


Original Research

Development and Internal Validation of a Mid-Pregnancy Nomogram to Predict Large-for-Gestational-Age Infants in Women With Hyperglycemia in Pregnancy

Yucheng Huo¹, Yuan Li¹, Xia Jin^{2,3}, Xuejing Kuang¹, Fang Li¹, Fenglan Sun¹, Ping Zhou^{1,*} , Risheng Wang⁴¹Department of Obstetrics and Gynecology, The First Hospital of Anhui University of Science & Technology, 232001 Huainan, Anhui, China²Department of Obstetrics and Gynecology, Maternal and Child Health Hospital Affiliated to Anhui Medical University, 230001 Hefei, Anhui, China³Department of Obstetrics and Gynecology, The Fifth Clinical College of Anhui Medical University, 230032 Hefei, Anhui, China⁴School of Computer Science, Huainan Normal University, 232038 Huainan, Anhui, China*Correspondence: FY2019075@aust.edu.cn; zhouping0708@163.com (Ping Zhou)

Academic Editors: Begoña Martínez de Tejada and Michael H. Dahan

Submitted: 27 November 2025 Revised: 4 January 2026 Accepted: 20 January 2026 Published: 17 April 2026

Abstract

Background: Women with hyperglycemia in pregnancy (HIP) have an increased risk of delivering large-for-gestational-age (LGA) infants, a condition associated with both short- and long-term adverse outcomes for mothers and offspring. The identification of women with HIP at high risk of LGA using routinely available clinical indicators could help optimize antenatal management. **Methods:** This retrospective cohort study included 675 women diagnosed with HIP who delivered at Anhui Maternal and Child Health Hospital in Hefei, China, between January 2017 and December 2019. Maternal demographic characteristics and biochemical parameters measured at 25–29⁺⁶ weeks of gestation were extracted from medical records. LGA was defined as birthweight above the 90th percentile for gestational age and sex, according to the INTERGROWTH-21st international newborn standards. Candidate predictors were selected based on clinical relevance and correlation analysis. A multivariable logistic regression model with stepwise selection was used to develop a prediction model for LGA, followed by the construction of a nomogram. Internal validation was performed using 1000 bootstrap resamples. Discrimination, calibration, and clinical utility were assessed using the area under the receiver operating characteristic curve (AUC), Brier score, Hosmer-Lemeshow goodness-of-fit test, calibration plots, and decision curve analysis (DCA). **Results:** Among the 675 women with HIP, 113 (16.7%) delivered LGA infants. The final multivariable model retained 8 predictors: maternal height, serum albumin, lactate dehydrogenase, triglycerides, total cholesterol, blood urea nitrogen, homocysteine, and fasting plasma glucose. The model showed acceptable discrimination, with an AUC of 0.7333 (95% confidence interval [CI]: 0.6856–0.7810) in the original dataset and an optimism-corrected AUC of 0.7132 after 1000 bootstrap resamples. Calibration was acceptable (Brier score 0.1260; Hosmer-Lemeshow $p = 0.0644$), and DCA indicated a net clinical benefit of using the model across a wide range of threshold probabilities. **Conclusions:** We developed and internally validated a simple mid-pregnancy prediction model to estimate the risk of LGA in infants born to women with HIP. This nomogram, based on routinely available clinical and biochemical indicators, may help clinicians identify high-risk women in mid-pregnancy, thus facilitating individualized antenatal management. However, external validation in independent populations is required before broader clinical implementation.

Keywords: hyperglycemia in pregnancy; large-for-gestational-age infants; risk factor; prediction model; nomogram

1. Introduction

Hyperglycemia in pregnancy (HIP), also known as pregnancy with diabetes, includes pregestational diabetes (PGDM) and gestational diabetes mellitus (GDM) [1]. According to the International Diabetes Federation (IDF) Global Diabetes Map 2021, 21.1 million (16.7%) live births worldwide are affected by HIP [2]. In China, approximately 17.5% of pregnant women will experience hyperglycemia. Several specific risks are associated with uncontrolled diabetes during pregnancy, including an increased likelihood of having large for gestational age (LGA) infants [3]. LGA infants have a higher risk of adverse perinatal outcomes, such as birth asphyxia, shoulder dystocia,

and neonatal morbidity. LGA infants also face increased risks of infant death, future obesity, hypertension, and type 2 diabetes [4]. Therefore, identification of the risk of LGA infants is important, as this may provide opportunities for early intervention. Several studies have focused on predicting the occurrence of LGA before delivery. Late pregnancy ultrasound is currently an important tool for estimating LGA [5,6]. However, prenatal ultrasound prediction allows only a short period of time for intervention before delivery. Moreover, its clinical value is limited due to racial disparities in different countries. Therefore, the aim of this study was to identify routine biochemical indicators during the mid-pregnancy period that are associated with LGA in women with HIP. These indicators were then used to de-



velop and internally validate a nomogram to predict LGA risk.

2. Materials and Methods

2.1 Data Source

This retrospective cohort study was conducted at Anhui Maternal and Child Health Hospital, a tertiary maternity hospital in Anhui Province, China. Medical records of pregnant women with HIP who delivered at the hospital between January 2017 and December 2019 were reviewed.

HIP in this study referred to GDM diagnosed during the current pregnancy. Women with known pregestational type 1 or type 2 diabetes were managed in a different clinical pathway and were not included in this cohort. GDM was diagnosed between 24 and 28 weeks of gestation based on a one-step, 75-g oral glucose tolerance test (OGTT) in accordance with the American Diabetes Association Standards of Medical Care in Diabetes, 2022 [1,3].

For the present analysis, blood biochemistry indices measured at 25–29⁺⁶ weeks of gestation were extracted from the hospital electronic database. A total of 675 women with HIP who delivered singleton live births at ≥ 38 weeks of gestation were included in the study. We focused on term and late-term deliveries to reduce heterogeneity arising from preterm birth-related complications, and to ensure that birthweight reflected fetal growth potential rather than early delivery. Based on neonatal birthweight, participants were classified into an LGA group ($n = 113$) or a non-LGA group ($n = 562$). Neonatal weight status was defined according to the INTERGROWTH-21st international standards for newborn weight by gestational age and sex [7].

The inclusion criteria were: (1) singleton live birth; and (2) gestational age at delivery ≥ 38 weeks. The exclusion criteria were: (1) twin or multiple pregnancy; and (2) incomplete clinical or laboratory information.

