 0.5-12.5) Caesarean section (all): RR 0.7 (95% CI 0.2-2.2) Caesarean section (emergency): RR 1.6 (95% CI 0.9-2.7) Shoulder dystocia: RR 0.5 (95% CI 0.1-2.6) Preeclampsia: RR 0.7 (95% CI 0.2-2.2)	-Small sample size (<200) -2-hour OGTT values slightly higher for insulin group. -Participants were not blinded, but providers were.

Reference	Study Design	Total and Stratified N	Study Population	Results	Limitations
Barrett et al 2013	RCT	Metformin Group N=236 Insulin Group N=242 Total N=478 samples collected (733 enrolled in original MiG trial)	-Same study population as Rowan et al. 2008 (MiG trial) -Subset of women with available cord plasma samples.	Metabolic markers were similar between neonates exposed to metformin and insulin. Glucose in cord plasma was similar (4.38 mmol/l [95% CI 4.15-4.63] for metformin-exposed; 4.46 mmol/l [95% CI 4.28-4.65] for insulin-exposed). Higher maternal values of HbA1c and triglycerides were associated with birthweight >90th percentile in both treatment groups. Glucose values from 36 weeks gestation to 6-8 weeks postpartum increased more for women treated with insulin (mean 4.29 to 5.04) than metformin (mean 4.43 to 4.94).	-Sample size varied based on availability of samples. -There is little information on the baseline characteristics of the subset of mothers with available measurements, and comparisons between metformin and insulin treated groups do not appear to control for these differences.
Waheed et al. 2013	RCT	Metformin Group N=34 Insulin Group N=34 Total N=68	-Women seen at a hospital in Islamabad, Pakistan beyond 14 weeks gestation -Diagnosed with diabetes by fasting blood sugar and random blood sugar tests	There were no statistically significant differences (p<0.05) between the two groups with regard to post-treatment blood sugar levels and HBA1C results.	-Small sample size (<100) -Sample size included both overt diabetes and GDM cases.
Mesdaghinia et al. 2013	RCT	Metformin Group N=100 Insulin Group N=100 Total N=200	-Women with singleton pregnancies aged 18 to 45 years -Between 24 and 34 weeks gestation -Receiving care at one hospital in Isfahan, Iran -Diagnosed with GDM with two-step criteria	Results of chi-squared test showed no significant differences (p<0.05) with regard to birth weight, LGA, hypoglycemia, birth defects, shoulder dystocia, stillbirth, sepsis, or Apgar scores. Significant differences were found for jaundice, respiratory distress,	-Small sample size (200). -Participants weren't blinded but providers were. -Very little information about outcome definitions (i.e., spontaneous versus induced preterm birth,

Reference	Study Design	Total and Stratified N	Study Population	Results	Limitations
			(glucose challenge test followed by glucose tolerance test)	NICU admission, and preterm labor.	cutpoints for hypoglycemia)
Ruholamin et al. 2013	RCT	Metformin Group N=50 Insulin Group N=50 Total N=100 (119 enrolled)	-Women with singleton pregnancies aged 18 to 45 years -Between 24 and 33 weeks of gestation -Diagnosed with GDM according to the Australasian Diabetes in Pregnancy Society guidelines	There were no statistically significant differences (p<0.05) between the two groups with regard to hyperbilirubinemia, respiratory distress syndrome, NICU admissions, shoulder dystocia, hypoglycemia, Apgar scores, macrosomia, small for gestational age, and stillbirth. The average birth weight and postnatal blood sugar were higher for the metformin treated women but not significantly.	-Small sample size (<200) -Study says participants were double-blinded but unclear how they were blinded to receiving metformin (pill) versus insulin (injection)
Goh et al. 2011	Observational	Metformin Group N=465 Insulin Group N=399 Total N=1,269 (includes diet-controlled)	-Women with singleton pregnancies with GDM delivering after 20 weeks of gestation -Seen at National Women's Health in Auckland, NZ -Diagnosed with GDM based on fasting glucose of 5.5 mmol/L or higher or 2-h glucose of 9.0 mmol/L or higher	Women receiving metformin were less likely to have LGA infants, preterm birth (iatrogenic), NICU admissions, and use of intravenous glucose to the infant.	-Small sample size (<1,000 for medication users) -No information on treatment duration, glucose control achieved after treatment. -Women receiving insulin were more likely to be European; women receiving metformin more likely to be Maori or defined as Other Asian. -Effect estimates were not estimated that controlled for confounding.

Reference	Study Design	Total and Stratified N	Study Population	Results	Limitations
Terrti et al. 2008	Observational	Metformin Group N=45 Insulin Group N=45 Total N=173 (includes diet controls)	-Women with singleton pregnancies diagnosed with GDM from risk factor-based screening -Treated at one hospital in Finland	Macrosomia: OR 0.6 (95% CI 0.2-1.9) Hypoglycemia: OR 0.4 (95% CI 0.2-0.9) Hyperbilirubinemia: OR 1.1 (95% CI 0.4-2.9) Respiratory distress: NS 0.0 (too small) Caesarean section delivery: OR 1.6 (95% CI 0.6-4.2)	-Insulin users had higher OGTT values. They were more likely to have labor induced and were more likely to be primipara. -Patients were matched on pre-pregnancy BMI (categories) and age (categories) but analyses did not control for other confounders.
Baani et al. 2009	Observational	Metformin Group N=100 Insulin Group N=100 Total N=200 (27 in Metformin switched to insulin or stopped)	-Women with inadequately controlled GDM (diagnosed as 2-h glucose value ≥ 7.8mmol/l at 28 weeks). Inadequate control defined as three or more tests out of range (<6.0 mmol/l fasting, <8.0 mmol/l 1-h, <7.0 mmol/l 2-h) within 2 weeks or within 1 week of any treatment change -Seen at antenatal clinic in Surrey, UK hospital between January 2007 and July 2008	Women receiving metformin were significantly less likely to have preeclampsia (2 vs 9%); preterm birth (0 vs 10%); or admission to the NICU (6 vs 19%) than those receiving insulin. Infants exposed to metformin were less likely to be LGA >90th percentile (14 vs 25%) and had smaller birthweight overall (3372 ± 474 vs 3511 ± 511); less likely to have jaundice (8 vs 30%), hypoglycemia (9 vs 18%), and NICU admission (6 vs 19%).	-Small sample size (200 total). -Women in insulin group more likely to have prior gestational diabetes (25%) than metformin (13%), and had higher OGTT values overall. -Gestational age at entry (defined as inadequate glucose control) was earlier for insulin group (30 weeks) than metformin (32 weeks). -Proportions and mean values were compared without controlling for confounders.

Reference	Study Design	Total and Stratified N	Study Population	Results	Limitations
Child Health Outcomes					
Rowan et al. 2010	RCT (follow-up)	Metformin Group N=154 Insulin Group N=164 Total N=318 (731 from MiG trial)	-2-year old children born to women participating in the MiG trial (see Rowan et al. 2008 above for details on enrolled mothers)	Children exposed to metformin and insulin in the MiG trial had similar anthropometric measurements. Children exposed to metformin had larger chest circumference (52.1 ± 3.0 vs 51.6 ± 3.0), upper-arm circumference (17.2 ± 1.5 vs 16.7 ± 1.5), subscapular skinfold thickness (6.3 ± 1.9 vs 6.0 ± 1.7) and biceps skinfold thickness (6.0 ± 1.9 vs 5.6 ± 1.7) in centimeters as compared with insulin-exposed children. In children with available DEXA measurements (57 per group), children exposed to metformin had a smaller % of abdominal fat mass (5.3 ± 1.3 vs 5.6 ± 1.3) than insulin-exposed children but not significantly.	-Small sample size (318 overall, 114 with DEXA measurements) -Mothers of children participating in this follow-up study who had received metformin were more likely to be European/Caucasian (57.1 vs 47.6%) than those who had received insulin and were more likely to have smoked during pregnancy (10.4 vs 6.1) though not statistically significantly. -Statistical significance partly driven by comparing continuous outcomes; differences may not be meaningful.
Ijäs et al. 2015	RCT (follow-up)	Metformin Group N=45 Insulin Group N=48 Total N=93 (100 from original trial)	-18-month old children born to women participating in a previous trial (see Ijas et al 2010 above for details on enrolled mothers)	At 12 and 18 months, children exposed to metformin were more likely to have height ≥95% percentile (15.6 vs 2.1% at 12 months, 13.3 vs 4.2% at 18 months) than insulin-exposed children. At all three time points, metformin-exposed children	-Small sample size (<100) -Mothers of children randomized to metformin were less likely to be smokers (2.2 vs 8.3%) than insulin-assigned mothers. -Authors note that

Reference	Study Design	Total and Stratified N	Study Population	Results	Limitations
				weighed more than insulin-exposed children. Metformin-exposed infants had higher birthweight overall at the time of delivery. Metformin- and insulin-exposed were similar with respect to development.	maternal pre-pregnancy BMI was a strong predictor of child weight at 18 months but do not control for this or other confounders in their primary analysis.
Terrti et al. 2015	Observational	Metformin Group N=75 Insulin Group N=71 Total N=146 (217 from original trial)	-2-year old children born to women participating in a previous trial (see Terrti et al. 2008 above for details on enrolled mothers)	Children had lower scores on the Bayley-III language scale than the normative Finnish sample. Differences in scoring between metformin- and insulin-exposed infants were small and inconsistent.	-Small sample size (<200) -Metformin-treated mothers were less likely to be smokers (5.5 vs 12.9%). Information on the normative Finnish sample was limited.

