Association of maternal lipid levels with birth weight and cord blood insulin

Wang, Jingya; Kuang, Yashu; Shen, Songying; Price, Malcolm James; Lu, Jinhua; Sattar, Naveed; He, Jianrong; Pittavino, Marta; Xia, Huimin; Thomas, G Neil; Qiu, Xiu; Cheng, Kar Keung; Nirantharakumar, Krishnarajah

DOI:
10.1136/bmjopen-2022-064122

BMJ Open

Association of maternal lipid levels with birth weight and cord blood insulin: a Bayesian network analysis

Jingya Wang ,[1,2] Yashu Kuang,[1,3,4] Songying Shen,[1,3,4] Malcolm James Price,[2] Jinhua Lu,[1,3,4] Naveed Sattar ,[5] Jianrong He,[1,3,4] Marta Pittavino,[6] Huimin Xia,[1,4] G.Neil Thomas,[2] Xiu Qiu,[1,4,7] Kar Keung Cheng,[2] Krishnarajah Nirantharakumar [2]

To cite: Wang J, Kuang Y, Shen S, *et al*. Association of maternal lipid levels with birth weight and cord blood insulin: a Bayesian network analysis. *BMJ Open* 2022;**12**:e064122. doi:10.1136/bmjopen-2022-064122

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-064122).

XQ and KN are joint senior authors.

Received 09 May 2022
Accepted 21 November 2022

For numbered affiliations see end of article.

Correspondence to
Dr Xiu Qiu; qxiu0161@163.com,
Dr Krishnarajah Nirantharakumar; k.nirantharan@bham.ac.uk and
Dr Jingya Wang; J.Wang.6@bham.ac.uk

ABSTRACT

Objective To assess the independent association of maternal lipid levels with birth weight and cord blood insulin (CBI) level.

Setting The Born in Guangzhou Cohort Study, Guangzhou, China.

Participants Women who delivered between January 2015 and June 2016 and with umbilical cord blood retained were eligible for this study. Those with prepregnancy health conditions, without an available fasting blood sample in the second trimester, or without demographic and glycaemic information were excluded. After random selection, data from 1522 mother–child pairs were used in this study.

Exposures and outcome measures Additive Bayesian network analysis was used to investigate the interdependency of lipid profiles with other metabolic risk factors (prepregnancy body mass index (BMI), fasting glucose and early gestational weight gain) in association with birth weight and CBI, along with multivariable linear regression models.

Results In multivariable linear regressions, maternal triglyceride was associated with increased birth weight (adjusted β=67.46, 95% CI 41.85 to 93.06 g per mmol/L) and CBI (adjusted β=0.89, 95% CI 0.06 to 1.72 µU/mL per mmol/L increase), while high-density lipoprotein cholesterol was associated with decreased birth weight (adjusted β=−45.29, 95% CI −85.49 to −5.09 g per mmol/L). After considering the interdependency of maternal metabolic risk factors in the Network analysis, none of the maternal lipid profiles was independently associated with birth weight and CBI. Instead, prepregnancy BMI was the global strongest factor for birth weight and CBI directly and indirectly.

Conclusions Gestational dyslipidaemia appears to be secondary to metabolic dysfunction with no clear association with metabolic adverse outcomes in neonates. Maternal prepregnancy overweight/obesity appears the most influential upstream metabolic risk factor for both maternal and neonatal metabolic health; these data imply weight management may need to be addressed from the preconception period and during early pregnancy.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ To the best of our knowledge, this is the first study to date focusing on maternal lipids during pregnancy that weighs key maternal metabolic risk factors systematically using an innovative additive Bayesian network analytical tool.
⇒ Given the practical constraints, maternal insulin resistance status during pregnancy, as a potential significant explanatory factor, was not included in the current analysis, although maternal fasting glucose level was included.
⇒ Due to the limited ability of detecting feedback loops, partial results of Bayesian network analysis should be interpreted with a degree of caution.
⇒ Risk of bias due to self-reported prepregnancy weight can not be excluded.

INTRODUCTION

Unfavourable fetal growth is often considered to attribute to an adverse intrauterine nutritional exposure that is largely dependent on maternal metabolic status.[1 2] Other than those established maternal metabolic risk factors (prepregnancy body mass index (BMI), gestational weight gain (GWG) and gestational hyperglycaemia) for adverse birthweight outcomes,[3–5] maternal lipid levels during pregnancy are increasingly recognised as ignored risk factors for adverse pregnancy outcomes in the last few years.[6–8] With the evidence from 42 observational studies using classic statistical methods, our recent systematic review found that increased maternal triglycerides and decreased high-density lipoprotein cholesterol (HDL-C) were positively associated with high birth weight.[9]

However, the majority of previous studies only focused on the association between maternal lipid levels and birth weight, but ignored the underlying interdependency between lipids and other metabolic factors.[10 11] Using data from 400 The Hyperglycemia and Adverse Pregnancy Outcome Study Caucasian mother–child pairs, two network analyses investigated the associations of maternal metabolites with maternal BMI, glucose level, birth weight or cord blood c-peptide.[2 12] These studies only considered the interdependency of maternal metabolites

and their association with a single phenotype rather than putting a broad range of metabolic traits and phenotypes together to draw a panoramic picture. Understanding how maternal lipid profiles influence neonatal metabolic health taking into consideration the clinical metabolic network into account is crucial to understand any potential underlying mechanism and future interventional studies.

In a birth cohort from Chinese pregnant women and their newborn babies, this study aimed to investigate the association of maternal lipid levels with birth weight and cord blood insulin (CBI) (indicators of neonatal developmental and metabolic conditions), with the consideration of the interdependency of maternal lipids with other maternal metabolic risk factors, including prepregnancy BMI, fasting glucose and GWG.

MATERIALS AND METHODS

Participants

This study is part of the Born in Guangzhou Cohort Study (BIGCS), a prospective cohort study conducted in the Guangzhou Women and Children's Medical Centre (GWCMC). The design and methods of BIGCS have been described previously.[13] In brief, eligible women with Chinese nationality, living in Guangzhou who are <20 weeks gestation and who intend to deliver at one of the two GWCMC campuses were recruited into BIGCS. This study was conducted in a planned subgroup of BIGCS in whom maternal and cord blood were analysed for metabolic parameters separately. Pregnant women attending BIGCS with a singleton pregnancy who delivered at GWCMC between January 2015 and June 2016 and had umbilical cord blood retained are eligible for this study. Women were excluded if (1) maternal blood samples were unavailable at 14–27 gestation week; (2) no records of maternal fasting glucose at 20–28 gestation week; (3) lacking maternal demographic information and (4) diagnosed with one or more health conditions prior to pregnancy, including type 1 or type 2 diabetes, thyroid dysfunction, hypertension, virus hepatitis and renal diseases. Sample sizes were calculated according to the association between maternal triglycerides (the potential weakest risk factors among maternal metabolic traits) with birth weight according to literature (online supplemental file S1). The eligible mother–child pairs were then selected for this study by computer-generated randomisation.

Maternal data collection

Maternal information, including age, height, prepregnancy weight, parity, date of last menstrual period, monthly income, education levels and ethnicity, was collected through semistructured questionnaire. Maternal BMI was calculated by dividing weight in kilograms by height in metres squared. Based on the recommendations of the China Obesity Task Force of the Chinese Ministry of Health, maternal prepregnancy BMI is classified into two groups: lean group (<24 kg/m^2) and overweight group (≥24 kg/m^2).[14] At 22–28 weeks gestation, women attending their second prenatal visit underwent a standard 2 hours 75 g oral glucose tolerance test (OGTT). Women with OGTT results which met or exceeded at least one threshold of the International Association of Diabetes and Pregnancy Study Groups criteria (fasting plasma glucose ≥5.1 mmol/L, 1-hour glucose ≥10 mmol/L and 2 hours glucose ≥8.5 mmol/L) were diagnosed as having gestational diabetes mellitus (GDM).[15] Maternal fasting glucose concentration was obtained from OGTT test zero-time value in hospital records. Maternal second-trimester weight was measured to the nearest 0.1 kg using an electronic scale. Maternal early GWG was calculated by subtracting prepregnancy weight from maternal second-trimester weight, with documentation of the gestational age at measurement.