Women with HIP were managed at our institution according to national and international guidelines. Initial management consisted of individualized medical nutrition therapy and advice on moderate physical activity. Women were instructed to monitor capillary blood glucose at home. When fasting or postprandial glucose values repeatedly exceeded the target levels (fasting < 5.3 mmol/L, 1-hour postprandial < 7.8 mmol/L, 2-hour postprandial < 6.7 mmol/L), insulin therapy was initiated at the discretion of the treating obstetrician or endocrinologist. However, because complete longitudinal records of self-monitored blood glucose and insulin doses were not consistently available in the electronic database, overall glycemic control during pregnancy could not be formally quantified in this retrospective study.

2.2 Variables and Data Collection

General characteristics and biochemical indexes were obtained from medical records. The variables included maternal age, height, gravidity, parity, and the following

laboratory parameters: total protein (TP), albumin (ALB), globulin (GLO), albumin-to-globulin ratio (A/G), alanine aminotransferase (ALT), aspartate aminotransferase (AST), total bilirubin (TBIL), total bile acid (TBA), indirect bilirubin (IBIL), direct bilirubin (DBIL), alkaline phosphatase (ALP), γ -glutamyl transferase (GGT), lactate dehydrogenase (LDH), low-density lipoprotein cholesterol (LDL), high-density lipoprotein cholesterol (HDL), triglycerides (TG), total cholesterol (CHOL), uric acid (UA), creatinine (CREA), blood urea nitrogen (BUN), cystatin C (CysC), creatine kinase (CK), α -hydroxybutyrate dehydrogenase (HBDH), homocysteine (HCY), and fasting plasma glucose (GLU). In total, 29 candidate predictors were considered.

All biochemical measurements were performed on fasting venous blood samples in the same hospital laboratory following standard operating procedures. Continuous variables were summarized as the median (P25–P75) and compared between groups using the Wilcoxon rank-sum test. Categorical variables were expressed as n (%) and compared using the chi-square test or Fisher's exact test, as appropriate. A two-sided p -value < 0.05 was considered statistically significant.

Information on pre-pregnancy weight, body mass index (BMI), gestational weight gain, and glycated hemoglobin (HbA1c) was not systematically recorded in the electronic database and therefore could not be included as candidate predictors in this analysis.

2.3 Model Development

Pearson correlation analyses were performed to explore collinearity among candidate predictors. When two variables were highly correlated (absolute correlation coefficient > 0.7), the variable with stronger clinical interpretability and/or a stronger univariable association with LGA was retained and the other was excluded to reduce multicollinearity. The linearity of each continuous predictor with the outcome logit was then examined using restricted cubic splines.

Subsequently, univariable logistic regression was used to examine the association between each candidate predictor and LGA. Predictors with $p < 0.10$ in the univariable analyses, together with clinically important factors, were entered into a multivariable logistic regression model. Stepwise selection was applied to obtain the final model and reported odds ratio (OR) with 95% confidence interval (CI). To evaluate the models' robustness, multicollinearity was determined using variance inflation factors (VIFs), and influential observations were identified with Cook's distance. Using regression coefficients from the final model, we constructed a nomogram to estimate the probability of delivering an LGA infant among women with HIP. Cases with missing values in any candidate predictor were excluded. With 113 LGA events and 8 predictors retained, the events-per-variable ratio exceeded 10, indicating adequate statistical power for logistic regression.