APPENDIX 2. CONCEPTUAL DIAGRAM DEPICTING PROPOSED RELATIONSHIPS BETWEEN STUDY EXPOSURE, OUTCOMES, AND CONFOUNDERS
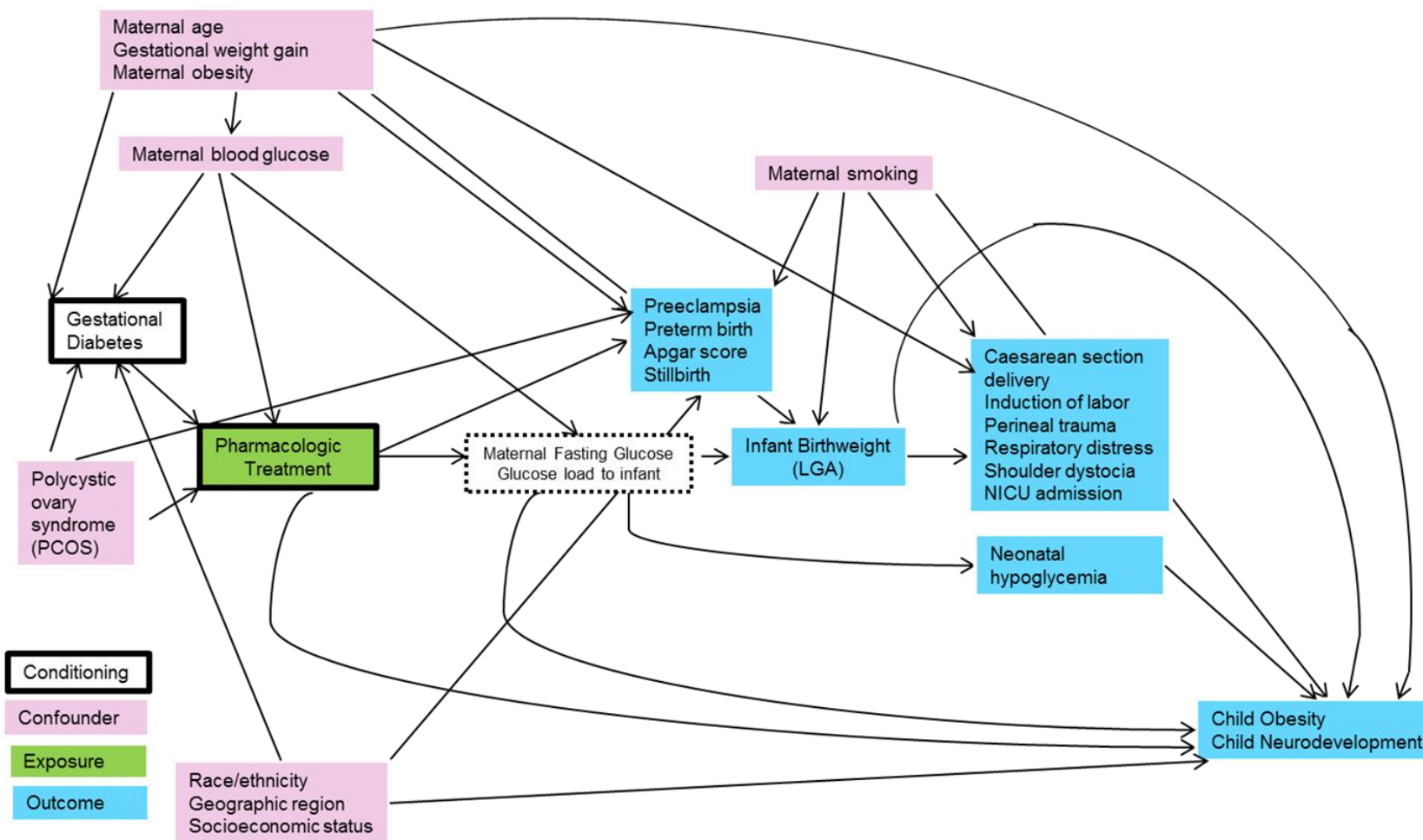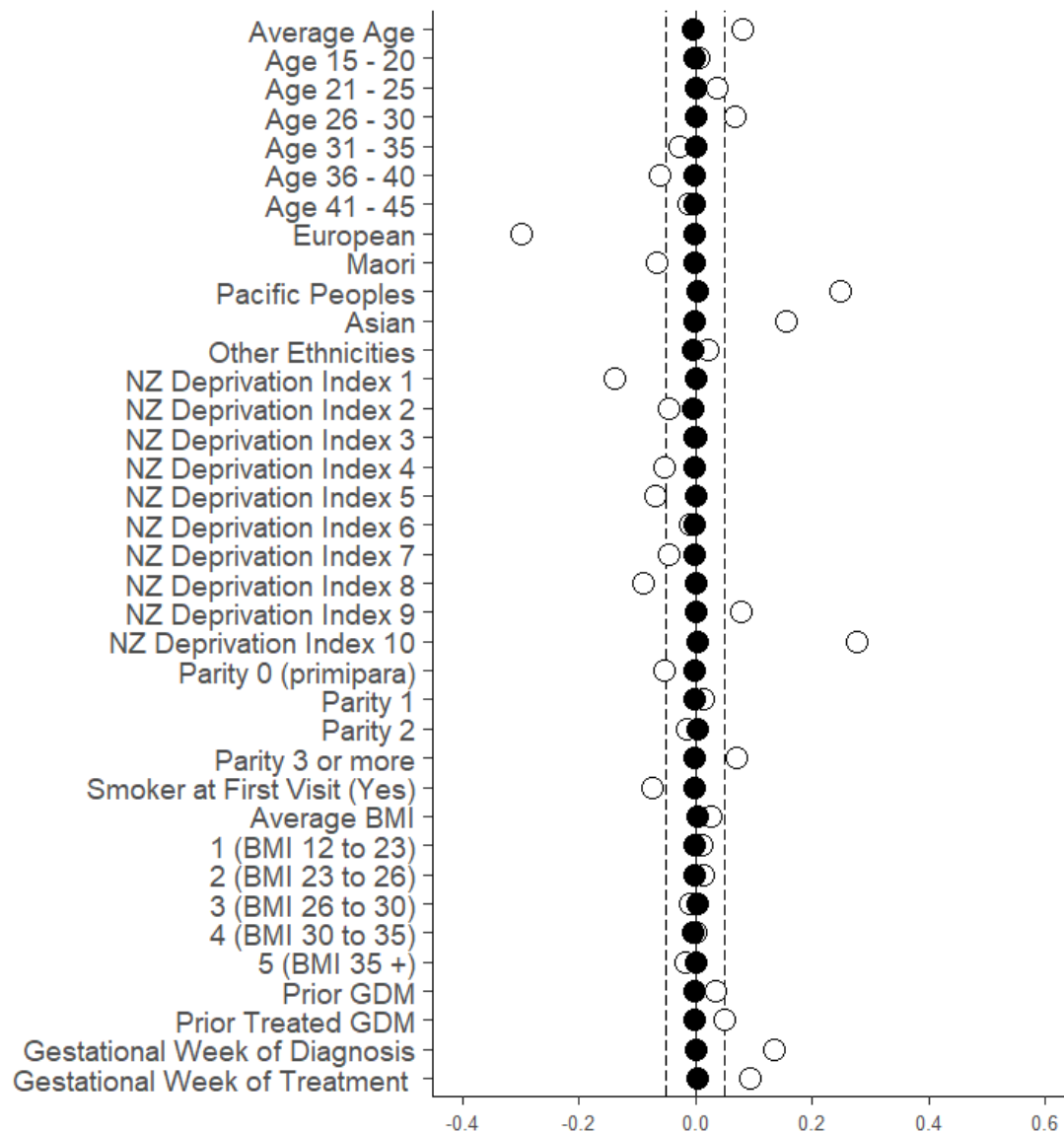
APPENDIX 4. TABLE OF MATERNAL AND NEONATAL OUTCOMES, STRATIFIED BY PRIORITIZED ETHNICITY

European

	Event Counts, (%)		Unadjusted		IPTW-Adjusted[a]	
Maternal Outcome	Metformin (N=1085)	Insulin (N=1468)	RR (95% CI)	RD (95% CI)	RR (95% CI)	RD (95% CI)
Elective c-section	249 (23.0)	408 (27.8)	0.83 (0.72, 0.95)	-4.8 (-8.3, -1.4)	0.83 (0.72, 0.95)	-4.7 (-8.2, -1.1)
Emergency c-section	193 (17.8)	255 (17.4)	1.03 (0.86, 1.22)	0.4 (-2.6, 3.5)	1.02 (0.85, 1.21)	0.5 (-2.6, 3.6)
Induction	545 (50.2)	786 (53.5)	0.94 (0.87, 1.01)	-3.2 (-7.2, 0.7)	0.93 (0.86, 1.00)	-4.2 (-8.3, -0.0)
Infant Outcome						
Birthweight (g) (Mean, SD)[b b,c]	3438.0 (503.8)	3440.5 (525.1)	-1.92 (-42.59, 38.74) p=0.93		-7.92 (-50.22, 34.38) p=0.71	
Large-for-gestational-age[b,d]	154 (14.2)	268 (18.3)	0.78 (0.65, 0.94)	-4.0 (-6.9, -1.1)	0.78 (0.65, 0.94)	-4.4 (-7.4, -1.4)
Macrosomia[b,e]	131 (12.1)	174 (11.9)	1.02 (0.82, 1.26)	0.2 (-2.4, 2.8)	1.06 (0.85, 1.32)	0.1 (-2.5, 2.7)
Neonatal hypoglycemia[f]	160 (14.8)	342 (23.3)	0.64 (0.54, 0.75)	-8.4 (-11.6, -5.5)	0.69 (0.58, 0.82)	-7.0 (-10.0, -3.9)
NICU admission[g]	178 (16.4)	441 (30.0)	0.56 (0.48, 0.66)	-13.1 (-16.3, -9.8)	0.68 (0.59, 0.79)	-6.7 (-9.5, -3.9)