Outcomes

The outcomes of this study are neonatal birth weight and CBI. We considered birth weight as it is an important indicator of prenatal developmental conditions for newborns. Birth weight has been associated with both short-term and long-term health outcomes, including stillbirth, infant mortality, obesity, type 2 diabetes and cardiovascular diseases.[16–18] Similarly, CBI, which is synthesised only by fetal pancreas,[19] is also considered as an important biomarker that could reflect insulins resistance status in neonates, since insulin plays a central role in fetal growth and development.[20]

Biochemical test

Mothers were asked to keep overnight fasting status before taking blood samples in the second trimester. Maternal real fasting status was asked again and recorded before collecting samples. Maternal blood samples during the second trimester were used for testing the level of lipid profiles. Venous umbilical cord blood plasma samples were collected for insulin tests. Sample collection, delivery, pretreatment and measurements were blinded. All blood samples were stored and delivered to the pretreatment laboratory centre. Blood samples were then separated into serum and plasma by immediate centrifugation and were stored in EDTA tubes in the Biobank at −80°C until analysis. Maternal plasma lipids, including total cholesterol (TC), HDL-C, low-density lipoprotein cholesterol (LDL-C), and triglycerides, and CBI levels were measured using commercial kits in fully automated clinical analyser (Roche Diagnostics, Mannheim, Germany). Intraday and interday coefficients of variation (CVs) were consistently less than 2% for all assays.

Neonatal anthropometry data collection

For participating children, birth information, including birth characteristics, delivery mode and perinatal outcomes, was obtained from routine medical records. Gestational age was estimated from ultrasound examination during the first or second trimester. Birth weight and

other information, including gestational age at delivery, mode of delivery, neonatal sex and pregnancy complications, were obtained from hospital records. Birth weight was measured to the nearest 50 g using an electronic scale by midwives immediately after delivery. Birth weight Z-Score and percentile adjusted for gestational age at delivery and neonatal sex were calculated using Intergrowth 21st Newborn Size Standard and Tools.[21] Large-for-gestational age (LGA) was defined as a birth weight larger than the 90th percentile for gestational age by sex, while small-for-gestational age (SGA) was defined as a birth weight smaller than the 10th percentile based on the same birth weight reference.

Patient and public involvement

No patient involved.

Statistical analysis

Classic statistical methods

Continuous variables were summarised as mean±SD or median (IQR), and categorical variables were summarised as counts with percentages. Pearson's correlation was used to assess the impact of the long-term −80°C storage on insulin concentrations in EDTA tube. Multiple imputation was used to handle missing data. Adjustments were then made to account for any degradation by correcting the initial value using linear regression methods (online supplemental file S2). Similarly, maternal lipid levels were adjusted for gestational age using regression model to account for timing of blood sampling (online supplemental file S3).[22]

Initially, multivariable linear regression models were used to estimate the association of maternal lipid levels with neonatal birth weight and CBI level. Covariates in the regression model included maternal age, ethnic group, parity, gestational age, maternal fasting status, neonatal sex and early pregnancy cigarette exposures. Delivery mode and sample storage duration were added in the model for CBI level. Thereafter, same model was used to investigate the association of maternal lipid Z-Scores with birth weight Z-Score and CBI Z-Score. To investigate the independent associations of maternal lipid levels with neonatal birth weight and CBI, model was further adjusted for maternal prepregnancy BMI Z-Score, maternal fasting glucose Z-score, GWG Z-score and gestational age of maternal weight measurement during pregnancy. Sensitivity analyses were conducted to compare the estimate differences between GDM and non-GDM participants, primiparous women and non-primiparous women, lean and overweight group, as well as before and after multiple imputations. All statistical tests were two tailed and a $p<0.05$ was considered statistically significant. Statistical analyses were performed in Stata V.14.0.

Additive Bayesian networks analysis

To further assess the interdependency between maternal metabolic risk factors and their association with birth weight and CBI, additive Bayesian Network (ABN) model—an unsupervised machine learning method—was conducted. Bayesian network analysis is a form of structure discovery statistical modelling that derives, from empirical data, a graphical network describing the dependency structure between variables, shown as directed acyclic graphs (DAGs).[23] ABNs comprise DAGs where each node in the graph comprises a generalised linear model or a generalised linear mixed model. ABN model is suitable for analysing highly complex epidemiological data comprising many interdependent variables.[24]

Ten variables were chosen for ABN based on prior knowledge gained from literature and findings of the classical statistical analyses. These 10 variables were maternal age, maternal prepregnancy BMI, maternal fasting glycaemia in OGTT, early GWG, maternal fasting plasma HDL-C and triglycerides in the second trimester, birthweight Z-Score, CBI, gestational age at delivery and neonatal sex. GWG was adjusted for gestational age at weight measurement in mid-pregnancy. CBI was adjusted for sample storage duration. All continuous variables were standardised to Z-Scores to eliminate the influence of different measurement units. Mother–child pairs with missing data were excluded (n=93/1522, 6%).

First, an optimal DAG with the best goodness of fit (highest log marginal likelihood) was identified. Next, parametric bootstrapping (12 800 samples) was performed to address the potential overfitting. Full technical details are provided in online supplemental file S4, supplemental figure S1–S5. ABN analysis was conducted in R V.3.4.4 (The R Foundation for Statistical Computing) using 'abn' package.[24]

RESULTS

Figure 1 shows the flow diagram of study. A total of 5497 pregnant women attending BIGCS between January 2015 and June 2016 are eligible for this study. Women whose blood samples were unavailable at 14–27 gestation week (n=902), who had no records of maternal fasting glucose at 20–28 gestation week (n=343), lacked maternal demographic information (n=39) and diagnosed with health condition prior to pregnancy (n=118) were excluded. The remaining 4039 women were subjected to randomisation, resulting in 1522 women were selected for maternal lipid profiles and CBI testing. The baseline characteristics of

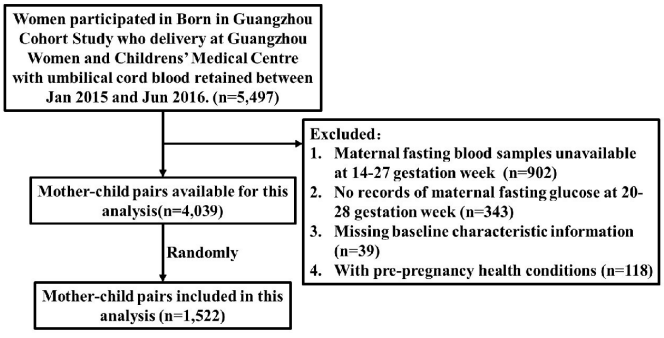


Figure 1 Flow chart.

Table 1 Baseline characteristics table

Characteristics	Included participants (n=1522)
Maternal baseline information	
Maternal age at enrolment (years)	29.50±3.30
Ethnic Han	1486 (97.70)
Primiparous	1223 (80.35)
Spontaneous delivery	1239 (81.41)
Early pregnancy cigarette exposure	436 (28.68)
Maternal metabolic profile	
GDM	181 (11.89)
Fasting glucose (mmol/L)	4.25±0.42
Gestational age of OGTT test (weeks)	25.60 (1.38)
Prepregnancy BMI (kg/m^2)	20.47±3.85
Early gestational weight gain (kg)	4.21±8.42
Total cholesterol (mmol/L)	5.47±0.90
HDL-C (mmol/L)	2.07±0.43
LDL-C (mmol/L)	3.06±0.77
Triglycerides (mmol/L) *	1.71 (1.39–2.15)
Gestational age of blood sampling (weeks)*	19 (17–24)
Neonatal information	
Gestational age (days)*	275 (270–281)
Preterm delivery	66 (4.34)
Male	820 (53.88)
Birth weight (g)	3,203±411
LGA	96 (6.31)
SGA	106 (6.96)
Cord blood insulin (μU/mL) *	7.43 (4.34–12.61)

Data are mean±SD or n (%).
*Median (IQR).
BMI, body mass index; GDM, gestational diabetes mellitus; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; LGA, large-for-gestational age; OGTT, the oral glucose tolerance test; SGA, small-for-gestational age.

participants are shown in table 1. The majority (91.20%) of maternal blood samples were collected after overnight fasting. Maternal mid-pregnancy weight, fasting glucose and lipids profile were measured at a mean of 20.0 (SD=4.0), 24.6 (SD=1.4) and 20.5 (SD=3.5) gestation weeks, respectively. Cord blood samples were stored for a median of 488 (IQR 394–707) days before analysis.

Table 2 presents the associations between the association of maternal mid-pregnancy lipid levels with neonatal birth weight and CBI level. The elevated maternal triglycerides level was associated with increased birth weight (adjusted β=67.46, 95% CI 41.85 to 93.06 g per mmol/L) and CBI level (adjusted β=0.89, 95% CI 0.06 to 1.72 μU/mL per mmol/L). Conversely, the increased maternal HDL-C level was negatively associated with decreased birth weight (adjusted β=−45.29, 95% CI −85.49 to −5.09 g per mmol/L), but not associated with CBI level (adjusted β=−0.82, 95% CI −2.12 to 0.48 μU/mL per mmol/L). No statistically significant association was observed between maternal TC and LDL-C levels and birth weight/CBI.