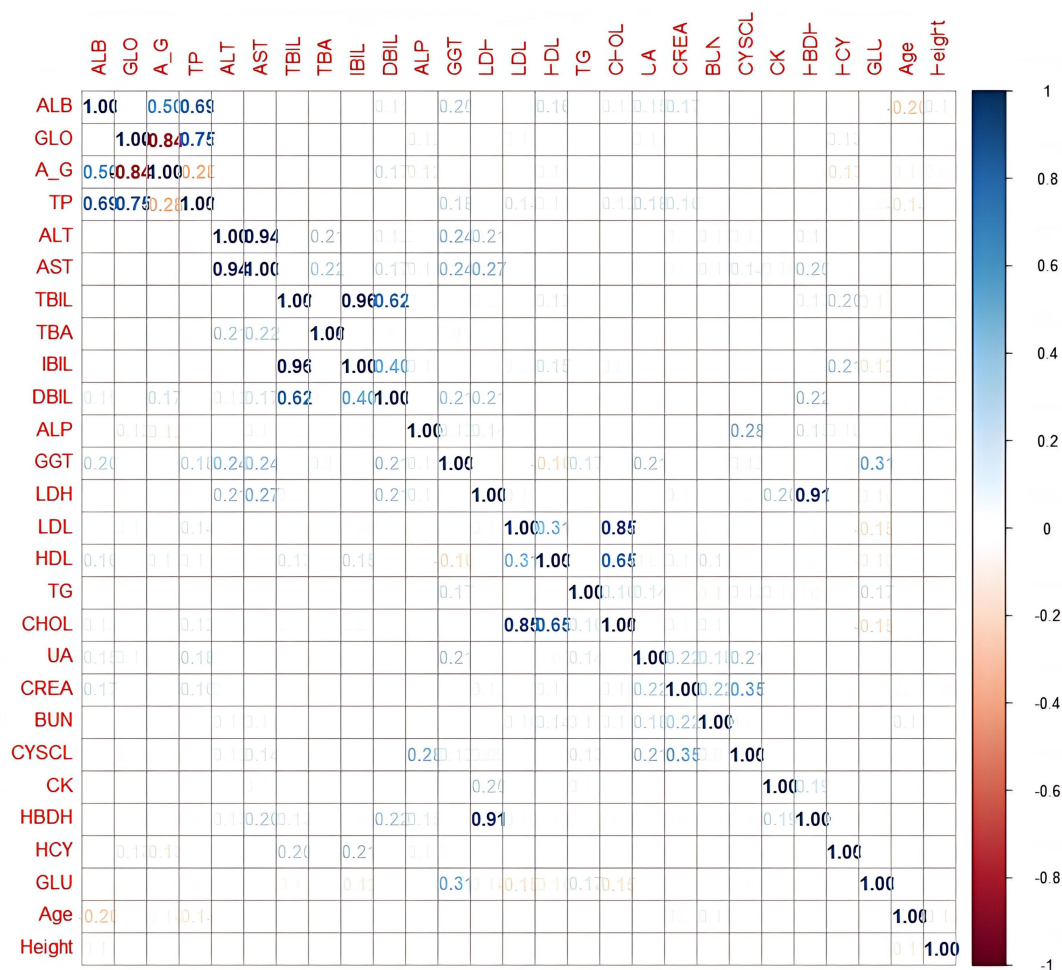


Fig. 1. Pairwise correlation matrix of candidate predictors included in the model development. Darker colors indicate stronger positive or negative correlations. ALB, albumin; GLO, globulin; A_G, albumin-to-globulin ratio; TP, total protein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TBIL, total bilirubin; TBA, total bile acid; IBIL, indirect bilirubin; DBIL, direct bilirubin; ALP, alkaline phosphatase; GGT, γ -glutamyl transferase; LDH, lactate dehydrogenase; LDL, low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; TG, triglycerides; CHOL, total cholesterol; UA, uric acid; CREA, creatinine; BUN, blood urea nitrogen; CysC, cystatin C; CK, creatine kinase; HBDH, α -hydroxybutyrate dehydrogenase; HCY, homocysteine; GLU, fasting plasma glucose.

2.4 Model Evaluation and Statistical Analysis

Model discrimination was evaluated by calculating the area under the receiver operating characteristic curve (AUC). Calibration was assessed using the Brier score, the Hosmer-Lemeshow goodness-of-fit test, and calibration plots that compared predicted and observed risks across deciles of predicted probability. For internal validation, we ran bootstrap resampling with 1000 repetitions and reported optimism-corrected performance estimates. Clinical usefulness was examined by decision curve analysis (DCA), which quantifies the net benefit of applying the model across a range of threshold probabilities relative to treating all patients or none. All analyses were conducted in R (version 4.1.3; R Foundation for Statistical Computing, Vienna, Austria). Model development and validation were reported in accordance with the Transparent Reporting of a

multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement.

3. Results

3.1 Baseline Characteristics of the Study Population

The correlation coefficients between A/G and GLO, TP and GLO, AST and ALT, IBIL and TBIL, HBDH and LDH, and CHOL and LDL were all >0.7 (Fig. 1). With each highly correlated pair, the variable that was more clinically interpretable or more strongly associated with LGA was retained, while the other variable was excluded (A/G, TP, AST, TBIL, HBDH, and LDL). This was carried out to reduce multicollinearity.

Table 1. Basic characteristics of the HIP patients.

Indicator	non-LGA group (n = 562)	LGA group (n = 113)	<i>p</i>
ALB, g/L	39.70 (38.10, 41.50)	39.00 (37.30, 40.50)	0.003
GLO, g/L	26.50 (24.70, 28.40)	25.80 (24.80, 27.40)	0.038
ALT, IU/L	13.40 (10.60, 18.70)	13.10 (9.50, 18.10)	0.200
TBA, μmol/L	2.24 (1.50, 3.26)	2.38 (1.72, 3.71)	0.130
IBIL, μmol/L	6.14 (5.04, 7.79)	5.87 (4.66, 7.30)	0.064
DBIL, μmol/L	1.94 (1.62, 2.33)	1.94 (1.55, 2.33)	0.700
ALP, IU/L	77.25 (63.30, 95.00)	77.70 (58.60, 93.00)	0.500
GGT, IU/L	14.00 (10.70, 19.60)	13.40 (10.10, 21.20)	0.600
LDH, IU/L	160.75 (143.30, 181.05)	168.50 (145.20, 188.90)	0.130
HDL, mmol/L	2.19 (1.90, 2.48)	2.05 (1.81, 2.29)	<0.001
TG, mmol/L	2.72 (2.21, 3.42)	3.09 (2.42, 4.04)	<0.001
CHOL, mmol/L	6.42 (5.62, 7.15)	6.02 (5.49, 6.59)	<0.001
UA, μmol/L	230.85 (199.62, 270.00)	230.50 (204.40, 263.30)	0.800
CREA, μmol/L	39.70 (36.30, 43.77)	38.30 (35.40, 42.50)	0.130
BUN, mmol/L	2.88 (2.44, 3.44)	2.64 (2.34, 3.01)	0.001
CysC, mg/L	5.80 (5.20, 6.40)	6.00 (5.40, 6.60)	0.058
CK, IU/L	39.70 (30.90, 53.68)	39.70 (30.70, 58.60)	0.400
HCY, μmol/L	4.00 (2.70, 4.90)	3.70 (2.60, 4.30)	0.010
GLU, mmol/L	5.01 (4.60, 5.28)	5.07 (4.69, 5.39)	0.022
Age, years	31.58 (29.42, 35.50)	32.25 (29.33, 35.58)	0.500
Gravidity			0.600
Primigravid	207 (36.8%)	41 (36.3%)	
Repregnant	355 (63.2%)	72 (63.7%)	
Parity			0.800
Multiparity	263 (46.8%)	53 (46.9%)	
Primiparity	299 (53.2%)	60 (53.1%)	
Height, cm	161.00 (158.00, 165.00)	163.00 (160.00, 165.00)	0.004

The results of the univariable analysis comparing the two groups are shown in Table 1. Among the general characteristics, only maternal height was significantly different between the LGA and non-LGA groups ($p < 0.05$). For the biochemical indices, TG and GLU levels were significantly higher in the LGA group compared with the non-LGA group, whereas ALB, GLO, HDL, CHOL, BUN, and HCY levels were all significantly lower ($p < 0.05$). No significant differences between the two groups were observed for the remaining laboratory parameters.

Table 2. Multivariate regression analysis of LGA.

Characteristic	OR	95% CI	<i>p</i> -value
ALB	0.86	0.78, 0.94	0.002
LDH	1.01	1.00, 1.02	0.010
TG	1.25	1.07, 1.46	0.005
CHOL	0.77	0.63, 0.94	0.011
BUN	0.65	0.47, 0.89	0.009
HCY	0.83	0.72, 0.94	0.004
GLU	1.35	1.01, 1.77	0.032
Height	1.09	1.04, 1.14	<0.001

Abbreviations: OR, odds ratio; CI, confidence interval.