Māori

	Event Counts, (%)		Unadjusted		IPTW-Adjusted[a]	
Maternal Outcome	Metformin (N=555)	Insulin (N=585)	RR (95% CI)	RD (95% CI)	RR (95% CI)	RD (95% CI)
Elective c-section	77 (13.9)	122 (20.9)	0.68 (0.52, 0.88)	-6.7 (-11.0, -2.3)	0.69 (0.52, 0.90)	-5.7 (-10.3, -1.2)
Emergency c-section	92 (16.6)	108 (18.5)	0.90 (0.70, 1.16)	-1.9 (-6.3, 2.5)	0.91 (0.70, 1.19)	-1.7 (-6.2, 2.9)
Induction	329 (59.3)	350 (59.8)	1.00 (0.91, 1.10)	-0.0 (-5.7, 5.7)	1.01 (0.91, 1.11)	0.2 (-5.7, 6.1)
Infant Outcome						
Birthweight (g) (Mean, SD)[b b,c]	3486.2 (519.4)	3534.8 (570.5)	-43.85 (-107.64, 19.95) p=0.18		-45.97 (-111.82, 19.87) p=0.17	
Large-for-gestational-age[b,d]	120 (21.6)	165 (28.2)	0.77 (0.63, 0.95)	-6.4 (-11.5, -1.3)	0.79 (0.64, 0.98)	-6.0 (-11.3, -0.7)

	Event Counts, (%) Metformin	Event Counts, (%) Insulin	Unadjusted RR (95% CI)	Unadjusted RD (95% CI)	IPTW-Adjusted[a] RR (95% CI)	IPTW-Adjusted[a] RD (95% CI)
Macrosomia[b,e]	79 (14.2)	111 (19.0)	0.77 (0.59, 1.00)	-4.5 (-8.8, -0.1)	0.75 (0.57, 0.98)	-3.7 (-8.1, 0.6)
Neonatal hypoglycemia[f]	88 (15.9)	147 (25.1)	0.63 (0.50, 0.80)	-9.4 (-14.0, -4.7)	0.66 (0.52, 0.85)	-8.2 (-12.9, -3.4)
NICU admission[g]	95 (17.1)	198 (33.9)	0.54 (0.43, 0.67)	-15.2 (-20.6, -9.9)	0.54 (0.43, 0.68)	-15.1 (-20.4, -9.8)

Pacific Peoples

	Event Counts, (%)		Unadjusted		IPTW-Adjusted[a]	
Maternal Outcome	Metformin (N=781)	Insulin (N=393)	RR (95% CI)	RD (95% CI)	RR (95% CI)	RD (95% CI)
Elective c-section	135 (17.3)	74 (18.8)	0.92 (0.72, 1.18)	-1.5 (-6.1, 3.1)	1.15 (0.87, 1.52)	1.7 (-3.4, 6.7)
Emergency c-section	137 (17.5)	82 (20.9)	0.85 (0.66, 1.08)	-3.2 (-8.0, 1.6)	0.83 (0.63, 1.08)	-3.4 (-8.8, 1.9)
Induction	425 (54.4)	233 (59.3)	0.91 (0.82, 1.01)	-5.2 (-11.2, 0.7)	0.87 (0.78, 0.97)	-8.3 (-14.8, -1.8)
Infant Outcome						
Birthweight (g) (Mean, SD)[b,c]	3551.3 (604.6)	3581.8 (613.7)	-23.76 (-97.66, 50.15) p=0.53		-64.67 (-147.29, 17.95) p=0.13	
Large-for-gestational-age[b,d]	136 (17.4)	92 (23.4)	0.78 (0.61, 1.00)	-5.0 (-10.1, 0.2)	0.72 (0.55, 0.94)	-5.7 (-11.5, 0.0)
Macrosomia[b,e]	158 (20.2)	89 (22.7)	0.90 (0.72, 1.13)	-2.4 (-7.3, 2.6)	0.84 (0.66, 1.08)	-3.1 (-8.6, 2.5)
Neonatal hypoglycemia[f]	110 (14.1)	99 (25.2)	0.57 (0.44, 0.72)	-11.0 (-16.0, -6.0)	0.65 (0.50, 0.84)	-7.8 (-13.1, -2.5)
NICU admission[g]	67 (8.6)	105 (26.7)	0.32 (0.24, 0.43)	-18.0 (-22.9, -13.2)	0.65 (0.50, 0.86)	-4.3 (-8.1, -0.5)

Asian

	Event Counts, (%)		Unadjusted		IPTW-Adjusted[a]	
Maternal Outcome	Metformin (N=1,279)	Insulin (N=909)	RR (95% CI)	RD (95% CI)	RR (95% CI)	RD (95% CI)
Elective c-section	220 (17.2)	183 (20.1)	0.85 (0.71, 1.02)	-3.0 (-6.3, 0.4)	0.93 (0.77, 1.13)	-1.3 (-4.8, 2.2)
Emergency c-section	215 (16.8)	173 (19.0)	0.88 (0.74, 1.06)	-2.2 (-5.5, 1.1)	0.90 (0.74, 1.09)	-1.8 (-5.2, 1.6)
Induction	615 (48.1)	461 (50.7)	0.94 (0.87, 1.03)	-2.9 (-7.2, 1.4)	0.92 (0.84, 1.00)	-4.5 (-9.0, 0.1)
Infant Outcome						

Birthweight (g) (Mean, SD)[b][b,c]	3154.0 (460.7)	3145.7 (498.9)	9.38 (-31.80, 50.56) p=0.66		40.25 (-4.84, 85.34) p=0.08	
Large-for-gestational-age[b,d]	129 (10.1)	118 (13.0)	0.77 (0.61, 0.97)	-3.0 (-5.7, -0.3)	0.89 (0.69, 1.14)	-1.0 (-3.8, 1.8)
Macrosomia[b,e]	31 (2.4)	38 (4.2)	0.54 (0.34, 0.87)	-2.0 (-3.5, -0.4)	0.79 (0.46, 1.36)	-1.2 (-2.8, 0.4)
Neonatal hypoglycemia[f]	169 (13.2)	152 (16.7)	0.89 (0.64, 0.96)	-3.6 (-6.7, -0.5)	0.97 (0.79, 1.19)	0.2 (-2.8, 3.2)
NICU admission[g]	99 (7.7)	191 (21.0)	0.37 (0.29, 0.46)	-13.4 (-16.5, -10.4)	0.59 (0.47, 0.74)	-4.1 (-6.6, -1.5)

[a]Inverse probability of treatment weights (IPTW) were estimated using a propensity score model containing the following covariates: maternal age, maternal ethnicity, maternal NZ deprivation decile score, parity, BMI prior to prescription initiation, smoking status, timing of GDM diagnosis and treatment, history of GDM and history of GDM treatment. The IPTW adjusted model uses imputed values for parity, BMI, and smoking status. These models are also adjusted for geographic region.
[b]Birthweight values were missing for 63 infants (29 metformin, 34 insulin).
[c]Linear regression was used to compare mean birthweight in infants treated initially with metformin versus insulin.
[d]Large-for-gestational-age is defined as a birthweight greater than the 90th percentile. These percentiles were calculated using a sample of New Zealand infants born between 2001 and 2015 with available data on birthweight and gestational age (n=850,506). These percentiles are additionally sex- and ethnicity-specific.
[e]Macrosomia is defined as birthweight greater than 4,000g.
[f]Hospitalization data for infants were missing for 39 delivery events (18 metformin, 21 insulin).
[g]NICU admission is defined as a hospitalization event with a health specialty code for a Level II or Level III neonatal intensive care unit.

APPENDIX 5. TABLE OF MATERNAL AND NEONATAL OUTCOMES FOR DELIVERIES AFTER 2008

	Events (%)		Unadjusted		IPTW-Adjusted[a]	
Maternal Outcome	Metformin (N=3,646)	Insulin (N=2,717)	RR (95% CI)	RD (95% CI)	RR (95% CI)	RD (95% CI)
Elective c-section	683 (18.7)	635 (23.4)	0.80 (0.73, 0.88)	-4.6 (-6.7, -2.6)	0.89 (0.80, 0.99)	-2.4 (-4.5, -0.2)
Emergency c-section	625 (17.1)	485 (17.9)	0.96 (0.86, 1.07)	-0.7 (-2.6, 1.2)	1.00 (0.89, 1.12)	0.0 (-2.0, 2.0)
Induction	1885 (51.7)	1502 (55.3)	0.93 (0.89, 0.98)	-3.7 (-6.1, -1.2)	0.91 (0.86, 0.95)	-5.3 (-8.0, -2.6)
Infant Outcome						
Birthweight (g) (Mean, SD)[b,c]	3367.0 (535.5)	3365.4 (545.8)	3.28 (-23.73, 30.28)		20.36 (-9.15, 49.88)	
Macrosomia[b,d]	383 (10.5)	295 (10.9)	0.96 (0.83, 1.11)	-0.4 (-2.0, 1.1)	1.02 (0.88, 1.20)	0.3 (-1.3, 2.0)
Large-for-gestational-age[b,e]	524 (14.4)	481 (17.7)	0.81 (0.73, 0.91)	-3.3 (-5.2, -1.5)	0.86 (0.76, 0.97)	-2.4 (-4.4, -0.4)
Neonatal hypoglycemia[f]	520 (14.3)	575 (21.2)	0.67 (0.60, 0.75)	-7.0 (-8.9, -5.1)	0.77 (0.69, 0.86)	-4.1 (-6.1, -2.1)
NICU admission[g]	429 (11.8)	803 (29.6)	0.40 (0.36, 0.45)	-17.6 (-19.7, -15.6)	0.60 (0.54, 0.66)	-6.9 (-8.7, -5.1)

[a]Inverse probability of treatment weights (IPTW) were estimated using a propensity score model containing the following covariates: maternal age, maternal ethnicity, maternal NZ deprivation decile score, parity, BMI prior to prescription initiation, smoking status, timing of GDM diagnosis and treatment, history of GDM and history of GDM treatment. The IPTW adjusted model uses imputed values for parity, BMI, and smoking status. These models are also adjusted for geographic region.