Following further adjustment of prepregnancy BMI, GWG and maternal glucose level, only the association between maternal triglyceride Z-Score and birth weight Z-Score remains statistically significant (standardised β=0.07, 95% CI 0.03 to 0.11) (online supplemental table S1 and S2). No statistically significant result was observed in the sensitivity analysis (online supplemental table S3). It is worth noticing that the association estimate of maternal glucose Z-Score with birth weight Z-Score differed between non-GDM and GDM participants (standardised β: 0.06 vs 0.16, p=0.06), while the association estimate of maternal prepregnancy BMI Z-Score with birth weight Z-Score differed between primiparous and multiparous women (standardised β: 0.18 vs 0.28, p=0.06). Estimates for the association of maternal HLD-C Z-Score (standardised β: −0.03 vs −0.13, p=0.17), prepregnancy BMI Z-Score (standardised β: 0.03 vs 0.14, p=0.38) and glucose Z-Score (standardised β: 0.10 vs 0.19, p=0.13) with CBI Z-Score differed between non-GDM and GDM women. Between lean and overweight women, a difference in the association between maternal HDL-C Z-Score and CBI Z-Score was observed (standardised β: −0.03 vs −0.12, p=0.24). All estimates remain the same before and after multiple imputations.

Figure 2 and online supplemental table S4 show the optimal summary DAGs inferred by ABN analysis. Neither triglycerides nor HDL-C were linked to birth weight or CBI in the ABN results after accounting for the interdependency of maternal metabolic factors (including lipids). Maternal prepregnancy BMI was associated with all other maternal metabolic parameters in pregnancy (glycaemia: β=0.14, 95% CI 0.09 to 0.19; early GWG: β=−0.12, 95% CI −0.17 to −0.06; triglycerides: β=0.23, 95% CI 0.18 to 0.28; HDL-C: β=−0.12, 95% CI −0.17 to −0.07) and birth weight (β=0.27, 95% CI 0.22 to 0.32). An indirect association between prepregnancy BMI and CBI was also observed. Early GWG was associated with birth weight Z-Score (β=0.17, 95% CI 0.12 to 0.22). Maternal glycaemia was associated with CBI (β=0.12, 95% CI 0.07 to 0.17). Birth weight was also associated with CBI (β=0.24, 95% CI 0.19 to 0.29).

DISCUSSION

To the best of our knowledge, this is the first large prospective birth cohort study that considered the interdependency of maternal lipids profiles with other metabolic risk factors (maternal prepregnancy weight, GWG and gestational glycaemia level) when assessing the association between maternal circulating lipids levels and birth outcomes. We found that maternal mid-pregnancy HDL-C and triglycerides levels may not independently associated with birth weight and CBI level using Bayesian network analysis, although statistically significant

Table 2 Association of maternal lipid levels with birth weight and cord blood insulin level

	TC (mmol/L)	HDL-C (mmol/L)	LDL-C (mmol/L)	TG (mmol/L)
Birth weight (g)				
β (95% CI)*	−0.42 (−20.32 to 18.76)	**−45.29 (−85.49 to −5.09)¶**	−0.24 (−22.82 to 22.34)	**67.46 (41.85 to 93.06)**
Standardised β (95% CI)†	0.00 (−0.04 to 0.04)	−0.01 (−0.05 to 0.03)	0.00 (−0.04 to 0.04)	**0.07 (0.03 to 0.11)**
Cord blood insulin (μU/mL)				
β (95% CI)‡	−0.15 (−0.77 to 0.48)	−0.82 (−2.12 to 0.48)	−0.15 (−0.88 to 0.58)	**0.89 (0.06 to 1.72)**
Standardised β (95% CI) §	−0.02 (−0.07 to 0.03)	−0.01 (−0.07 to 0.04)	−0.03 (−0.08 to 0.02)	0.04 (−0.01 to 0.09)

*Adjusted for maternal age, ethnic group, parity, gestational age, neonatal sex, early pregnancy cigarette exposures and maternal fasting status.
†Standardised β for maternal lipid Z-Scores and birth weight Z-Score. Adjusted for maternal age, ethnic group, parity, gestational age, neonatal sex, early pregnancy cigarette exposures, maternal prepregnancy BMI Z-Score, maternal glucose Z-Score, gestational weight gain Z-Score, gestational age of maternal weight measurement during pregnancy and maternal fasting status.
‡Adjusted for maternal age, ethnic group, parity, gestational age, neonatal sex, early pregnancy cigarette exposures, delivery mode, sample storage duration and maternal fasting status.
§Standardised β for maternal lipid Z-Scores and cord blood insulin Z-Score. Adjusted for maternal age, ethnic group, parity, gestational age, neonatal sex, early pregnancy cigarette exposures, delivery mode, sample storage duration, maternal prepregnancy BMI Z-Score, maternal glucose Z-Score, gestational weight gain Z-Score, gestational age of maternal weight measurement during pregnancy and maternal fasting status.
¶Boldface values refer to results with statistical significance.
BMI, body mass index; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, total cholesterol; TG, triglycerides.

associations—consistent with previous studies—were observed using conventional multivariable linear regression models.

Our results demonstrated that lipid pathways may not be meaningfully involved in the metabolic network pathway between mothers and neonates, and instead be a proxy measure for maternal metabolic health. Similarly, a previous study using Mendelian randomisation, another established causal inference method, analysing data from 30 487 women in 18 studies concluded that genetically higher maternal fasting HDL-C/triglycerides was not potentially causally associated with higher birth weight.[25] Thus, both detailed pathways analyses in this paper and genetic finding suggest that lipid pathways are not causally related to birth weight.

Although the underlying mechanism of the absence of association between maternal lipid profiles and birth weight and CBI remains unclear, it is biologically

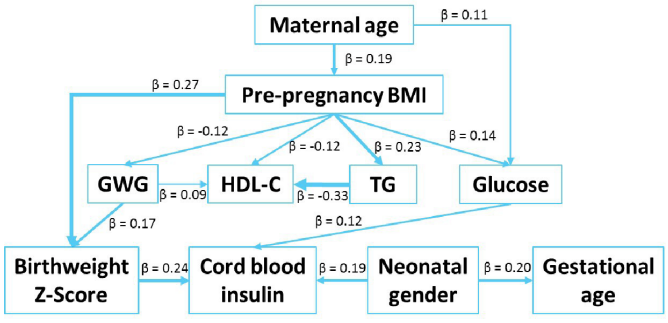


Figure 2 Additive Bayesian network graph. β, standardised regression coefficient; BMI, body mass index; GWG, gestational weight gain; HDL-C, high-destiny-lipoprotein cholesterol; TG; triglycerides.

plausible. Pregnant women with a higher prepregnancy BMI, increased GWG or/and glucose intolerance often have associated dyslipidaemia.[1 26–29] The transportation efficiency of free fatty acids and glycerol broken down from triglycerides across placenta is significantly influenced by maternal insulin level, although it cannot cross the placenta.[19 30] Elevated insulin concentration in fetus triggered by the increased maternal hyperglycaemia/hyperinsulinaemia could promote the circulating free fatty acids, triacylglycerol and glucose uptake in adipose and muscle tissue, block glycogenolysis and gluconeogenesis in the liver, and stimulate glycogen synthesis, resulting in enlarged adipose tissue in fetus.[31–35] Conversely, the enlarged adipocytes could further induce insulin resistance, leading to an elevated insulin level in fetus.[31 36] Therefore, although the glycerol and free fatty acids delivered from the maternal side are the essential nutrient substrate for neonatal de novo lipogenesis, it is plausible that the maternal circulating lipid levels, the placental transport of lipids and the uptake of lipids in the neonatal side are mainly determined by the maternal insulin resistance status driven by the body weight, insulin resistance status and glycaemic level in mothers.

Since 1950s, Pedersen hypothesis suggested that the increase in glucose transport from GDM mothers to fetus induces fetal hyperinsulinaemia, and results in subsequently adipose accumulation in neonates.[37–39] In turn, the enlarged adipocyte will gradually become resistant to insulin to avoid further expansion, therefore, contributing to the increased insulin secretion in neonates.[40] Hence, there may be a bidirectional relationship between birth weight and CBI. Due to the nature of DAG and lacking dynamic data, the ABN

analysis might have a limited ability on exploring feedback loops. That means that the current DAG generated by the ABN analysis missed an arc from CBI level to birth weight, and that may be the reason why no arc from maternal glucose level to neonatal birth weight was observed.