3.2 Variable Selection and Final Prediction Model

The Pearson correlation test was conducted on 29 independent variables. The ‘rms’ package was employed on the R 4.1.3 platform to fit a regression model via stepwise regression. The final mathematical equation derived from the multivariable logistic regression model to predict LGA infants was: $\text{logit}(p) = -9.5305 - 0.1564 \times \text{ALB (g/L)} + 0.0100 \times \text{LDH (IU/L)} + 0.2201 \times \text{TG (mmol/L)} - 0.2551 \times \text{CHOL (mmol/L)} - 0.4320 \times \text{BUN (mmol/L)} - 0.1922 \times \text{HCY (μmol/L)} + 0.3008 \times \text{GLU (mmol/L)} + 0.0856 \times \text{height (cm)}$ (Table 2). To present the prediction model, we constructed a nomogram that allows quantitative estimation of the risk of LGA in infants born to women with HIP (Fig. 2). Additionally, we developed a website with a calculation tool (<https://China-ahsfybjy-model.shinyapps.io/dynnomapp/>). Restricted cubic splines were used to explore potential non-linear relationships between each continuous predictor and the log odds of LGA (Fig. 3). For all predictors in the final model, the *p*-values for the non-linear components were >0.05 , indicating no strong statistical evidence of departure from linearity within the observed ranges. The variables were also evaluated using the VIF and Cook’s distance. The VIF of the models was significantly below 10 (range 1.0395–1.0976), which suggests there was

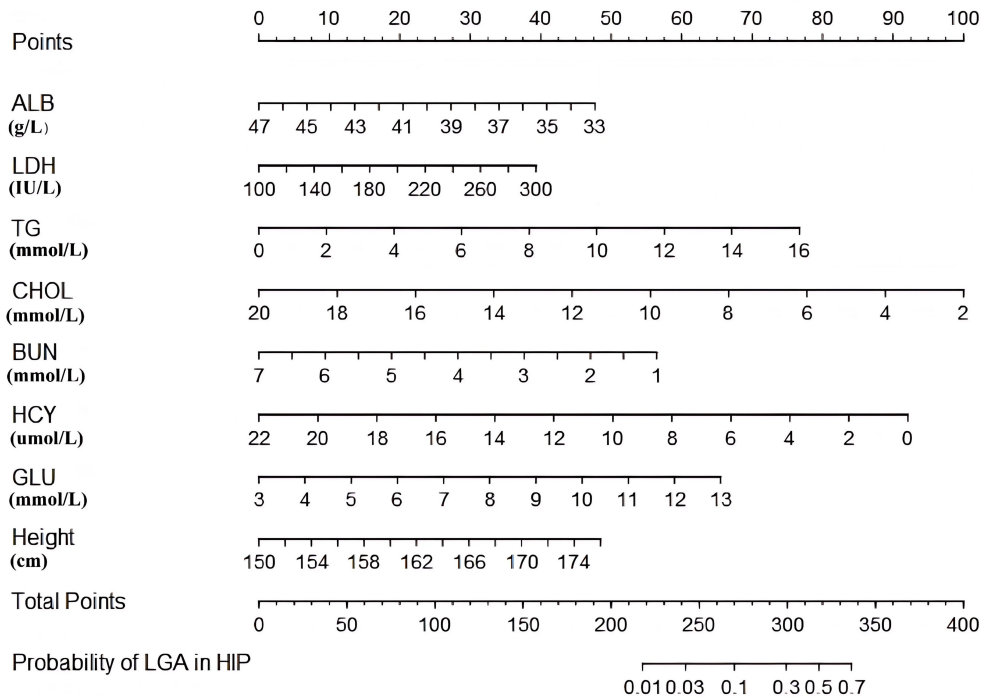


Fig. 2. Nomogram for predicting the risk of delivering a large-for-gestational-age infant in women with hyperglycemia in pregnancy based on maternal height and mid-pregnancy biochemical indicators.

no covariance among the variables. The point-of-influence analysis demonstrated that Cook's distance of the models was considerably <1 , indicating minimal influence of abnormal values in these variables (Fig. 4).

3.3 Discrimination and Calibration of the Model

We assessed discrimination using the area under the receiver operating characteristic (ROC) curve. The nomogram achieved an AUC of 0.7333 (95% CI: 0.6856–0.7810), indicating moderate discriminatory ability (Fig. 5). Model calibration was primarily evaluated using the calibration plot, which showed close agreement between the predicted and observed probabilities of LGA across deciles of predicted risk (Fig. 6). The Brier score was 0.1260, indicating acceptable overall accuracy. The Hosmer Lemeshow goodness-of-fit test yielded $\chi^2 = 14.7410$ (df = 8, $p = 0.0644$), providing no strong evidence of lack of fit. However, this test has important sample size-related limitations and should thus only be interpreted as supplementary information to the graphical assessment.

DCA was used to assess the model's clinical effectiveness, with the results shown in Fig. 7. Application of the nomogram prediction model yields a higher net benefit for women with HIP when the predicted probability ranges from 0.0 to 0.65.

3.4 Model Performance and Internal Validation

Bootstrap resampling was performed 1000 times to validate the model. The ROC curve after resampling (Fig. 8) shows an AUC of 0.7132 (95% CI: 0.7052–0.7329). The Brier score after resampling was calculated as 0.1302. The calibration chart after resampling is depicted in Fig. 9. Hence, the model demonstrates acceptable predictive capability.

4. Discussion

LGA infants are traditionally predicted using late-pregnancy ultrasound, most commonly through estimated fetal weight equations such as the Hadlock formula, or by specific sonographic parameters such as fetal anterior abdominal wall thickness in women with GDM [5,6]. Because these assessments are usually performed in the third trimester, the window for effective intervention is short, and their performance may vary across ethnic populations. More recently, several studies have explored the predictive value of maternal lipid profiles or other biochemical markers for LGA in general obstetric populations [8,9]. Although a study by Oben demonstrated that glycosylated hemoglobin collected at 24–28 weeks gestation can predict the risk of adverse perinatal outcomes in HIP pregnant women, it did not specifically predict the occurrence of LGA alone [10]. However, data that is specific to women with HIP is still scarce. In this context, our study extends the existing literature by developing a nomogram based