[b]Birthweight values were missing for 54 infants (27 metformin, 27 insulin).

[c]Linear regression was used to compare mean birthweight in infants treated initially with metformin versus insulin.

[d]Macrosomia is defined as birthweight greater than 4,000g.

[e]Large-for-gestational-age is defined as a birthweight greater than the 90th percentile. These percentiles were calculated using a sample of New Zealand infants born between 2001 and 2015 with available data on birthweight and gestational age (n=850,506). These percentiles are additionally sex- and ethnicity-specific.

[f]Hospitalization data for infants were missing for 35 delivery events (17 metformin, 18 insulin).

[g]NICU admission is defined as a hospitalization event with a health specialty code for a Level II or Level III neonatal intensive care unit.

APPENDIX 6. TABLE OF MATERNAL AND NEONATAL OUTCOMES, STRATIFIED BY INFANT SEX

	Boys			Girls		
	Metformin N=2013 N (%)	**Insulin N=1838 N (%)**	**Adjusted Risk Difference (95% CI)[a]**	**Metformin N=1785 N (%)**	**Insulin N=1584 N (%)**	**Adjusted Risk Difference (95% CI)[a]**
Macrosomia[b,c]	237 (11.8)	246 (13.4)	-0.6 (-2.8, 1.6)	168 (9.4)	169 (10.7)	-1.0 (-3.2, 1.2)
Large-for-gestational-age[b,d]	265 (13.2)	343 (18.7)	-4.8 (-7.3, -2.3)	284 (15.9)	310 (19.6)	-2.0 (-4.8, 0.8)
Neonatal hypoglycemia	295 (14.7)	423 (23.0)	-6.3 (-9.0, -3.7)	248 (13.9)	331 (20.8)	-3.5 (-6.1, -0.9)
Respiratory distress	114 (5.7)	147 (8.0)	-1.2 (-2.9, 0.4)	80 (4.5)	97 (6.1)	-0.5 (-2.2, 1.1)
NICU admission[d]	244 (12.1)	521 (28.4)	-6.3 (-8.6, -4.1)	214 (12.0)	441 (27.7)	-5.3 (-7.5, -3.0)
Preterm birth[d]	153 (7.6)	175 (9.4)	-1.6 (-3.5, 0.2)	116 (6.5)	136 (8.5)	-1.5 (-3.4, 0.4)

[a]Inverse probability of treatment weights (IPTW) were estimated using sex-specific propensity score models containing the following covariates: maternal age, maternal ethnicity, maternal NZ deprivation decile score, parity, BMI prior to prescription initiation, smoking status, timing of GDM diagnosis and treatment, history of GDM and history of GDM treatment. The IPTW adjusted model uses imputed values for parity, BMI, and smoking status. These models are also adjusted for geographic region.
[b]Birthweight values were missing for 63 infants (29 metformin, 34 insulin).
[c]Macrosomia is defined as birthweight greater than 4,000g.
[d]Large-for-gestational-age is defined as a birthweight greater than the 90th percentile. These percentiles were calculated using a sample of New Zealand infants born between 2001 and 2015 with available data on birthweight and gestational age (n=850,506). These percentiles are additionally sex- and ethnicity-specific.
[d]NICU admission is defined as a hospitalization event with a health specialty code for a Level II or Level III neonatal intensive care unit.
[e]Preterm birth is defined as gestational age <37 weeks.

APPENDIX 7. FIGURE OF SEX-STRATIFIED ADJUSTED RISK DIFFERENCE ESTIMATES

RD (95% CI)

Macrosomia		
Boys	-0.6 (-2.8, 1.6)	
Girls	-1.0 (-3.2, 1.2)	
Large-for-Gestational-Age		
Boys	-4.8 (-7.3, -2.3)	
Girls	-2.0 (-4.8, 0.8)	
Neonatal Hypoglycemia		
Boys	-6.3 (-9.0, -3.7)	
Girls	-3.5 (-6.1, -0.9)	
Respiratory Distress		
Boys	-1.2 (-2.9, 0.4)	
Girls	-0.5 (-2.2, 1.1)	
NICU Admission		
Boys	-6.3 (-8.6, -4.1)	
Girls	-5.2 (-7.5, -3.0)	
Preterm Birth		
Boys	-1.6 (-3.5, 0.2)	
Girls	-1.6 (-3.4, 0.4)	

-10   -5   0   5   10
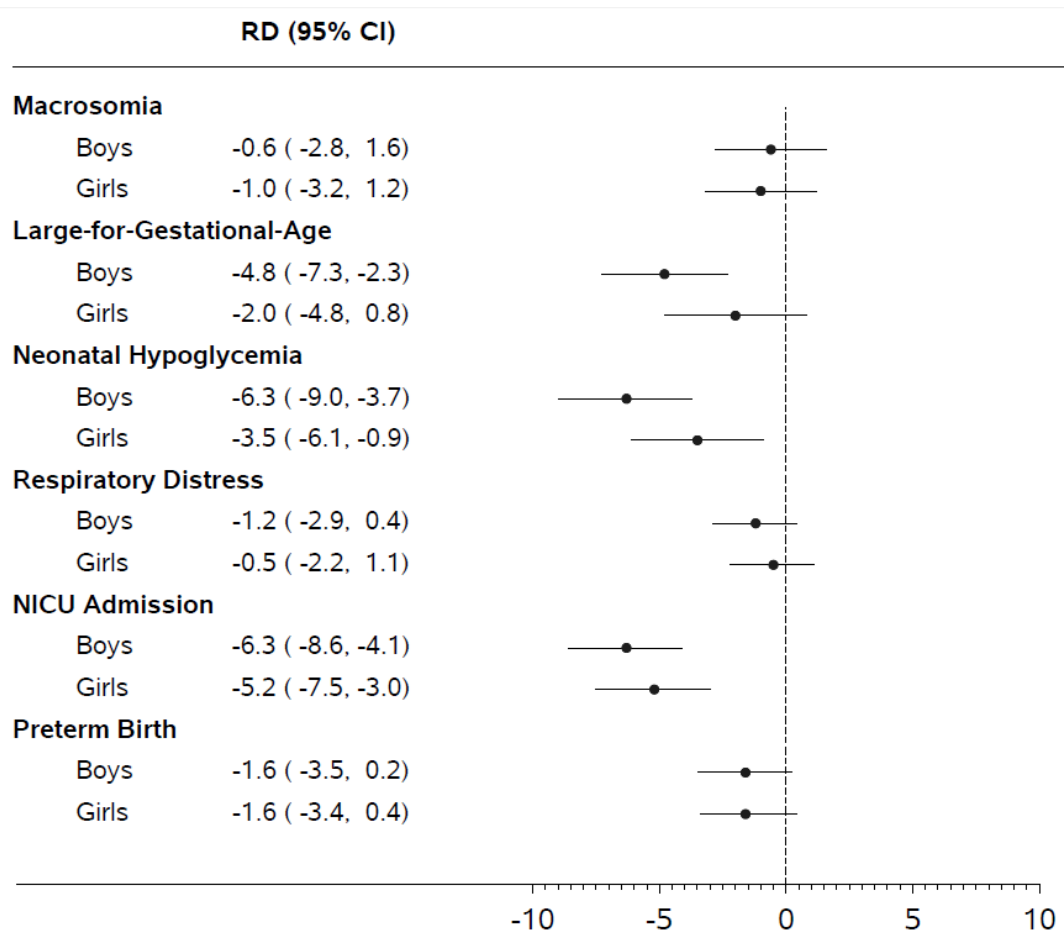
APPENDIX 8. COMPARISON OF PREVIOUS STUDY ESTIMATES

	RR (95% CI)	
Neonatal hypoglycemia		
Rowan et al 2008 (n=733)	0.81 (0.59,1.12)	
Tertti et al 2008 (n=90)	0.40 (0.20,0.90)	
Balani et al 2009* (n=200)	0.50 (0.24,1.06)	
Ijas et al 2011 (n=100)	0.70 (0.23,1.89)	
Goh et al 2011* (n=659)	0.91 (0.61,1.37)	
Niromanesh et al 2012 (n=160)	1.50 (0.30,8.70)	
Spaulonci et al 2013* (n=92)	0.30 (0.09,1.02)	
Tertti et al 2013 (n=216)	1.00 (0.50,1.90)	
Mesdaghinia et al 2013* (n=200)	0.67 (0.32,1.41)	
Current study (n=7229)	0.74 (0.66,0.82)	
Macrosomia		
Moore et al 2007* (n=63)	0.62 (0.15,2.23)	
Tertti et al 2008 (n=90)	1.30 (0.70,2.50)	
Balani et al 2009* (n=200)	0.79 (0.38,1.65)	
Ijas et al 2011 (n=100)	0.90 (0.40,1.91)	
Niromanesh et al 2012 (n=160)	0.40 (0.10,1.40)	
Tertti et al 2008 (n=90)	0.60 (0.20,1.90)	
Mesdaghinia et al 2013* (n=200)	0.61 (0.30,1.23)	
Current study (n=7205)	0.91 (0.80,1.06)	
NICU admission		
Moore et al 2007* (n=63)	0.48 (0.10,2.46)	
Rowan et al 2008 (n=733)	0.89 (0.66,1.19)	
Tertti et al 2013 (n=216)	0.90 (0.50,1.40)	
Balani et al 2009* (n=200)	0.32 (0.13,0.76)	
Ijas et al 2011 (n=100)	0.70 (0.29,1.60)	
Goh et al 2011* (n=659)	0.56 (0.37,0.84)	
Niromanesh et al 2012 (n=160)	2.50 (0.50,12.5)	
Tertti et al 2008 (n=90)	0.50 (0.20,1.10)	
Mesdaghinia et al 2013* (n=200)	0.42 (0.24,0.74)	
Current study (n=7229)	0.63 (0.57,0.70)	