Our ABN analysis suggested that the maternal prepregnancy BMI is the overweight/obesity is the most influential upstream metabolic risk factor for both maternal and neonatal metabolic health. Previous evidence shows maternal overweight/obesity has a persistent impact on the quality and content of oocytes as well as the epigenetic and translation profile of the developing embryo, thereby leading to altered phenotypes in offspring.[41–44] Compared with lean women, women with a higher prepregnancy BMI tend to have increased risks of excessive GWG as well as GDM and often present with more severe dyslipidaemia during pregnancy.[45–47] The hyperinsulinaemia and systemic low-grade inflammation driven by maternal obesity, gestational hyperglycaemia and gestational dyslipidaemia could potentially induce functional changes in the placenta, leading to elevated nutrients uptake and transport.[48 49] Stress induced by the overwhelming nutrient exposure in utero can systematically shape fetal development, leading to permanent changes of metabolic function in offspring.[50]

Implications

If replicated, the findings in this study may have implications for antenatal/prenatal counselling and managing to avoid adverse weight-related birth outcomes. Most current clinical guidelines on preconception and antenatal care only focus on weight management during pregnancy. Recommendations on prepregnancy weight management are limited and ambiguous.[51–53] Our results provide further important evidence on the clinical importance of maternal prepregnancy high BMI for both maternal and neonatal health outcomes. Dietary and/or physical activity interventions initiated before or in early pregnancy to reduce prepregnancy BMI, GWG and insulin resistance, rather than maternal lipid levels, would likely be most effective, although the effects need further investigation in future randomised trials.

Strengths and limitations

The major strengths of this study are the prospective design based on relatively large sample size, standardisation of strength of association for the comparison among maternal metabolic risk factors, and the use of powerful analytical tools for interpretation of multidimensional data. The study also has some limitation. First, given the practical constraints, maternal fasting glucose and triglycerides levels were measured only once during pregnancy. Therefore, we could not investigate the dynamic long-term influences of maternal metabolic risk factors in detail, although such levels generally track well over gestation. Second, the average prepregnancy BMI of included women and incidence of LGA/SGA babies in this study were significantly lower than for people living in the northern part of China. The relative healthiness of our cohort suggests that our results might underestimate the true impact of maternal metabolic disorders on neonatal health outcomes if extrapolated to this wider population. Third, the prepregnancy weight was self-reported, which might potentially underestimate the true value. However, evidence suggests that utilisation of self-reported or measured prepregnancy weight for prepregnancy BMI classification results in identical categorisation for most women.[52] Fourth, as we mentioned earlier, maternal insulin resistance might be a significant explanatory factor in the whole pathway but was not estimated in the current analysis due to practical constraints. Although we include maternal fasting glucose level in the second trimester as a proxy of maternal insulin resistance, it may not sufficiently reflect the reality. Further analyses including maternal insulin resistance status are warranted.

CONCLUSION

In this cohort study, we found that mid-pregnancy maternal HDL-C and triglycerides concentrations was not independently associated with birth weight or CBI level after taking the interdependency of maternal lipids profiles with other metabolic risk factors into account. Pregnancy BMI, fasting glucose and GWG are three metabolic risk factors independently associated with increased birth weight and/or insulin secretion in neonates, with the prepregnancy BMI being the most influential upstream risk factor. Interventions initiated before or in early pregnancy to reduce prepregnancy BMI, GWG and insulin resistance would likely be most effective and need further investigation in future randomised trials.

Author affiliations
[1]Division of Birth Cohort Study, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China
[2]Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[3]Guangdong Provincial Key Laboratory of Research in Structure Birth Defect Disease, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China
[4]Guangdong Provincial Clinical Research Center for Child Health, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China
[5]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK
[6]University of Geneva, Geneva, Switzerland
[7]Department of Women's Health, Guangdong Provincial Key Clinical Specialty of Woman and Child Health, Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

Twitter Naveed Sattar @MetaMedTeam

Contributors KN, XQ, KKC and JW conceived the research question. JW defined the question, designed the study and conducted data analysis. KN and MJP advised on study design and contributed to data analysis. KKC, H-MX, GNT and XQ

provided input for study design. YK, JH, SS, J-HL and JW contributed to the sample collection, storage and biochemical tests, and data preprocessing. MP checked the ABN analysis externally. All authors contributed to the interpretation of the results. JW led the writing of the manuscript with critical input from all other authors (YK, SS, MJP, J-HL, NS, JH, MP, H-MX, GNT, XQ, KKC and KN). JW, XQ and KN had full access to all data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. JW is the guarantor.

Funding This work was supported by National Natural Science Foundation of China [grant numbers 82173525 and 82273642]. JW was supported by the LiSiguang scholarship provided by the University of Birmingham and the China Scholarship Council jointly (award number: 201508060007).

Disclaimer The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Consent obtained directly from patient(s).

Ethics approval Ethical permission for the study was granted by the Guangzhou Women and Children's Medical Center Ethics Committee (Reference number: 2013053107). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as online supplemental information. No additional data available.

ORCID iDs
Jingya Wang http://orcid.org/0000-0003-1498-2693
Naveed Sattar http://orcid.org/0000-0002-1604-2593
Krishnarajah Nirantharakumar http://orcid.org/0000-0002-6816-1279