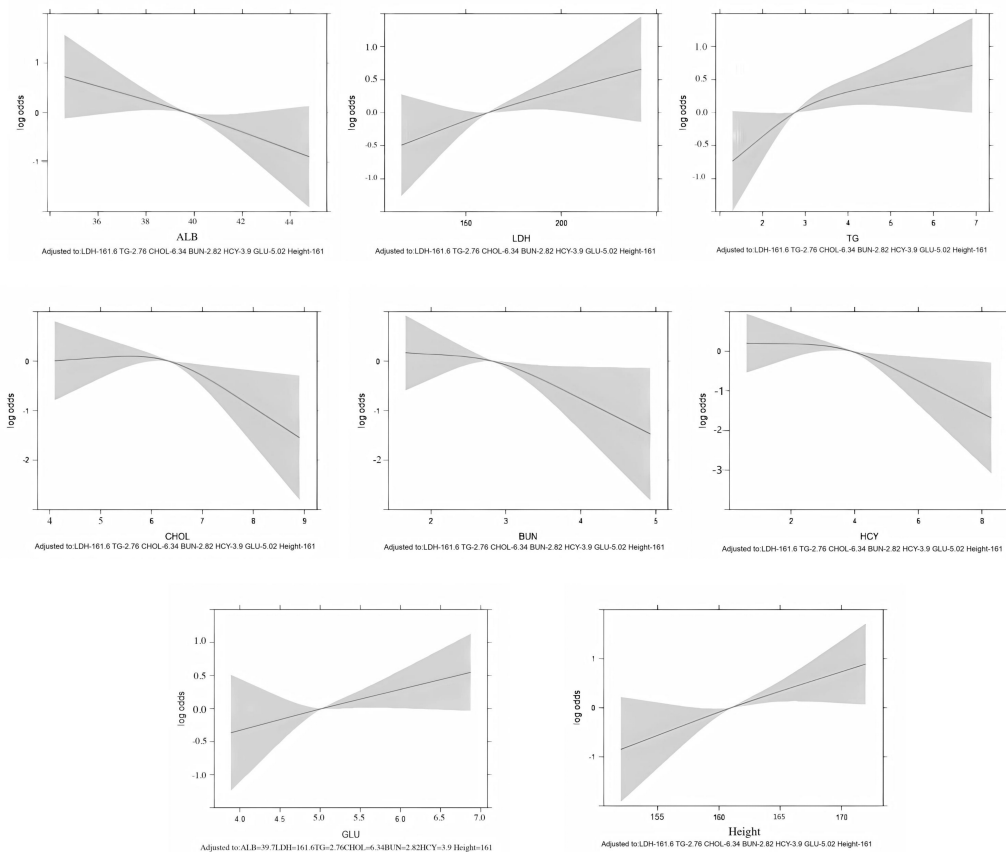


Fig. 3. Restricted cubic spline plots showing the relationship between each continuous predictor and the log odds of large-for-gestational-age birth. No strong evidence of non-linearity was observed.

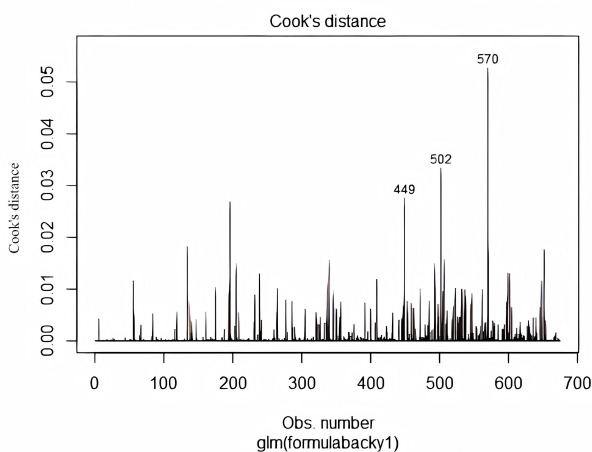


Fig. 4. Cook's distance plots for the final logistic regression model, indicating the influence of individual observations.

solely on routinely measured mid-pregnancy biochemical indicators and height in women with HIP. By focusing on a high-risk group and using indicators that are widely available in clinical practice, our model offers a pragmatic tool for earlier risk stratification and individualized antenatal management.

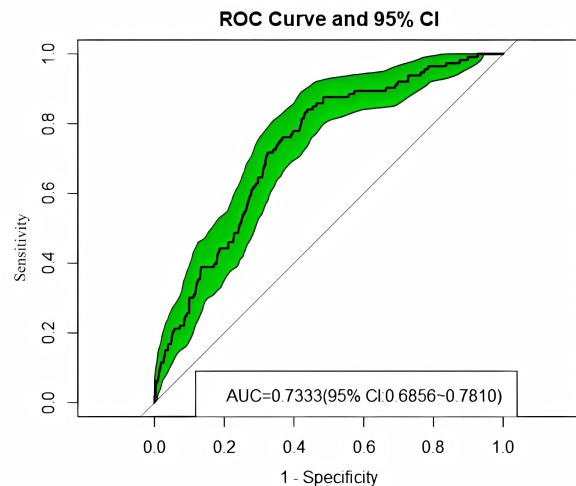


Fig. 5. Receiver operating characteristic (ROC) curve of the prediction model in the original dataset. The shaded area represents the 95% confidence interval for the AUC. AUC, area under the receiver operating characteristic curve.

4.1 Discrimination of the Prediction Model

Multiple logistic regression (MLR) is widely used in health outcomes research, with applications in medicine,

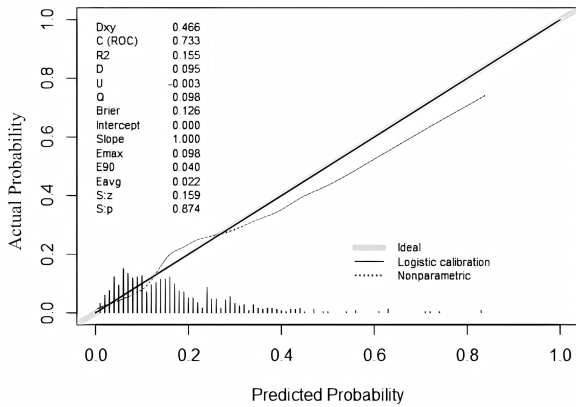


Fig. 6. Calibration plot comparing predicted and observed risks of large-for-gestational-age birth across deciles of predicted probability in the original dataset.