0.0625   16

*Indicates a risk ratio calculated from reported proportions
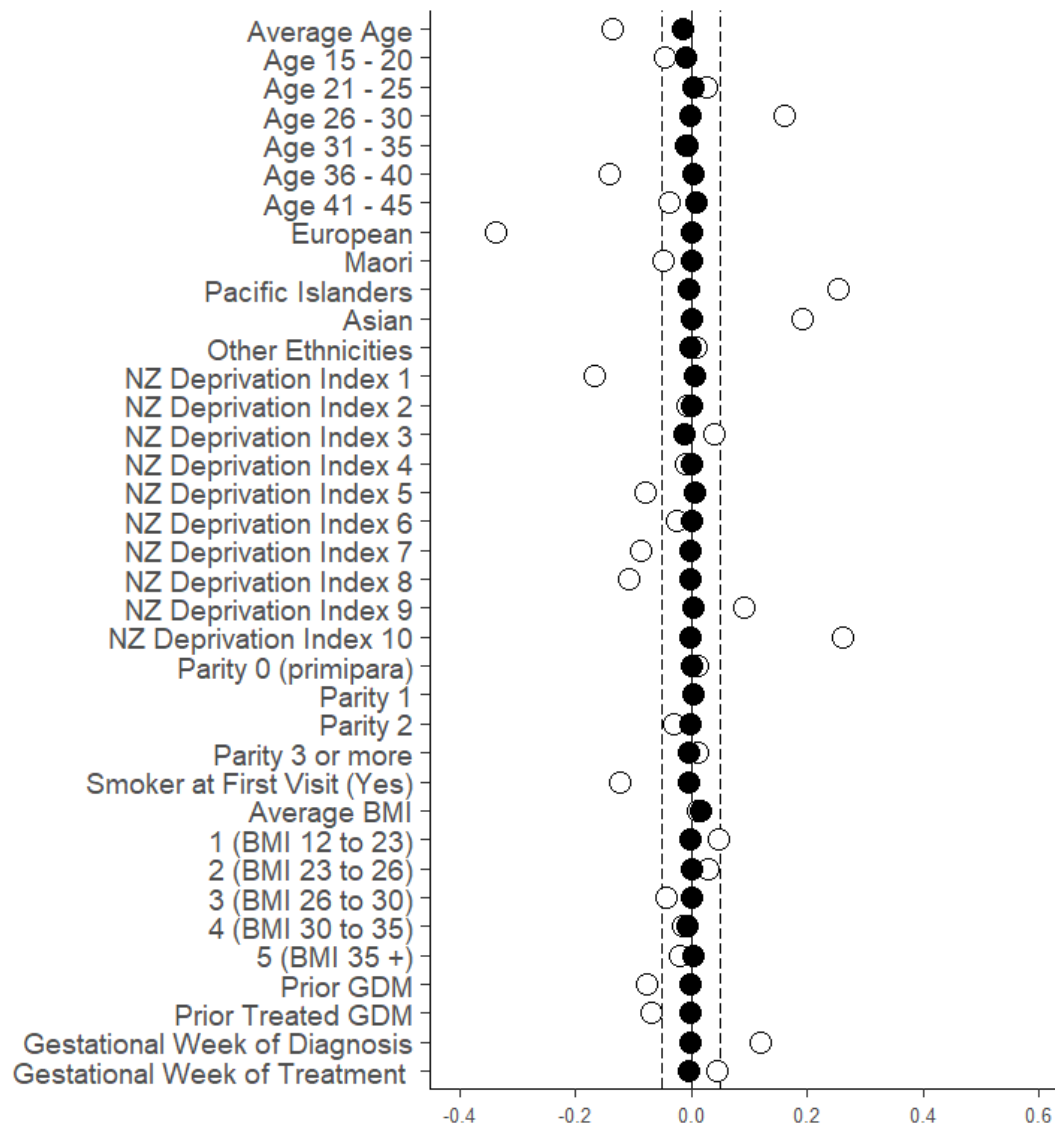
APPENDIX 10. TABLE OF SEX-STRATIFIED CHILD GROWTH AND DEVELOPMENT OUTCOMES FROM ADJUSTED LINEAR REGRESSION MODELS

	Boys			Girls		
	Metformin N=1082 Mean (SD)	Insulin N=1044 Mean (SD)	Adjusted Estimate (95% CI)	Metformin N=914 Mean (SD)	Insulin N=888 Mean (SD)	Adjusted Estimate (95% CI)
Child Growth Outcomes						
Weight z-score	0.8 (1.3)	0.9 (1.2)	-0.125(-0.272, 0.022)	0.7 (1.1)	0.6 (1.2)	0.014 (-0.122, 0.151)
Height z-score	0.4 (1.2)	0.3 (1.0)	0.006(-0.113, 0.125)	0.3 (1.1)	0.1 (1.1)	0.080 (-0.049, 0.209)
Weight for Length z-score	0.8 (1.4)	0.9 (1.3)	-0.125(-0.272, 0.022)	0.7 (1.3)	0.8 (1.3)	0.015 (-0.122, 0.151)
BMI z-score	0.8 (1.5)	1.0 (1.4)	-0.192(-0.352, -0.032)	0.7 (1.3)	0.8 (1.3)	-0.050 (-0.198, 0.098)
Strengths and Difficulties Questionnaire						
SDQ-P Difficulties Score[f,g]	7.2 (4.9)	6.9 (5.2)	0.235(-0.312, 0.783)	6.2 (4.4)	6.1 (4.4)	0.041 (-0.501, 0.583)
SDQ-T Difficulties Score [f,h]	4.7 (4.9)	4.3 (4.9)	0.266(-0.475, 1.007)	3.3. (3.7)	3.0 (3.9)	0.333 (-0.216, 0.882)
SDQ-P Prosocial Behaviour Score[g]	8.3 (1.8)	8.4 (1.8)	-0.062(-0.251, 0.127)	8.7 (1.5)	8.7 (1.6)	0.047 (-0.140, 0.235)
SDQ-T Prosocial Behaviour Score[h]	7.6 (2.4)	7.8 (2.5)	-0.140(-0.495, 0.215)	8.5 (2.0)	8.7 (2.1)	-0.276(-0.578, 0.025)

[a]Inverse probability of treatment weights (IPTW) were estimated using sex-specific propensity score models containing the following covariates: maternal age, maternal ethnicity, maternal NZ deprivation decile score, parity, BMI prior to prescription initiation, smoking status, timing of GDM diagnosis and treatment, history of GDM and history of GDM treatment. The IPTW adjusted model uses imputed values for parity, BMI, and smoking status. These models are also adjusted for geographic region.

[b]Weight was avaliable for 3,156 children (80.4%). Weight measurements were missing for 21.5% of insulin-exposed children and 17.9% of metformin exposed children. Weight was missing for 421 boys (24.7%) and 351 girls (24.2%).

[c]Z-scores and percentiles were calculated based on a WHO 2006 reference standard.

[d]Height was available for 3,154 children (80.3%). Height measurements were missing for 21.5% of insulin-exposed children and 17.9% of metformin exposed children. Height was missing for 422 boys (24.7%) and 352 girls (24.2%).
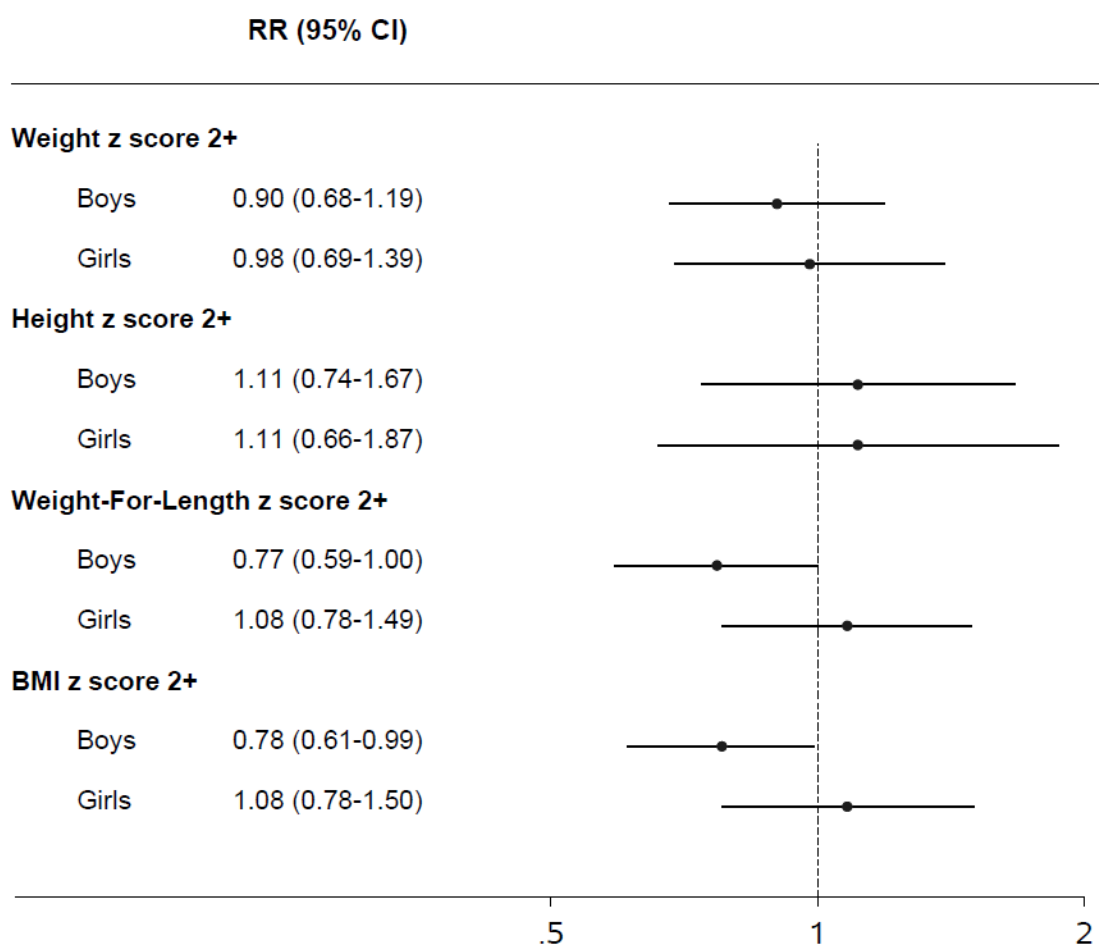
[e] Weight for length and BMI measurements were calculated for children with non-missing weight and height (N=3,154, 80.3%). Height was rounded to the nearest 0.5 for determining weight-for-length percentiles.