REFERENCES

1 Heerwagen MJR, Miller MR, Barbour LA, *et al*. Maternal obesity and fetal metabolic programming: a fertile epigenetic soil. *Am J Physiol Regul Integr Comp Physiol* 2010;299:R711–22.
2 Scholtens DM, Bain JR, Reisetter AC, *et al*. Metabolic networks and metabolites underlie associations between maternal glucose during pregnancy and newborn size at birth. *Diabetes* 2016;65:2039–50.
3 HAPO Study Cooperative Research Group, Metzger BE, Lowe LP, *et al*. Hyperglycemia and adverse pregnancy outcomes. *N Engl J Med* 2008;358:1991–2002.
4 Villamor E, Cnattingius S. Interpregnancy weight change and risk of adverse pregnancy outcomes: a population-based study. *Lancet* 2006;368:1164–70.
5 LifeCycle Project-Maternal Obesity and Childhood Outcomes Study Group, Voerman E, Santos S, *et al*. Association of gestational weight gain with adverse maternal and infant outcomes. *JAMA* 2019;321:1702–15.
6 McIntyre HD, Dekker Nitert M, McIntyre HDBarrett, *et al*. Normalizing metabolism in diabetic pregnancy: is it time to target lipids? *Diabetes Care* 2014;37:1484–93.
7 Barbour LA. Metabolic Culprits in Obese Pregnancies and Gestational Diabetes Mellitus: Big Babies, Big Twists, Big Picture : The 2018 Norbert Freinkel Award Lecture. *Diabetes Care* 2019;42:718–26.
8 Busso D, Rigotti A. Blood lipids during pregnancy: a progressively appreciated subject in basic and clinical research. *Atherosclerosis* 2018;276:163–5.
9 Wang J, Moore D, Subramanian A, *et al*. Gestational dyslipidaemia and adverse birthweight outcomes: a systematic review and meta-analysis. *Obes Rev* 2018;19:1256–68.
10 Kulkarni SR, Kumaran K, Rao SR, *et al*. Maternal lipids are as important as glucose for fetal growth: findings from the Pune maternal nutrition study. *Diabetes Care* 2013;36:2706–13.
11 Schaefer-Graf UM, Graf K, Kulbacka I, *et al*. Maternal lipids as strong determinants of fetal environment and growth in pregnancies with gestational diabetes mellitus. *Diabetes Care* 2008;31:1858–63.
12 Sandler V, Reisetter AC, Bain JR, *et al*. Associations of maternal BMI and insulin resistance with the maternal metabolome and newborn outcomes. *Diabetologia* 2017;60:518–30.
13 Qiu X, Lu J-H, He J-R, *et al*. The born in Guangzhou cohort study (BIGCS). *Eur J Epidemiol* 2017;32:337–46.
14 Chen C, Lu FC, Department of Disease Control Ministry of Health, PR China. The guidelines for prevention and control of overweight and obesity in Chinese adults. *Biomed Environ Sci* 2004;17 Suppl:1–36.
15 Metzger BE, Gabbe SG, Persson B, *et al*. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. *Diabetes Care* 2010;33:e98–682.
16 Harder T, Rodekamp E, Schellong K, *et al*. Birth weight and subsequent risk of type 2 diabetes: a meta-analysis. *Am J Epidemiol* 2007;165:849–57.
17 McIntire DD, Bloom SL, Casey BM, *et al*. Birth weight in relation to morbidity and mortality among newborn infants. *N Engl J Med* 1999;340:1234–8.
18 Yu ZB, Han SP, Zhu GZ, *et al*. Birth weight and subsequent risk of obesity: a systematic review and meta-analysis. *Obes Rev* 2011;12:525–42.
19 Ruiz-Palacios M, Ruiz-Alcaraz AJ, Sanchez-Campillo M, *et al*. Role of insulin in placental transport of nutrients in gestational diabetes mellitus. *Ann Nutr Metab* 2017;70:16–25.
20 Fowden AL. The role of insulin in fetal growth. *Early Hum Dev* 1992;29:177–81.
21 Villar J, Papageorghiou AT, Pang R, *et al*. The likeness of fetal growth and newborn size across non-isolated populations in the INTERGROWTH-21st project: the fetal growth longitudinal study and newborn cross-sectional study. *Lancet Diabetes Endocrinol* 2014;2:781–92.
22 Edison RJ, Berg K, Remaley A, *et al*. Adverse birth outcome among mothers with low serum cholesterol. *Pediatrics* 2007;120:723–33.
23 Bielza C, Larrañaga P. Bayesian networks in neuroscience: a survey. *Front Comput Neurosci* 2014;8:131.
24 Pittavino M, Lewis F, Furrer R. Abn: an R package for modelling multivariate data using additive Bayesian networks. *The Comprehensive R Archive Network* 2016:1–37.
25 Tyrrell J, Richmond RC, Palmer TM, *et al*. Genetic evidence for causal relationships between maternal obesity-related traits and birth weight. *JAMA* 2016;315:1129–40.
26 Klop B, Elte JWF, Cabezas MC. Dyslipidemia in obesity: mechanisms and potential targets. *Nutrients* 2013;5:1218–40.
27 Carr MC, Brunzell JD. Abdominal obesity and dyslipidemia in the metabolic syndrome: importance of type 2 diabetes and familial combined hyperlipidemia in coronary artery disease risk. *J Clin Endocrinol Metab* 2004;89:2601–7.
28 Ryckman KK, Spracklen CN, Smith CJ, *et al*. Maternal lipid levels during pregnancy and gestational diabetes: a systematic review and meta-analysis. *BJOG* 2015;122:643–51.
29 Petersen MC, Shulman GI. Mechanisms of insulin action and insulin resistance. *Physiol Rev* 2018;98:2133–223.
30 Brett KE, Ferraro ZM, Yockell-Lelievre J, *et al*. Maternal-Fetal nutrient transport in pregnancy pathologies: the role of the placenta. *Int J Mol Sci* 2014;15:16153–85.
31 Varlamov O, Somwar R, Cornea A, *et al*. Single-Cell analysis of insulin-regulated fatty acid uptake in adipocytes. *Am J Physiol Endocrinol Metab* 2010;299:E486–96.
32 Saltiel AR, Kahn CR. Insulin signalling and the regulation of glucose and lipid metabolism. *Nature* 2001;414:799–806.
33 Czech MP, Tencerova M, Pedersen DJ, *et al*. Insulin signalling mechanisms for triacylglycerol storage. *Diabetologia* 2013;56:949–64.
34 Lawlor DA, West J, Fairley L, *et al*. Pregnancy glycaemia and cord-blood levels of insulin and leptin in Pakistani and white British

mother–offspring pairs: findings from a prospective pregnancy cohort. *Diabetologia* 2014;57:2492–500.
35 Huhtala MS, Rönnemaa T, Pellonperä O, *et al*. Cord serum metabolome and birth weight in patients with gestational diabetes treated with metformin, insulin, or diet alone. *BMJ Open Diabetes Res Care* 2021;9:e002022.
36 Wang J, Shen S, Price MJ, *et al*. Glucose, insulin, and lipids in cord blood of neonates and their association with birthweight: differential metabolic risk of large for gestational age and small for gestational age babies. *J Pediatr* 2020;220:64–72.
37 Petersen MC, Shulman GI. Mechanisms of insulin action and insulin resistance. *Physiol Rev* 2018;98:2133–223.
38 Lawlor DA, West J, Fairley L, *et al*. Pregnancy glycaemia and cord-blood levels of insulin and leptin in Pakistani and white British mother-offspring pairs: findings from a prospective pregnancy cohort. *Diabetologia* 2014;57:2492–500.
39 Pedersen J. Diabetes mellitus and pregnancy: present status of the hyperglycaemia--hyperinsulinism theory and the weight of the newborn baby. *Postgrad Med J* 1971;:66–7.
40 Rasmussen KM, Yaktine AL. *Weight gain during pregnancy: reexamining the guidelines*, 2009.
41 Luzzo KM, Wang Q, Purcell SH, *et al*. High fat diet induced developmental defects in the mouse: oocyte meiotic aneuploidy and fetal growth retardation/brain defects. *PLoS One* 2012;7:e49217.
42 Igosheva N, Abramov AY, Poston L, *et al*. Maternal diet-induced obesity alters mitochondrial activity and redox status in mouse oocytes and zygotes. *PLoS One* 2010;5:e10074.
43 Minge CE, Bennett BD, Norman RJ, *et al*. Peroxisome proliferator-activated receptor-gamma agonist rosiglitazone reverses the adverse effects of diet-induced obesity on oocyte quality. *Endocrinology* 2008;149:2646–56.
44 Lane M, Zander-Fox DL, Robker RL, *et al*. Peri-conception parental obesity, reproductive health, and transgenerational impacts. *Trends Endocrinol Metab* 2015;26:84–90.
45 Heerwagen MJR, Miller MR, Barbour LA, *et al*. Maternal obesity and fetal metabolic programming: a fertile epigenetic soil. *Am J Physiol Regul Integr Comp Physiol* 2010;299:R711–22.
46 Guelinckx I, Devlieger R, Beckers K, *et al*. Maternal obesity: pregnancy complications, gestational weight gain and nutrition. *Obes Rev* 2008;9:140–50.
47 Torloni MR, Betrán AP, Horta BL, *et al*. Prepregnancy BMI and the risk of gestational diabetes: a systematic review of the literature with meta-analysis. *Obes Rev* 2009;10:194–203.
48 Magnusson-Olsson AL, Hamark B, Ericsson A, *et al*. Gestational and hormonal regulation of human placental lipoprotein lipase. *J Lipid Res* 2006;47:2551–61.
49 Challier JC, Basu S, Bintein T, *et al*. Obesity in pregnancy stimulates macrophage accumulation and inflammation in the placenta. *Placenta* 2008;29:274–81.
50 Ellis BJ, Del Giudice M. *Developmental adaptation to stress: an evolutionary perspective*, 2019.
51 Weight management before, during and after pregnancy. Available: https://www.nice.org.uk/Guidance/PH27
52 Bannon AL, Waring ME, Leung K, *et al*. Comparison of self-reported and measured Pre-pregnancy weight: implications for gestational weight gain counseling. *Matern Child Health J* 2017;21:1469–78.
53 Care of women with obesity in pregnancy (Green-top guideline No. 72). Available: https://www.rcog.org.uk/en/guidelines-research-services/guidelines/gtg72/

Supplementary File

S1 Sample size calculation

We powered the study for the potentially least associated maternal metabolic risk factor (triglycerides) for birthweight. Knopp et al. reported a correlation between maternal triglycerides and birthweight of r=0.09 ($p<0.05$) in non-GDM women.(1) We conservatively assumed an effect size of 0.08. STATA 14.0 was used to calculate the sample size. After using 'Fisher's z tests comparing one correlation to a reference value' tool, a sample of 1225 will give 80% power to detect a correlation of 0.08 at 5% significance level (two sided). We conservatively assumed 20% attrition rate due to missing data and loss to follow up, thus giving a sample size of 1,531.

Wang J, *et al. BMJ Open* 2022; 12:e064122. doi: 10.1136/bmjopen-2022-064122

S2 Testing for degradation for cord blood insulin

No prior literature reported the impact of long-term -80 °C storage on plasma insulin. We therefore fitted a regression model to detect potential degradation for cord blood insulin. The median storage duration of cord blood sample is 488 (IQR 394 to 707) days. Cord blood insulin was found to be slightly degraded over time (r=-0.07, p=0.01). In the multivariate regression model, we included sample storage time as a covariate. In the additive Bayesian Network analysis, adjustments were made to account for any degradation by correcting the initial value using linear regression methods (adjusted cord blood insulin = initial value of cord blood insulin + (mean value of sample storage time – sample storage time) * β, β=-0.0044446).