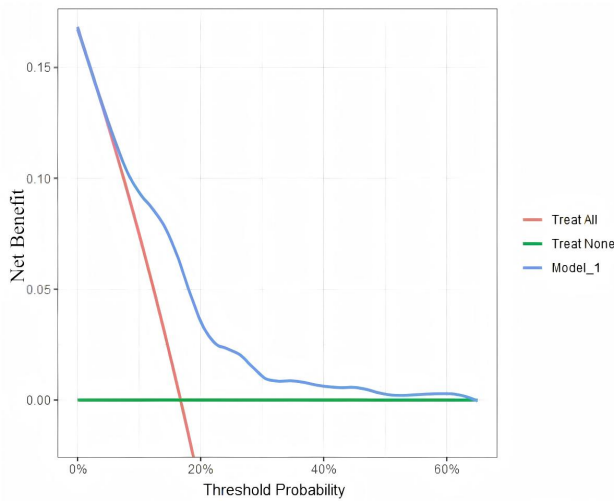


Fig. 7. Decision curve analysis of the prediction model, showing net benefit across a range of threshold probabilities compared with treat-all and treat-none strategies.

statistics, and machine learning (computer science). MLR involves fitting multiple parameters in a predictive model by assuming a linear or additive relationship between predictors and outcomes [11]. In the current study, we built a predictive model for LGA using general information and biochemical indicators for HIP patients extracted from hospital electronic medical records. The AUC of this model was 0.7333 (95% CI: 0.6856–0.7810), indicating moderate discriminative ability. To validate the model’s accuracy, we performed bootstrap validation 1000 times. The AUC value of our model after bootstrapping was 0.7132 (95% CI: 0.7052–0.7329), demonstrating acceptable predictive performance and reasonable stability. Univariate analysis results revealed statistically significant differences in height, ALB, GLO, HDL, TG, CHOL, BUN, HCY, and GLU between the LGA and non-LGA groups. Furthermore, mul-

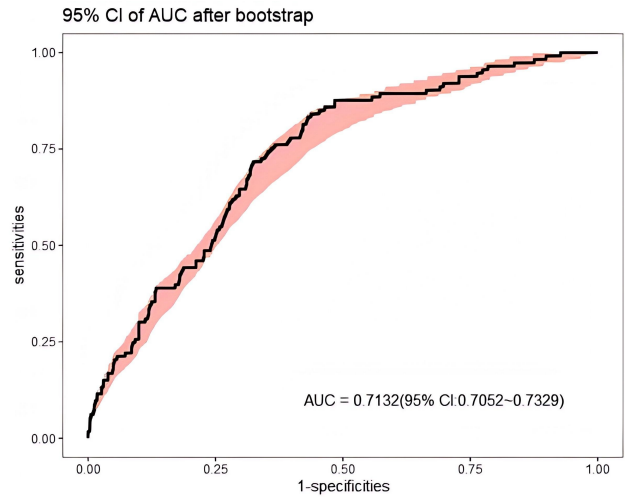


Fig. 8. ROC curve of the model after 1000 bootstrap resamples for internal validation.

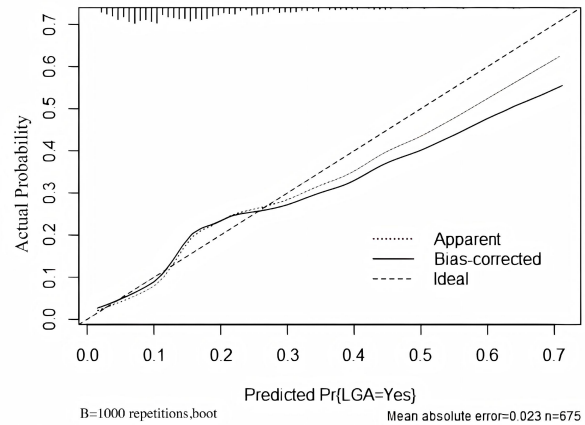


Fig. 9. Calibration plot of the model after 1000 bootstrap resamples for internal validation.

tivariate logistic regression analysis indicated that height, ALB, LDH, TG, CHOL, BUN, HCY, and GLU are significant predictors of LGA risk in HIP women.

4.2 Analysis of Influencing Factors

Some authors have proposed that LGA may be linked to factors such as pre-pregnancy overweight and obesity, excessive weight gain during pregnancy, abnormal glucose metabolism, abnormal lipid metabolism, and genetic factors [12–14]. The results of multivariate regression analysis indicated that maternal height (OR = 1.09, $p < 0.001$) was a risk factor for LGA in HIP patients. Children born to taller HIP women had higher birth weights, which aligns with findings from numerous studies conducted in various countries [15–17]. Miletic *et al.* [18] suggested that the uterine environment of shorter mothers is smaller, thereby limiting the space available for intrauterine growth. It is well known that the weight of pregnant women can vary,

whereas their height remains constant. Therefore, maternal height may be a better predictor of fetal birth weight than maternal weight [19].

Several studies have emphasized the importance of fasting blood glucose (FBG) as a screening tool for GDM and as a predictor of adverse neonatal outcomes [20]. A meta-analysis revealed a positive association between elevated fasting glucose levels in OGTT trials and adverse pregnancy outcomes, including LGA [21]. Fetal glucose exposure and subsequent insulin secretion can lead to excessive conversion of glucose to fat, resulting in fetal weight gain [22]. This finding is consistent with results from the present study.

Substantial literature exists on the relationship between blood lipid levels and GDM. However, there is still some uncertainty regarding the reported results on abnormal blood lipid status in diabetic pregnant women [23]. Some studies have found that women with GDM have elevated triglycerides, LDL, and total cholesterol (CHOL), while HDL levels are decreased [24,25]. A meta-analysis of 4168 women with GDM and 9718 healthy pregnant women showed that compared to women with standard glucose tolerance, those with GDM had higher triglyceride levels and lower HDL levels. However, CHOL and LDL levels were similar between the two groups [26]. Our study also found that high triglyceride levels are a risk factor for LGA in women with HIP (OR = 1.25, $p = 0.005$), in line with the findings of Shi *et al.* [27–29]. A possible reason could be that excess triglycerides in the mother are absorbed, metabolized by the placenta, and transported to the fetus in various forms, leading to excessive fetal growth. However, it remains unclear whether abnormal lipid levels in the blood directly affect fetal body mass.

Additionally, our study found that CHOL was a protective factor against LGA in women with HIP (OR = 0.77, $p = 0.011$). In contrast, Kaneko *et al.* [30] found that among Japanese mothers without GDM who delivered full-term babies, lower and higher maternal CHOL levels in the second trimester were associated with small-for-gestational-age (SGA) and LGA infants, respectively. In contrast, Shi *et al.* [27] found no significant correlations between CHOL or LDL concentrations and perinatal outcomes, including LGA, in either GDM or non-GDM women. These varying research findings may be attributed to pre-pregnancy differences in BMI, age, diet, region, and race. Our study also identified LDH as a risk factor for LGA in women with HIP (OR = 1.01, $p = 0.010$). There is still only limited domestic and international research on the relationship between LDH and LGA. Increased LDH activity indicates metabolic disturbances in glucose utilization and lactate production, which can lead to diabetic complications [31]. This may explain why elevated LDH levels contribute to LGA in pregnant women with HIP.