[f]SDQ Difficulties Scores were calculated by summing scores from four scales: emotional problems, conduct problems, hyperactivity/inattention, and peer relationship problems.

[g]SDQ-P values were available for 3,129 children (79.7%) (1,692 boys, 1,437 girls). Scores could not be calculated for 22.3% of children exposed to insulin, and 18.4% of children exposed to metformin.

[h]SDQ-T values were available for 1,681children (42.8%) (888 boys, 793 girls). Scores could not be calculated for 47.1% of children exposed to insulin, and 38.7% of children exposed to metformin.

APPENDIX 11. FIGURE OF SEX-STRATIFIED ADJUSTED RISK RATIO ESTIMATES FOR SELECTED CHILD GROWTH OUTCOMES

RR (95% CI)

Weight z score 2+

Boys	0.90 (0.68-1.19)
Girls	0.98 (0.69-1.39)

Height z score 2+

Boys	1.11 (0.74-1.67)
Girls	1.11 (0.66-1.87)

Weight-For-Length z score 2+

Boys	0.77 (0.59-1.00)
Girls	1.08 (0.78-1.49)

BMI z score 2+

Boys	0.78 (0.61-0.99)
Girls	1.08 (0.78-1.50)

.5   1   2

REFERENCES

1.   Barbour LA, McCurdy CE, Hernandez TL, Kirwan JP, Catalano PM, Friedman JE. Cellular mechanisms for insulin resistance in normal pregnancy and gestational diabetes. Diabetes Care. 2007;30 Suppl 2:S112-9.

2.   International Diabetes Federation. IDF Diabetes Atlas, Eighth Edition. Brussels, Belgium: International Diabetes Federation; 2017.

3.   NHS UK. Gestational diabetes - causes London, UK2014 [Available from: http://www.nhs.uk/Conditions/gestational-diabetes/Pages/Causes.aspx.

4.   Petry CJ. Gestational diabetes: risk factors and recent advances in its genetics and treatment. Br J Nutr. 2010;104(6):775-87.

5.   DeSisto CL, Kim SY, Sharma AJ. Prevalence estimates of gestational diabetes mellitus in the United States, Pregnancy Risk Assessment Monitoring System (PRAMS), 2007-2010. Prev Chronic Dis. 2014;11:E104.

6.   Getahun D, Nath C, Ananth CV, Chavez MR, Smulian JC. Gestational diabetes in the United States: temporal trends 1989 through 2004. Am J Obstet Gynecol. 2008;198(5):525 e1-5.

7.   National Institute for Health and Care Excellence. Diabetes in pregnancy: Management of diabetes and its complications from pre-conception to the postnatal period London, UK: National Institute for Health and Care Excellence; 2015 [Available from: https://www.nice.org.uk/guidance/cg63.

8.   Australian Institute of Health and Welfare. Incidence of diabetes Canberra, Australia: Australian Institute of Health and Welfare,; 2015 [Available from: http://www.aihw.gov.au/diabetes-indicators/incidence/.

9.   Coppell KJ, Mann JI, Williams SM, Jo E, Drury PL, Miller JC, et al. Prevalence of diagnosed and undiagnosed diabetes and prediabetes in New Zealand: findings from the 2008/09 Adult Nutrition Survey. N Z Med J. 2013;126(1370):23-42.

10.   Ministry of Health. Screening, Diagnosis, and Management of Gestational Diabetes in New Zealand: A clinical practice guideline. Wellington, New Zealand: Ministry of Health, New Zealand; 2014.

11.   Seshiah V, Balaji V, Balaji MS, Sanjeevi CB, Green A. Gestational diabetes mellitus in India. J Assoc Physicians India. 2004;52:707-11.

12.   Crowther CA, Hiller JE, Moss JR, McPhee AJ, Jeffries WS, Robinson JS. Effect of treatment of gestational diabetes mellitus on pregnancy outcomes. The New England journal of medicine. 2005;352(24):2477-86.

13.   Horvath K, Koch K, Jeitler K, Matyas E, Bender R, Bastian H, et al. Effects of treatment in women with gestational diabetes mellitus: systematic review and meta-analysis. BMJ (Clinical research ed). 2010;340:c1395.

14.   Landon MB, Spong CY, Thom E, Carpenter MW, Ramin SM, Casey B, et al. A multicenter, randomized trial of treatment for mild gestational diabetes. The New England journal of medicine. 2009;361(14):1339-48.

15.   Coustan DR. Pharmacological management of gestational diabetes: an overview. Diabetes Care. 2007;30 Suppl 2:S206-8.

16.   Kelley KW, Carroll DG, Meyer A. A review of current treatment strategies for gestational diabetes mellitus. Drugs Context. 2015;4:212282.

17.   Gill GV, Yudkin JS, Keen H, Beran D. The insulin dilemma in resource-limited countries. A way forward? Diabetologia. 2011;54(1):19-24.

18.   Feghali MN, Miodovnik M, Umans JG. Pharmacologic considerations affecting hypoglycemic therapy during pregnancy. In: Langer O, editor. The Diabetes in Pregnancy Dilemma: Leading Change with Proven Solutions, 2e. Shelton, CT: PMPH-USA; 2015. p. 17-24.

19.   Eyal S, Easterling TR, Carr D, Umans JG, Miodovnik M, Hankins GD, et al. Pharmacokinetics of metformin during pregnancy. Drug Metab Dispos. 2010;38(5):833-40.

20.   Vanky E, Zahlsen K, Spigset O, Carlsen SM. Placental passage of metformin in women with polycystic ovary syndrome. Fertil Steril. 2005;83(5):1575-8.

21.   Lautatzis ME, Goulis DG, Vrontakis M. Efficacy and safety of metformin during pregnancy in women with gestational diabetes mellitus or polycystic ovary syndrome: a systematic review. Metabolism. 2013;62(11):1522-34.

22.   Moore LE, Briery CM, Clokey D, Martin RW, Williford NJ, Bofill JA, et al. Metformin and insulin in the management of gestational diabetes mellitus: preliminary results of a comparison. J Reprod Med. 2007;52(11):1011-5.

23.   Silva JC, Fachin DR, Coral ML, Bertini AM. Perinatal impact of the use of metformin and glyburide for the treatment of gestational diabetes mellitus. Journal of perinatal medicine. 2012;40(3):225-8.

24.   Moore LE, Clokey D, Rappaport VJ, Curet LB. Metformin compared with glyburide in gestational diabetes: a randomized controlled trial. Obstet Gynecol. 2010;115(1):55-9.

25.   Silva JC, Pacheco C, Bizato J, de Souza BV, Ribeiro TE, Bertini AM. Metformin compared with glyburide for the management of gestational diabetes. Int J Gynaecol Obstet. 2010;111(1):37-40.

26.   George A, Mathews JE, Sam D, Beck M, Benjamin SJ, Abraham A, et al. Comparison of neonatal outcomes in women with gestational diabetes with moderate hyperglycaemia on metformin or glibenclamide--a randomised controlled trial. Aust N Z J Obstet Gynaecol. 2015;55(1):47-52.

27.   Baptiste-Roberts K, Nicholson WK, Wang NY, Brancati FL. Gestational diabetes and subsequent growth patterns of offspring: the National Collaborative Perinatal Project. Matern Child Health J. 2012;16(1):125-32.

28.   Gillman MW, Rifas-Shiman S, Berkey CS, Field AE, Colditz GA. Maternal gestational diabetes, birth weight, and adolescent obesity. Pediatrics. 2003;111(3):e221-6.

29.   Kim C, Newton KM, Knopp RH. Gestational diabetes and the incidence of type 2 diabetes: a systematic review. Diabetes Care. 2002;25(10):1862-8.

30.   Pettitt DJ, McKenna S, McLaughlin C, Patterson CC, Hadden DR, McCance DR. Maternal glucose at 28 weeks of gestation is not associated with obesity in 2-year-old offspring: the Belfast Hyperglycemia and Adverse Pregnancy Outcome (HAPO) family study. Diabetes Care. 2010;33(6):1219-23.

31.   Thaware PK, McKenna S, Patterson CC, Hadden DR, Pettitt DJ, McCance DR. Untreated Mild Hyperglycemia During Pregnancy and Anthropometric Measures of Obesity in Offspring at Age 5-7 Years. Diabetes Care. 2015;38(9):1701-6.