S3 Adjusting gestational age at sampling for maternal lipid profile

The average sampling time for maternal overnight fasting blood sample at the second trimester was 20.46 gestation weeks. The table below shows the estimates of associations between maternal plasma lipid levels and gestational age when blood sampling was carried out.

lipids	Regression coefficient (β)
Total cholesterol	0.098
HDL-C	0.015
LDL-C	0.075
Triglycerides	0.059

Linear regression model

Adjustment Equation: Adjusted lipids = initial value + (20.46 – sampling time) * β

Wang J, *et al. BMJ Open* 2022; 12:e064122. doi: 10.1136/bmjopen-2022-064122

S4 Additive Bayesian Network methodologies

Introduction to Additive Bayesian Network analysis

A Bayesian network is a probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed acyclic graphs (DAGs).(2) It is a well-established unsupervised machine learning methodology that is typically referred to as structure discovery model for dealing with multidimensional data.(3) Unlike other widely used multivariate approaches, such as principal component analysis, propensity score matching analysis and multivariable regression model, graphical modelling does not involve any dimension reduction. Most graphical models, including path analysis and structural equation modelling, rely on a pre-specified structure, whereas Bayesian network is entirely data-driven.

Unlike the contingency table parameterization in standard Bayesian network models, Additive Bayesian networks (ABN) allow us to obtain interpretable DAGs where each node in graph comprises a generalized linear model (GLM) or a generalized linear mixed model (GLMM, if binary variable involved).(4, 5) There are two mutually dependent parts in ABN model: a network structure (i.e. the DAG) and a set of parameters. Each node (corresponding to the variables in the dataset) in the DAG is the equivalent of a potential dependent variable in a Bayesian GLM or GLMM regression model. While other DAG nodes where relevant as identified by the unsupervised learning act as covariates, having a role of corresponding parameters. Therefore, an ABN model is ideally suited to analysing highly complex epidemiological data comprising many inter-dependent variables.

The technical process of ABN

After an initial data preparation phase we used a three-step procedure to determine an optimal DAGs for our data.

Step 0 Data pre-processing

Ten variables were chosen for ABN based on our knowledge gained from prior literature and findings of the classical statistical analyses. These included maternal age, maternal pre-pregnancy BMI, maternal fasting glucose concentration in OGTT, early gestational weight gain (GWG, adjusted for gestational age at weight measurement), maternal fasting plasma high-density lipoprotein cholesterol (HDL-C, adjusted for gestational age at blood sampling) in 2nd trimester, maternal fasting plasma triglycerides in 2nd trimester (adjusted for gestational age at blood sampling), birthweight Z-Score (adjusted for gestational age at delivery and neonatal gender), cord blood insulin (CBI, adjusted for sample storage duration) concentration, gestational age at delivery, and neonatal gender. All continuous variables were standardized to Z-Scores to eliminate the influence of different measurement units (maternal triglycerides and cord blood insulin were log-transformed before standardization). Participants with data missing for at least one of these ten variables (6% of participants) were excluded from the analysis. The number of mother-child pairs that was finally included in ABN analysis is 1,429.

Step 1 Identification of the optimal model

The identification of the single optimal model is referred to as structure discovery. The purpose of this step is to combine all individual GLMs into a single, probabilistically cohesive model describing all the inter-dependent relationships via a DAG. We blocked all directions of arcs between variables that are biologically impossible to occur. This was done using the adjacency matrix in figure S1.

```
ban <- matrix( c(
  # 01 02 03 04 05 06 07 08 09 10
     0, 1, 1, 1, 1, 1, 1, 1, 1, 1, # 01 mage
     0, 0, 1, 1, 1, 1, 1, 1, 1, 1, # 02 prebmi
     0, 0, 0, 0, 0, 0, 1, 1, 1, 1, # 03 gwg
     0, 0, 0, 0, 0, 0, 1, 1, 1, 1, # 04 glu
     0, 0, 0, 0, 0, 0, 1, 1, 1, 1, # 05 hdl
     0, 0, 0, 0, 0, 0, 1, 1, 1, 1, # 06 tg
     0, 0, 0, 0, 0, 0, 0, 0, 0, 0, # 07 bwz
     0, 0, 0, 0, 0, 0, 0, 0, 0, 0, # 08 ins
     1, 1, 1, 1, 1, 1, 1, 1, 0, 1, # 09 sex
     0, 0, 0, 0, 0, 0, 0, 0, 0, 0 # 10 gaw
), byrow=TRUE, ncol=10)
```

Variable labels explanation: 01 mage, maternal age; 02 prebmi, maternal pre-pregnancy BMI; 03 gwg, gestational weight gain; 04 glu, maternal fasting glucose level; 05 hdl, maternal plasma high-density lipoprotein cholesterol level; 06 tg, maternal plasma triglyceride level; 07 bwz, birthweight Z-Score; 08 ins, cord blood insulin; 09 sex, neonatal gender; 10 gaw, gestational age at delivery. Same labels also apply to the numbers across the top of the matrix.
DAG definition. Rows are children nodes, columns are parent nodes. 1 represents block from parent node (column) towards child node (row), 0 represents unblock.

Figure S1 ABN block matrix definition

To find the DAG with the best goodness of fit (network score - log marginal likelihood), exact searches were conducted across the parent limits (the limit number of arcs from parent nodes to child node), starting from a minimum of 1 and reaching a maximum of 9. As shown in Figure S2, we found that the goodness of fit (maximum marginal likelihood=-19153.30) does not improve when the number of parent limit is greater than 4.

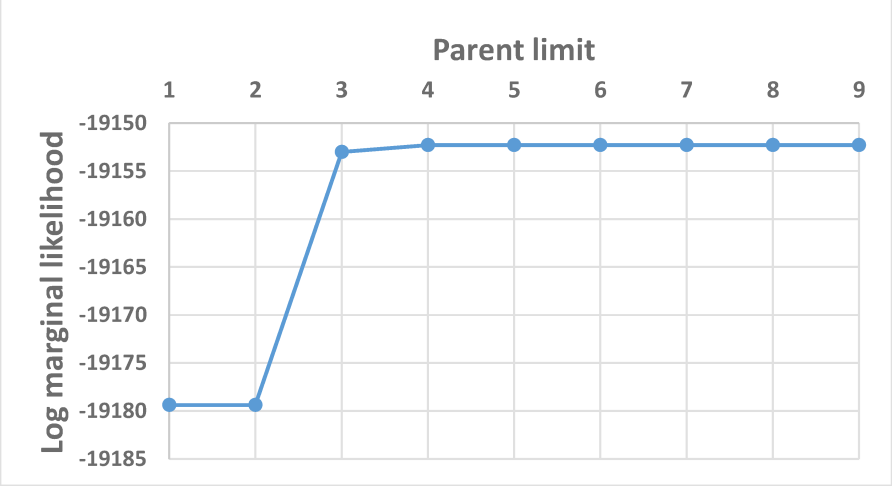
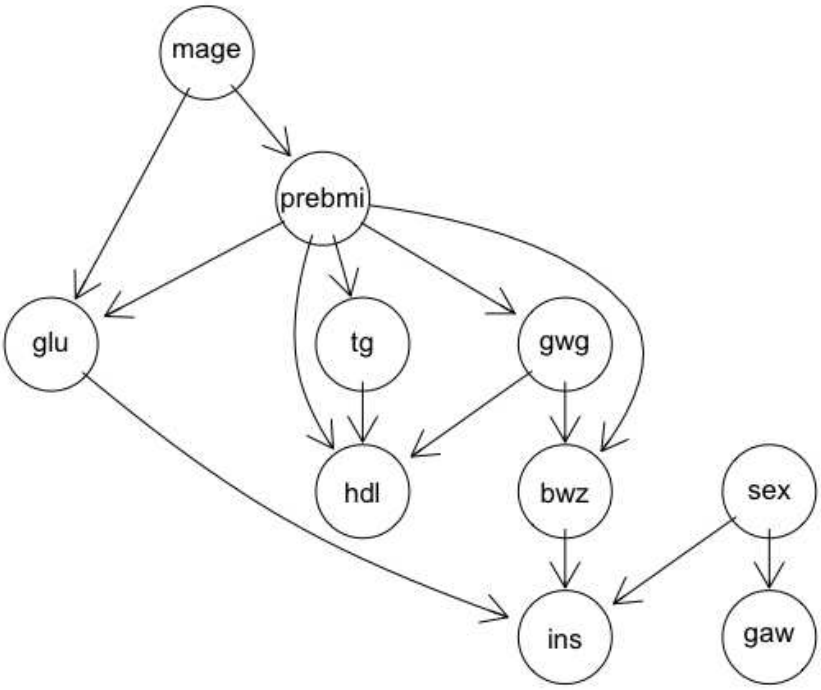