Human serum albumin (HSA) plays a crucial role in various physiological functions and is vital for maintaining

good health. Hypoalbuminemia occurs due to a combination of inflammation and insufficient nutritional intake, and is commonly observed in patients with chronic diseases. Inflammation is closely associated with vascular diseases and can result in damage to the vascular endothelium [32]. Clinical and biomedical research has indicated a positive regulatory cycle between hyperglycemia and interstitial inflammation. Hyperglycemia triggers the synthesis of proinflammatory cytokines, chemokines, and adipokines in both the placenta and peripheral blood, ultimately affecting fetal growth [33]. In the present study, higher serum albumin levels were inversely associated with the risk of LGA in women with HIP (OR = 0.86, $p = 0.002$). To our knowledge, this specific association has not been reported previously, although study has suggested that albumin may modulate systemic inflammation and oxidative stress [34]. It is therefore plausible that better-preserved albumin levels reflect a more favorable metabolic and inflammatory profile, which might contribute to a lower risk of LGA. However, this finding should be interpreted with caution and confirmed in future studies.

Several authors have investigated the association between GDM and homocysteine (HCY), but the findings have so far been inconsistent. Liu *et al.* [35] found that HCY levels in women with GDM were significantly higher than those in women with standard glucose tolerance. On the other hand, Radzicka *et al.* [36] found no significant correlation between HCY levels and glucose tolerance. In a study of 7587 participants from the Ottawa and Kingston birth cohorts, maternal HCY concentration was found to be associated with a higher incidence of placental-mediated complications, including SGA, preeclampsia, and placental abruption [37]. Our research found similar results, suggesting that HCY acts as a protective factor against LGA (OR = 0.83, $p = 0.004$). This may be attributed to the involvement of HCY in inflammatory reactions, leading to dysfunction of maternal and placental vascular endothelial cells [38], which in turn affects fetal growth and development.

Urea nitrogen is a byproduct of protein metabolism and is excreted by the kidneys. Previous studies have shown that urea nitrogen is an independent risk factor for GDM [39]. However, it remains unclear whether BUN (blood urea nitrogen) can cause LGA. The current study found that BUN was a protective factor for LGA in pregnant women with HIP (OR = 0.65, $p = 0.009$). A possible explanation is that the hyperglycemic state leads to glycosylation of plasma and tissue proteins, impairing their normal function and reducing the excretion of urea nitrogen by the kidneys. Consequently, the level of BUN increases, and excessive amounts enter the placenta, thereby affecting the growth and development of the fetus.

In our model, FBG measured at 25–29⁺⁶ weeks was retained as an independent predictor of LGA. Although fasting glucose is part of the diagnostic workup for HIP, the measurement used in this study reflects glycemic sta-

tus at mid-pregnancy, following the diagnosis and start of lifestyle management or insulin therapy in many women, rather than the diagnostic OGTT value *per se*. We were unable to include HbA1c as an additional marker of longer-term glycemic control because it was not routinely measured in all pregnant women with HIP during the study period. Future studies with more complete information on HbA1c and longitudinal glucose profiles may help to clarify the relative contributions of chronic versus short-term glycemic control to the risk of LGA.

HCY emerged as a significant predictor in our final model. HCY is measured as part of routine biochemical testing during mid-pregnancy at our hospital, thus facilitating its inclusion in the present analysis. However, testing for HCY is not universally available in all antenatal care settings. This may limit the direct transferability of our nomogram to some institutions, and also highlights the need to evaluate simplified models that rely only on universally available predictors.

4.3 Limitations

Our study has several limitations. First, because of the challenges in obtaining comprehensive biochemical data during pregnancy, the sample size was limited to 675 women, and the number of LGA events was modest. We therefore used bootstrap resampling rather than a split-sample approach for internal validation, and external validation in independent cohorts is still required. Second, the analysis was restricted to women who delivered at ≥ 38 weeks of gestation. This design choice reduced heterogeneity related to preterm birth, but may have systematically excluded some higher-risk HIP pregnancies with earlier delivery. This potentially limits the ability to generalize our model to preterm or early-term births.

Third, important clinical factors such as pre-pregnancy BMI, gestational weight gain, HbA1c, and detailed longitudinal information on self-monitored blood glucose and insulin use were not consistently available in the electronic database and could not be incorporated. Consequently, we were unable to formally quantify overall glycemic control during pregnancy or evaluate its interaction with the biochemical predictors, which may have led to residual confounding. Fourth, although several predictors in the final model (albumin, BUN, HCY) are biologically linked to endothelial dysfunction and hypertensive disorders of pregnancy, we could not reliably ascertain the occurrence of preeclampsia or other hypertensive disorders in this retrospective cohort. The observed associations should therefore be interpreted with caution, as they may partially reflect underlying placental pathology.

Fifth, we focused exclusively on LGA and did not analyze SGA or fetal growth restriction (FGR) as competing fetal growth outcomes. Some of the predictors associated with a lower risk of LGA in our model may in fact

be markers of placental dysfunction and increased risk of SGA/FGR. Future studies with larger sample sizes are warranted to simultaneously evaluate predictors of both excessive and restricted fetal growth in women with HIP. Finally, to handle multicollinearity, we used a simple, correlation-based strategy to choose between pairs of highly correlated predictors. Although this approach is transparent and easy to implement, it may not be optimal from a predictive modeling perspective. Instead, penalized regression methods such as LASSO or elastic net could be explored in future work to further refine variable selection.

5. Conclusions

Based on analysis of a retrospective cohort of 675 women with HIP, we developed and internally validated a nomogram that estimates the risk of delivering an LGA infant using only maternal height and routinely measured biochemical indicators at 25–29⁺⁶ weeks of gestation. The model demonstrated acceptable discrimination and calibration, as well as favorable clinical utility in decision curve analysis. Because the predictors are easily available during routine care, this tool may facilitate the early identification of high-risk women and support individualized antenatal management. External validation in larger, multicenter cohorts is warranted to confirm the generalizability of our model and refine its clinical application.