32.   Deierlein AL, Siega-Riz AM, Chantala K, Herring AH. The association between maternal glucose concentration and child BMI at age 3 years. Diabetes Care. 2011;34(2):480-4.

33.   Bennett WL, Robinson KA, Saldanha IJ, Wilson LM, Nicholson WK. High priority research needs for gestational diabetes mellitus. J Womens Health (Larchmt). 2012;21(9):925-32.

34.   Feig DS, Moses RG. Metformin Therapy During Pregnancy: Good for the goose and good for the gosling too? Diabetes Care. 2011;34(10):2329-30.

35.   Nieuwenhuis-Ruifrok AE, Kuchenbecker WK, Hoek A, Middleton P, Norman RJ. Insulin sensitizing drugs for weight loss in women of reproductive age who are overweight or obese: systematic review and meta-analysis. Hum Reprod Update. 2009;15(1):57-68.

36.   Jo H, Schieve LA, Sharma AJ, Hinkle SN, Li R, Lind JN. Maternal prepregnancy body mass index and child psychosocial development at 6 years of age. Pediatrics. 2015;135(5):e1198-209.

37.   Fraser A, Nelson SM, Macdonald-Wallis C, Lawlor DA. Associations of existing diabetes, gestational diabetes, and glycosuria with offspring IQ and educational attainment: the Avon Longitudinal Study of Parents and Children. Exp Diabetes Res. 2012;2012:963735.

38.   Kaiser JR, Bai S, Gibson N, Holland G, Lin TM, Swearingen CJ, et al. Association Between Transient Newborn Hypoglycemia and Fourth-Grade Achievement Test Proficiency: A Population-Based Study. JAMA Pediatr. 2015;169(10):913-21.

39.   McKinlay CD, Alsweiler JM, Anstice NS, et al. Association of neonatal glycemia with neurodevelopmental outcomes at 4.5 years. JAMA Pediatrics. 2017;171(10):972-83.

40. McKinlay CJ, Alsweiler JM, Ansell JM, Anstice NS, Chase JG, Gamble GD, et al. Neonatal Glycemia and Neurodevelopmental Outcomes at 2 Years. The New England journal of medicine. 2015;373(16):1507-18.

41. Rowan JA, Hague WM, Gao W, Battin MR, Moore MP. Metformin versus Insulin for the Treatment of Gestational Diabetes. New England Journal of Medicine. 2008;358(19):2003-15.

42. Barrett HL, Gatford KL, Houda CM, De Blasio MJ, McIntyre HD, Callaway LK, et al. Maternal and neonatal circulating markers of metabolic and cardiovascular risk in the metformin in gestational diabetes (MiG) trial: responses to maternal metformin versus insulin treatment. Diabetes Care. 2013;36(3):529-36.

43. Rowan JA, Rush EC, Obolonkin V, Battin M, Wouldes T, Hague WM. Metformin in gestational diabetes: the offspring follow-up (MiG TOFU): body composition at 2 years of age. Diabetes Care. 2011;34(10):2279-84.

44. Battin MR, Obolonkin V, Rush E, Hague W, Coat S, Rowan J. Blood pressure measurement at two years in offspring of women randomized to a trial of metformin for GDM: follow up data from the MiG trial. BMC Pediatr. 2015;15:54.

45. Wouldes TA, Battin M, Coat S, Rush EC, Hague WM, Rowan JA. Neurodevelopmental outcome at 2 years in offspring of women randomised to metformin or insulin treatment for gestational diabetes. Arch Dis Child Fetal Neonatal Ed. 2016.

46. Ijas H, Vaarasmaki M, Morin-Papunen L, Keravuo R, Ebeling T, Saarela T, et al. Metformin should be considered in the treatment of gestational diabetes: a prospective randomised study. BJOG. 2011;118(7):880-5.

47. Ijas H, Vaarasmaki M, Saarela T, Keravuo R, Raudaskoski T. A follow-up of a randomised study of metformin and insulin in gestational diabetes mellitus: growth and development of the children at the age of 18 months. BJOG. 2015;122(7):994-1000.

48. Mesdaghinia E, Samimi M, Homaei Z, Saberi F, Moosavi SG, Yaribakht M. Comparison of newborn outcomes in women with gestational diabetes mellitus treated with metformin or insulin: a randomised blinded trial. Int J Prev Med. 2013;4(3):327-33.

49. Ruholamin S, Eshaghian S, Allame Z. Neonatal outcomes in women with gestational diabetes mellitus treated with metformin in compare with insulin: A randomized clinical trial. J Res Med Sci. 2014;19(10):970-5.

50. Spaulonci CP, Bernardes LS, Trindade TC, Zugaib M, Francisco RP. Randomized trial of metformin vs insulin in the management of gestational diabetes. Am J Obstet Gynecol. 2013;209(1):34 e1-7.

51. Waheed S, Malik FP, Mazhar SB. Efficacy of metformin versus insulin in the management of pregnancy with diabetes. J Coll Physicians Surg Pak. 2013;23(12):866-9.

52.   Balsells M, Garcia-Patterson A, Sola I, Roque M, Gich I, Corcoy R. Glibenclamide, metformin, and insulin for the treatment of gestational diabetes: a systematic review and meta-analysis. BMJ (Clinical research ed). 2015;350:h102.

53.   Poolsup N, Suksomboon N, Amin M. Efficacy and safety of oral antidiabetic drugs in comparison to insulin in treating gestational diabetes mellitus: a meta-analysis. PloS one. 2014;9(10):e109985.

54.   Goh JE, Sadler L, Rowan J. Metformin for gestational diabetes in routine clinical practice. Diabetic medicine : a journal of the British Diabetic Association. 2011;28(9):1082-7.

55.   Tertti K, Ekblad U, Vahlberg T, Ronnemaa T. Comparison of metformin and insulin in the treatment of gestational diabetes: a retrospective, case-control study. Rev Diabet Stud. 2008;5(2):95-101.

56.   Balani J, Hyer SL, Rodin DA, Shehata H. Pregnancy outcomes in women with gestational diabetes treated with metformin or insulin: a case-control study. Diabetic medicine : a journal of the British Diabetic Association. 2009;26(8):798-802.

57.   Castillo WC, Boggess K, Stürmer T, Brookhart MA, Benjamin DK, Funk MJ. Association of Adverse Pregnancy Outcomes With Glyburide vs Insulin in Women With Gestational Diabetes. JAMA pediatrics. 2015;169(5):452-8.

58.   Cheng YW, Chung JH, Kurbisch-Block I, Inturrisi M, Shafer S, Caughey AB. Gestational weight gain and gestational diabetes mellitus: perinatal outcomes. Obstet Gynecol. 2008;112(5):1015-22.

59.   Langer O, Conway DL, Berkus MD, Xenakis EM, Gonzales O. A comparison of glyburide and insulin in women with gestational diabetes mellitus. The New England journal of medicine. 2000;343(16):1134-8.

60.   Ramos GA, Jacobson GF, Kirby RS, Ching JY, Field DR. Comparison of glyburide and insulin for the management of gestational diabetics with markedly elevated oral glucose challenge test and fasting hyperglycemia. Journal of perinatology : official journal of the California Perinatal Association. 2007;27(5):262-7.

61.   Langer O. Oral hypoglycemic agents: do the ends justify the means? Maternal Health, Neonatology and Perinatology. 2015;1(1):1.

62.   ACOG. Practice Bulletin No. 137: Gestational diabetes mellitus. Obstet Gynecol. 2013;122(2 Pt 1):406-16.

63.   ACOG. Practice Bulletin No. 180: Gestational Diabetes Mellitus. Obstet Gynecol. 2017;130(1):e17-e37.

64.   UNC Center for Maternal and Infant Health. Gestational Diabetes Class A2: Antepartum Management (glyburide) Chapel Hill, NC: UNC School of Medicine; 2004 [updated July 2004. Available from: http://www.mombaby.org/PDF/obalg_gestdiab_classA2_glyburide.pdf.

65. Group Health Cooperative. Gestational Diabetes Screening and Treatment Guideline Seattle, WA: Group Health Cooperative; 2011 [Available from: https://www.ghc.org/all-sites/guidelines/gestationalDiabetes.pdf.

66. Thompson D, Berger H, Feig D, Gagnon R, Kader T, Keely E, et al. Diabetes and pregnancy. Can J Diabetes. 2013;37 Suppl 1:S168-83.

67. Medsafe. New Zealand Data Sheet: Metformin Wellington, NZ: New Zealand Medicines and Medical Devices Safety Authority; 2011 [Available from: http://www.medsafe.govt.nz/profs/datasheet/m/metformintab.pdf.

68. National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). Gestational Diabetes Bethesda, MD: National Diabetes Information Clearinghouse (NDIC); 2017 [Available from: https://www.niddk.nih.gov/health-information/diabetes/overview/what-is-diabetes/gestational.

69. 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):1047-52.

70. Landon MB, Mele L, Spong CY, Carpenter MW, Ramin SM, Casey B, et al. The relationship between maternal glycemia and perinatal outcome. Obstet Gynecol. 2011;117(2 Pt 1):218-24.

71. ADA. 13. Management of Diabetes in Pregnancy. Diabetes Care. 2017;40(Supplement 1):S114-S9.

72. Alwan N, Tuffnell DJ, West J. Treatments for gestational diabetes. Cochrane Database Syst Rev. 2009(3):CD003395.

73. Ministry of Health. National collections and surveys Wellington, NZ: New Zealand Ministry of Health; 2012 [Available from: http://www.health.govt.nz/nz-health-statistics/national-collections-and-surveys/collections.

74. WHO. New Zealand Health System Review. Geneva, Switzerland: World Health Organization; 2014.

75. NICHD. When is a cesarean delivery necessary & what are the risks? Bethesda, MD: Eunice Kennedy Shriver National Institute of Child Health and Human Development; 2012 [Available from: http://www.nichd.nih.gov/health/topics/obstetrics/conditioninfo/pages/risks.aspx.