Figure S2 Comparison of goodness of fits for different parent limits



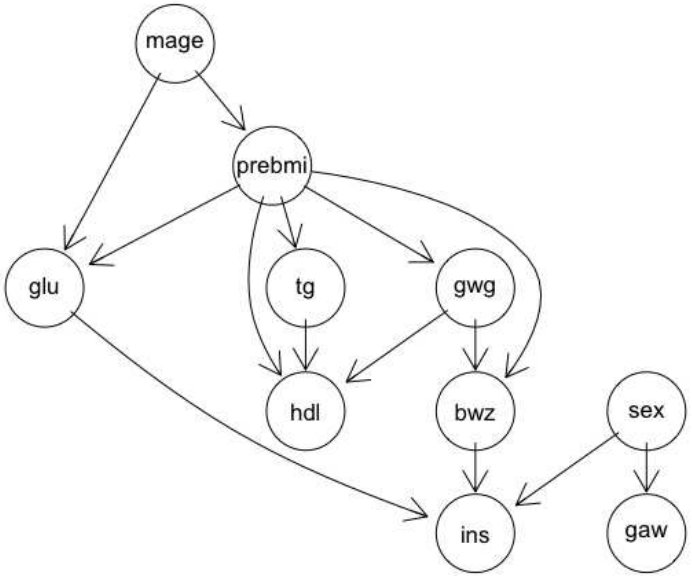
Variables explanation: mage, maternal age; prebmi, maternal pre-pregnancy BMI; gwg, gestational weight gain; glu, maternal fasting glucose level; hdl, maternal plasma high-density lipoprotein cholesterol level; tg, maternal plasma triglyceride level; bwz, birthweight Z-Score; ins, cord blood insulin; sex, neonatal gender; gaw, gestational age at delivery.

Figure S3 The identified optimal DAG from the initial search

Step 2 Adjustment for overfitting: parametric bootstrapping

We have identified the optimal DAG, but there is a risk of overfitting because of the combinatoric nature of Bayesian hypotheses. To address this, 12,800 independent parametric bootstrapping analyses were performed. This involves simulating data sets of the same size as the original dataset, and see how often the different structural features are recovered. Arcs present in less than 50% frequencies of the globally optimal DAGs estimated from the bootstrap data were considered not to be robust and need to be trimmed (removed) from the DAG generated in the first step.

The resulting optimal summary network was inferred from data with a total of 14 high-confidence arcs across 10 variables (Figure S3). The DAGs presented using pruning at 50% was constructed from 12,800 searches with a parent limit of four parents per node. Collating results across these 12,800 searches, all 14 arcs were recovered for at least 12,742 times, as resulting from the frequencies matrix at Figure S4.



Variables explanation: mage, maternal age; prebmi, maternal pre-pregnancy BMI; gwg, gestational weight gain; glu, maternal fasting glucose level; hdl, maternal plasma high-density lipoprotein cholesterol level; tg, maternal plasma triglyceride level; bwz, birthweight Z-Score; ins, cord blood insulin; sex, neonatal gender; gaw, gestational age at delivery.

Figure S4 Optimal final DAG (Containing 14 arcs after removal of arcs supported at less than 50% in bootstrapping)

```
> total.dag
        mage prebmi   gwg   glu hdl    tg   bwz ins   sex gaw
mage       0      0     0     0   0     0     0   0     0   0
prebmi 12800      0     0     0   0     0     0   0     0   0
gwg        0  12800     0     0   0     0     0   0     0   0
glu    12800  12800     0     0   0     0     0   0     0   0
hdl        0  12800 12742     0   0 12799     0   0     0   0
tg         0  12800     0     0   0     0     0   0     0   0
bwz        0  12800 12800     0   0     0     0   0     0   0
ins        0      0     0 12800   0     0 12800   0 12800   0
sex        0      0     0     0   0     0     0   0     0   0
gaw        0      0     0     0   0     0     0   0 12800   0
```

Variables explanation: mage, maternal age; prebmi, maternal pre-pregnancy BMI; gwg, gestational weight gain; glu, maternal fasting glucose level; hdl, maternal plasma high-density lipoprotein cholesterol level; tg, maternal plasma triglyceride level; bwz, birthweight Z-Score; ins, cord blood insulin; sex, neonatal gender; gaw, gestational age at delivery.
Rows are children nodes, columns are parent nodes. The number in each cell represents the frequencies at which each arc (from parent node towards child node) was recovered during 12,800 times of bootstrapping.

Figure S5 Frequencies at which each arc in the original DAG was recovered during bootstrapping

Step 3 Estimating marginal from the final DAG

Once the optimal DAG has been identified, we need to examine the strength of the various arcs in our analysis. This process is very similar to when estimating the marginal for the bootstrapping.

Table S1 Association of other maternal metabolic risk factors with birth weight, cord blood insulin level, and the risk of LGA/SGA.

	Pre-pregnancy BMI (Kg/m^2)	Early GWG (Kg)	Glucose (mmol/L)	TC (mmol/L)	LDL-C (mmol/L)
Regression Coefficients(95%CI)					
Birthweight(g)[#]	**29.25 (22.77, 35.73)**	**18.75 (13.06, 24.43)**	**84.32 (42.65, 125.98)**	-0.42 (-19.97, 19.12)	-0.24 (-22.82, 22.34)
Cord blood insulin[##](μU/mL)	0.2 (-0.02, 0.42)	0.08 (-0.11, 0.27)	**2.23 (0.89, 3.57)**	-0.15 (-0.77, 0.48)	-0.15 (-0.88, 0.58)
Odds Ratio (95%CI)					
LGA[§]	**1.24 (1.15, 1.32)**	**1.12 (1.04, 1.20)**	**2.06 (1.31, 3.24)**	1.00 (0.79, 1.25)	1.01 (0.78, 1.31)
SGA[§]	**0.86 (0.78, 0.94)**	0.94 (0.87, 1.00)	0.72 (0.44, 1.18)	0.91 (0.72, 1.15)	0.98 (0.75, 1.27)

BMI, body mass index; GWG, gestational weight gain; HDL-C, high-density lipoprotein cholesterol; TG, triglycerides; LGA, large-for-gestational age; SGA, small-for-gestational age.

[#] Adjusted for maternal age, ethnic group, parity, gestational age, neonatal sex, and early pregnancy cigarette exposures. For gestational weight gain, model was further adjusted for pre-pregnancy BMI and gestational age of maternal weight measurements during pregnancy. Maternal fasting status was further adjusted for TC and LDL-C.

[##] Adjusted for maternal age, ethnic group, parity, gestational age, neonatal sex, early pregnancy cigarette exposures, delivery mode, and sample storage duration. For gestational weight gain, model was further adjusted for pre-pregnancy BMI and gestational age of maternal weight measurements during pregnancy. Maternal fasting status was further adjusted for TC and LDL-C.

[§] Adjusted for maternal age, ethnic group, parity, and early pregnancy cigarette exposures. For gestational weight gain, model was further adjusted for pre-pregnancy BMI and gestational age of maternal weight measurements during pregnancy. Maternal fasting status was further adjusted for TC and LDL-C.

Table S2 **Association of other maternal metabolic parameter Z-Scores with birth weight Z-Score and cord blood insulin Z-Score**

	Pre-pregnancy BMI Z-Score	GWG Z-Score	Glucose Z-Score
Birth weight Z-Score			
Model 1	**0.20(0.15, 0.24)**	**0.17(0.12, 0.22)**	**0.08(0.04, 0.12)**
Model 2	**0.20(0.15, 0.24)**	**0.16(0.11, 0.22)**	**0.04(0.00, 0.09)**
Cord blood insulin Z-Score			
Model 3	**0.10(0.05, 0.15)**	0.05(-0.01, 0.12)	**0.13(0.08, 0.18)**
Model 4	**0.08(0.03, 0.14)**	0.05(-0.02, 0.11)	**0.11(0.06, 0.16)**

BMI, body mass index; GWG, gestational weight gain; HDL-C, high-density lipoprotein cholesterol; TG, triglycerides
Model 1: Adjusted for maternal age, ethic group, parity, and early pregnancy cigarette exposures. For gestational weight gain, model was further adjusted for pre-pregnancy BMI and gestational age of maternal weight measurements during pregnancy.
Model 2: Model 1 + pre-pregnancy BMI Z-Score + GWG Z-Score + Glucose Z-Score + HDL-C Z-Score + TG Z-Score + gestational age of maternal weight measurements during pregnancy.
Model 3: Adjusted for maternal age, ethic group, parity, early pregnancy cigarette exposures, gestational age, neonatal sex, delivery mode, and sample storage duration. For gestational weight gain, model was further adjusted for pre-pregnancy BMI and gestational age of maternal weight measurements during pregnancy.
Model 4: Model 3 + pre-pregnancy BMI Z-Score + GWG Z-Score + Glucose Z-Score + HDL-C Z-Score + TG Z-Score + gestational age of maternal weight measurements during pregnancy.