76. Ministry of Health. Maternity Tables 2013. Wellington, NZ: New Zealand Ministry of Health, Services A; 2015 October 6, 2015.

77. NICHD. What is induction of labor? Bethesa, MD: Eunice Kennedy Shriver National Institute for Child Health and Human Development; 2014 [Available from: https://www.nichd.nih.gov/health/topics/labor-delivery/topicinfo/Pages/induction.aspx.

78. Eskandar O, Shet D. Risk factors for 3rd and 4th degree perineal tear. J Obstet Gynaecol. 2009;29(2):119-22.

79.   Dahlen HG, Ryan M, Homer CS, Cooke M. An Australian prospective cohort study of risk factors for severe perineal trauma during childbirth. Midwifery. 2007;23(2):196-203.

80.   NICHD. Preeclampsia and Eclampsia: Condition Information Rockville, MD: Eunice Kennedy Shriver National Institute of Child Health and Human Development; 2013 [Available from: https://www.nichd.nih.gov/health/topics/preeclampsia/conditioninfo/Pages/default.aspx.

81.   Bryson CL, Ioannou GN, Rulyak SJ, Critchlow C. Association between Gestational Diabetes and Pregnancy-induced Hypertension. American Journal of Epidemiology. 2003;158(12):1148-53.

82.   Sacks DA, Black MH, Li X, Montoro MN, Lawrence JM. Adverse Pregnancy Outcomes Using The International Association of the Diabetes and Pregnancy Study Groups Criteria: Glycemic Thresholds and Associated Risks. Obstet Gynecol. 2015;126(1):67-73.

83.   Yogev Y, Xenakis EM, Langer O. The association between preeclampsia and the severity of gestational diabetes: the impact of glycemic control. Am J Obstet Gynecol. 2004;191(5):1655-60.

84.   Hadfield RM, Lain SJ, Simpson JM, Ford JB, Raynes-Greenow CH, Morris JM, et al. Are babies getting bigger? An analysis of birthweight trends in New South Wales, 1990-2005. Med J Aust. 2009;190(6):312-5.

85.   Harris DL, Weston PJ, Battin MR, Harding JE. A survey of the management of neonatal hypoglycaemia within the Australian and New Zealand Neonatal Network. J Paediatr Child Health. 2014;50(10):E55-62.

86.   Rozance PJ. Update on neonatal hypoglycemia. Curr Opin Endocrinol Diabetes Obes. 2014;21(1):45-50.

87.   Vignoles P, Gire C, Mancini J, Bretelle F, Boubli L, Janky E, et al. Gestational diabetes: a strong independent risk factor for severe neonatal respiratory failure after 34 weeks. Arch Gynecol Obstet. 2011;284(5):1099-104.

88.   Baxley EG, Gobbo RW. Shoulder dystocia. Am Fam Physician. 2004;69(7):1707-14.

89.   Ovesen PG, Jensen DM, Damm P, Rasmussen S, Kesmodel US. Maternal and neonatal outcomes in pregnancies complicated by gestational diabetes. a nation-wide study. J Matern Fetal Neonatal Med. 2015;28(14):1720-4.

90.   Mehta SH, Sokol RJ. Shoulder dystocia: risk factors, predictability, and preventability. Semin Perinatol. 2014;38(4):189-93.

91.   Politi S, D'Emidio L, Cignini P, Giorlandino M, Giorlandino C. Shoulder dystocia: an Evidence-Based approach. Journal of prenatal medicine. 2010;4(3):35-42.

92.   Wilcox AJ. Fertility and pregnancy: an epidemiologic perspective: Oxford University Press, USA; 2010.

93.   Holman N, Bell R, Murphy H, Maresh M. Women with pre-gestational diabetes have a higher risk of stillbirth at all gestations after 32 weeks. Diabetic medicine : a journal of the British Diabetic Association. 2014;31(9):1129-32.

94.   Reddy UM, Laughon SK, Sun L, Troendle J, Willinger M, Zhang J. Prepregnancy risk factors for antepartum stillbirth in the United States. Obstet Gynecol. 2010;116(5):1119-26.

95.   Robson S, Nolan C. Diabetes and stillbirth. O&G Magazine. 2013.

96.   Hutcheon JA, Kuret V, Joseph KS, Sabr Y, Lim K. Immortal time bias in the study of stillbirth risk factors: the example of gestational diabetes. Epidemiology. 2013;24(6):787-90.

97.   Watson D, Rowan J, Neale L, Battin MR. Admissions to neonatal intensive care unit following pregnancies complicated by gestational or type 2 diabetes. Aust N Z J Obstet Gynaecol. 2003;43(6):429-32.

98.   Rajput N, Tuohy P, Mishra S, Smith A, Taylor B. Overweight and obesity in 4-5-year-old children in New Zealand: results from the first 4 years (2009-2012) of the B4School Check programme. J Paediatr Child Health. 2015;51(3):334-43.

99.   Ministry of Health. The B4 School Check: A handbook for practitioners. Wellington, New Zealand: New Zealand Ministry of Health; 2008.

100.   WHO. Global Database on Child Growth and Malnutrition Geneva, Switzerland: World Health Organization; 2018 [Available from: http://www.who.int/nutgrowthdb/about/introduction/en/index5.html.

101.   Goodman R. The Strengths and Difficulties Questionnaire: a research note. J Child Psychol Psychiatry. 1997;38(5):581-6.

102.   Ministry of Health. Behaviour assessment: The Strengths and Difficulties Questionnaire Wellington, NZ: New Zealand Ministry of Health; 2013 [Available from: http://well-child.book.health.govt.nz/Section+5+B4+School+Check/Behaviour+assessment+The+Strengths+and+Difficulties+Questionnaire.

103.   Goodman R, Ford T, Richards H, Gatward R, Meltzer H. The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. J Child Psychol Psychiatry. 2000;41(5):645-55.

104.   Ministry of Health. A validation and norming study of the strengths and difficulties questionnaire in the New Zealand context Wellington, New Zealand: New Zealand Ministry of Health; 2015 [Available from: https://www.health.govt.nz/publication/validation-and-norming-study-strengths-and-difficulties-questionnaire-new-zealand-context.

105. Rajput N, McKinlay C, Purdie G, Filipovska J, Battin M, Patel H, et al. Community-based screening to detect school readiness problems in very preterm children. J Paediatr Child Health. 2017.

106. Ministry of Health. NZDep2013 Index of Deprivation Wellington, NZ: Department of Public Health, University of Otago; 2014 [Available from: http://www.health.govt.nz/publication/nzdep2013-index-deprivation.

107. Atkinson J, Salmond C, Crampton P. NZDep2013 Index of Deprivation. Wellington, New Zealand: University of Otago; 2014.

108. Stuart EA, Azur M, Frangakis C, Leaf P. Multiple imputation with large data sets: a case study of the Children's Mental Health Initiative. Am J Epidemiol. 2009;169(9):1133-9.

109. Tchetgen Tchetgen EJ, Glymour MM, Weuve J, Robins J. Specifying the correlation structure in inverse-probability- weighting estimation for repeated measures. Epidemiology. 2012;23(4):644-6.

110. Brown J, Martis R, Hughes B, Rowan J, Crowther CA. Oral anti-diabetic pharmacological therapies for the treatment of women with gestational diabetes. Cochrane Database Syst Rev. 2017;1:CD011967.

111. Ministry of Health. National collections and surveys Wellington, NZ: New Zealand Ministry of Health; 2015 [Available from: http://www.health.govt.nz/nz-health-statistics/national-collections-and-surveys?mega=Health%20statistics&title=National%20collections%20and%20surveys.

112. Hoffman L, Nolan C, Wilson JD, Oats JJ, Simmons D. Gestational diabetes mellitus--management guidelines. The Australasian Diabetes in Pregnancy Society. Med J Aust. 1998;169(2):93-7.

113. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol. 2008;168(6):656-64.

114. Lyons MR, Peterson LR, McGill JB, Herrero P, Coggan AR, Saeed IM, et al. Impact of sex on the heart's metabolic and functional responses to diabetic therapies. Am J Physiol Heart Circ Physiol. 2013;305(11):H1584-91.

115. Schutt M, Zimmermann A, Hood R, Hummel M, Seufert J, Siegel E, et al. Gender-specific Effects of Treatment with Lifestyle, Metformin or Sulfonylurea on Glycemic Control and Body Weight: A German Multicenter Analysis on 9 108 Patients. Exp Clin Endocrinol Diabetes. 2015;123(10):622-6.

116. Lund JL, Richardson DB, Sturmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. Curr Epidemiol Rep. 2015;2(4):221-8.

117. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. Am J Epidemiol. 2003;158(9):915-20.

118. American Diabetes Association. Management of diabetes in pregnancy. Diabetes Care. 2015;38 Suppl:S77-9.

119. Tertti K, Eskola E, Ronnemaa T, Haataja L. Neurodevelopment of Two-Year-Old Children Exposed to Metformin and Insulin in Gestational Diabetes Mellitus. J Dev Behav Pediatr. 2015;36(9):752-7.

120. Figueroa Gray M, Hsu C, Kiel L, Dublin S. "It's a Very Big Burden on Me": Women's Experiences Using Insulin for Gestational Diabetes. Matern Child Health J. 2017;21(8):1678-85.

121. Ministry of Health. B4 School Check Wellington, NZ: New Zealand Ministry of Health; 2015 [Available from: http://www.health.govt.nz/our-work/life-stages/child-health/b4-school-check.

122. Kapral N, Miller SE, Scharf RJ, Gurka MJ, DeBoer MD. Associations between birthweight and overweight and obesity in school-age children. Pediatr Obes. 2017.