Table S3 Sensitivity analysis of the association between maternal metabolic parameter Z-Score and birth weight Z-score and cord blood insulin Z-Score.

β(95%CI)	HDL-C Z-Score	TG Z-Score	Pre-pregnancy BMI Z-Score	GWG Z-Score	Glucose Z-Score
Birth weight Z-Score [#]					
Non-GDM	-0.04(-0.09, 0.01)	**0.13(0.08, 0.17)**	**0.20(0.15, 0.25)**	**0.18(0.12, 0.24)**	**0.06(0.01, 0.11)**
GDM	-0.07(-0.18, 0.04)	0.04(-0.09, 0.17)	**0.19(0.09, 0.29)**	0.11(-0.03, 0.25)	**0.16(0.07, 0.24)**
P for interaction	0.57	0.20	0.84	0.62	0.06
Lean	-0.02(-0.07, 0.02)	**0.11(0.06, 0.15)**	**0.23(0.16, 0.29)**	**0.15(0.09, 0.21)**	**0.08(0.03, 0.12)**
Overweight	-0.09(-0.22, 0.05)	0.13(-0.01, 0.27)	**0.20(0.03, 0.36)**	**0.13(0.01, 0.25)**	0.05(-0.07, 0.17)
P for interaction	0.36	0.75	0.73	0.79	0.67
Fasting	**-0.05(-0.10, -0.01)**	**0.12(0.08, 0.16)**	-	-	-
Non-fasting	0.01(-0.12, 0.14)	0.07(-0.07, 0.21)	-	-	-
P for interaction	0.39	0.50	-	-	-
Primiparous	-0.05(-0.09, 0.00)	**0.11(0.07, 0.16)**	**0.18(0.13, 0.22)**	**0.15(0.09, 0.21)**	**0.08(0.03, 0.12)**
Multiparous	-0.04(-0.13, 0.05)	**0.14(0.06, 0.22)**	**0.28(0.18, 0.37)**	0.12(-0.01, 0.25)	**0.11(0.01, 0.21)**
P for interaction	0.87	0.57	0.06	0.74	0.52
Before imputation	**-0.04(-0.09, -0.00)**	**0.12(0.08, 0.16)**	**0.20(0.15, 0.24)**	**0.18(0.12, 0.23)**	**0.08(0.04, 0.12)**
After imputation	**-0.05(-0.09, -0.00)**	**0.12(0.08, 0.16)**	**0.20(0.15, 0.24)**	**0.18(0.13, 0.23)**	**0.08(0.04, 0.12)**
Cord blood insulin Z-Score [##]					
Non-GDM	-0.03(-0.08, 0.03)	**0.06(0.01, 0.12)**	**0.03(0.01, 0.05)**	0.05(-0.01, 0.12)	**0.10(0.04, 0.16)**
GDM	-0.13(-0.26, 0.01)	0.03(-0.11, 0.17)	**0.14(0.03, 0.25)**	0.09(-0.07,0,26)	**0.19(0.09, 0.29)**
P for interaction	0.17	0.66	0.38	0.65	0.13
Lean	-0.03(-0.08, 0.03)	**0.07(0.02, 0.12)**	**0.15(0.07, 0.23)**	0.05(-0.01, 0.12)	**0.13(0.08, 0.19)**
Overweight	-0.12(-0.26, 0.03)	-0.03(-0.13, 0.13)	0.07(-0.09, 0.24)	0.10(-0.04, 0.25)	**0.11(0.00, 0.22)**
P for interaction	0.24	0.34	0.41	0.54	0.72
Fasting	-0.04(-0.09, 0.01)	**0.06(0.01, 0.11)**	-	-	-
Non-fasting	-0.01(-0.17, 0.14)	0.11(-0.04, 0.25)	-	-	-
P for interaction	0.73	0.56	-	-	-
Primiparous	-0.03(-0.08, 0.03)	**0.07(0.01, 0.12)**	**0.10(0.04, 0.15)**	0.06(-0.01, 0.13)	**0.14(0.09, 0.19)**
Multiparous	-0.09(-0.19, -0.00)	0.05(-0.05, 0.14)	**0.10(0.00, 0.20)**	-0.00(-0.12, 0.13)	0.09(-0.01,0.19)
P for interaction	0.26	0.68	0.89	0.39	0.42
Before imputation	-0.04(-0.09, 0.01)	**0.07(0.02, 0.12)**	**0.11(0.06, 0.16)**	**0.07(0.00, 0.13)**	**0.14(0.09, 0.19)**
After imputation	-0.04(-0.09, 0.01)	**0.06(0.01, 0.11)**	**0.10(0.05, 0.15)**	0.06(-0.01, 0.12)	**0.13(0.08, 0.18)**

[#] Adjusted for maternal age, ethic group, parity, early pregnancy cigarette exposures, and delivery mode. For gestational weight gain, model was further adjusted for pre-pregnancy BMI and gestational age of maternal weight measurements during pregnancy. Maternal fasting status was further adjusted for HDL-C and TG (except the analysis between fasting and non-fasting).

[##] Adjusted for maternal age, ethic group, parity, early pregnancy cigarette exposures, gestational age, neonatal gender, delivery mode, and sample storage duration. For gestational weight gain, model was further adjusted for pre-pregnancy BMI and gestational age of maternal weight measurements during pregnancy. Maternal fasting status was further adjusted for HDL-C and TG (except the analysis between fasting and non-fasting).

BMI, body mass index; GWG, gestational weight gain; HDL-C, high-density lipoprotein cholesterol; TG, triglycerides; GDM, gestational diabetes mellitus.

Table S4 Effect estimate of additive Bayesian network analysis.

Arcs	Effect estimate (β, 95%CI)	95% CI
Mage → prebmi	0.19	(0.14, 0.24)
Prebmi → gwg	-0.12	(-0.17, -0.06)
mage → glu	0.11	(0.06, 0.16)
Prebmi → glu	0.14	(0.09, 0.19)
Prebmi → hdl	-0.12	(-0.17, -0.07)
Gwg → hdl	0.09	(0.05, 0.14)
Tg → hdl	-0.33	(-0.38, -0.28)
Prebmi → tg	0.23	(0.18, 0.28)
Prebmi → bwz	0.27	(0.22, 0.32)
Gwg → bwz	0.17	(0.12, 0.22)
Glu → ins	0.12	(0.07, 0.17)
Bwz → ins	0.24	(0.19, 0.29)
Sex → ins	0.19	(0.09, 0.29)
Sex → gaw	0.20	(0.10, 0.31)

Variables explanation: mage, maternal age; prebmi, maternal pre-pregnancy BMI; gwg, gestational weight gain; glu, maternal fasting glucose level; hdl, maternal plasma high-density lipoprotein cholesterol level; tg, maternal plasma triglyceride level; bwz, birthweight Z-Score; ins, cord blood insulin; sex, neonatal gender; gaw, gestational age at delivery.

References

1. Knopp RH, Magee MS, Walden CE, Bonet B, Benedetti TJ. Prediction of infant birth weight by GDM screening tests. Importance of plasma triglyceride. Diabetes Care. 1992;15(11):1605-13.
2. Friedman N, Koller D. Being Bayesian about network structure. A Bayesian approach to structure discovery in Bayesian networks. Machine learning. 2003;50(1-2):95-125.
3. Koivisto M, Sood K. Exact Bayesian structure discovery in Bayesian networks. Journal of Machine Learning Research. 2004;5(May):549-73.
4. Pittavino M, Lewis F, Furrer R. abn: an R package for modelling multivariate data using additive Bayesian networks. The Comprehensive R Archive Network (CRAN), 1–37. 2016.
5. Jansen R, Yu H, Greenbaum D, Kluger Y, Krogan NJ, Chung S, et al. A Bayesian networks approach for predicting protein-protein interactions from genomic data. science. 2003;302(5644):449-53.