Availability of Data and Materials

The datasets used and analyzed in the current study are available from the corresponding author upon reasonable request.

Author Contributions

YCH, YL, XJ, XJK, FL, FLS, PZ, RSW contributed to the study concept and design. YL and XJ collected the data. RSW participated in data processing and analysis. YCH wrote the first draft of the manuscript. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

This retrospective study was conducted in accordance with the principles of the Declaration of Helsinki. This study protocol was approved by the ethical committee of Maternal and Child Health Hospital Affiliated to Anhui Medical University, with ethical approval numbers YYLL2022-2108085MH260-03-01. The requirement for written informed consent was waived for this retrospective study due to the complete anonymization of all patient clinical data.

Acknowledgment

The authors wish to thank all the staff members of the Department of Obstetrics and Gynecology at Anhui Maternal and Child Health Care Hospital for their strong support of this study.

Funding

This work was supported by the Anhui Province Key Research and Development Project (202204295107020050), Anhui Province Key Research and Development Project (No.201904a07020032), and the Higher Education Science Research Project of Anhui Province (NO. 2023AH053400).

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] American Diabetes Association Professional Practice Committee. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2022. *Diabetes Care*. 2022; 45: S17–S38. <https://doi.org/10.2337/dc22-S002>.
- [2] Magliano DJ, Boyko EJ. IDF Diabetes Atlas 10th edition scientific committee. IDF DIABETES ATLAS. 10th edn. International Diabetes Federation: Brussels. 2021.
- [3] American Diabetes Association. 14. Management of Diabetes in Pregnancy: *Standards of Medical Care in Diabetes-2021*. *Diabetes Care*. 2021; 44: S200–S210. <https://doi.org/10.2337/dc21-S014>.
- [4] Åmark H, Westgren M, Persson M. Prediction of large-for-gestational-age infants in pregnancies complicated by obesity: A population-based cohort study. *Acta Obstetrica et Gynecologica Scandinavica*. 2019; 98: 769–776. <https://doi.org/10.1111/aogs.13546>.
- [5] Hadlock FP, Harrist RB, Sharman RS, Deter RL, Park SK. Estimation of fetal weight with the use of head, body, and femur measurements—a prospective study. *American Journal of Obstetrics and Gynecology*. 1985; 151: 333–337. [https://doi.org/10.1016/0002-9378\(85\)90298-4](https://doi.org/10.1016/0002-9378(85)90298-4).
- [6] Bansal A, Thukral BB, Bagri N, Kanwar A, Khandelwal A, Bajaj B. Correlation of Fetal Anterior Abdominal Wall Thickness and Other Standard Biometric Ultrasound Measurements to Predict Fetal Macrosomia in Gestational Diabetes. *Journal of Medical Ultrasound*. 2023; 31: 29–34. https://doi.org/10.4103/JMU.JMU_57_21.
- [7] Villar J, Cheikh Ismail L, Victora CG, Ohuma EO, Bertino E, Altman DG, *et al.* International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. *Lancet (London, England)*. 2014; 384: 857–868. [https://doi.org/10.1016/S0140-6736\(14\)60932-6](https://doi.org/10.1016/S0140-6736(14)60932-6).
- [8] Zhu YT, Xiang LL, Chen YJ, Zhong TY, Wang JJ, Zeng Y. Developing and validating a predictive model of delivering large-for-gestational-age infants among women with gestational diabetes mellitus. *World Journal of Diabetes*. 2024; 15: 1242–1253. <https://doi.org/10.4239/wjd.v15.i6.1242>.
- [9] Wang N, Guo H, Jing Y, Zhang Y, Sun B, Pan X, *et al.* Development and validation of risk prediction models for large for gestational age infants using logistic regression and two machine learning algorithms. *Journal of Diabetes*. 2023; 15: 338–348. <https://doi.org/10.1111/1753-0407.13375>.
- [10] Oben A, Jauk V, Battarbee A, Longo S, Szychowski J, Tita A, *et al.* Value of HbA1c in Obese Women with Gestational Diabetes. *American Journal of Perinatology*. 2022; 39: 449–456. <https://doi.org/10.1055/s-0041-1740060>.
- [11] Sufriyana H, Husnayain A, Chen YL, Kuo CY, Singh O, Yeh TY, *et al.* Comparison of Multivariable Logistic Regression and Other Machine Learning Algorithms for Prognostic Prediction Studies in Pregnancy Care: Systematic Review and Meta-Analysis. *JMIR Medical Informatics*. 2020; 8: e16503. <https://doi.org/10.2196/16503>.
- [12] Dai RX, He XJ, Hu CL. Maternal pre-pregnancy obesity and the risk of macrosomia: a meta-analysis. *Archives of Gynecology and Obstetrics*. 2018; 297: 139–145. <https://doi.org/10.1007/s00404-017-4573-8>.
- [13] Tian C, Hu C, He X, Zhu M, Qin F, Liu Y, *et al.* Excessive weight gain during pregnancy and risk of macrosomia: a meta-analysis. *Archives of Gynecology and Obstetrics*. 2016; 293: 29–35. <http://doi.org/10.1007/s00404-015-3825-8>.
- [14] Sun Y, Shen Z, Zhan Y, Wang Y, Ma S, Zhang S, *et al.* Effects of pre-pregnancy body mass index and gestational weight gain on maternal and infant complications. *BMC Pregnancy and Childbirth*. 2020; 20: 390. <https://doi.org/10.1186/s12884-020-03071-y>.
- [15] Softa SM, Aldardeir N, Aloufi FS, Alshihabi SS, Khouj M, Radwan E. The Association of Maternal Height With Mode of Delivery and Fetal Birth Weight at King Abdulaziz University Hospital, Jeddah, Saudi Arabia. *Cureus*. 2022; 14: e27493. <https://doi.org/10.7759/cureus.27493>.
- [16] Kuritani Y, Hayashi S, Yamamoto R, Mitsuda N, Ishii K. Association between maternal height and mode of delivery in nulliparous Japanese women. *The Journal of Obstetrics and Gynecology Research*. 2020; 46: 2645–2650. <https://doi.org/10.1111/jog.14511>.
- [17] Henriksson P, Sandborg J, Blomberg M, Nowicka P, Petersson K, Bendtsen M, *et al.* Body mass index and gestational weight gain in migrant women by birth regions compared with Swedish-born women: A registry linkage study of 0.5 million pregnancies. *PloS One*. 2020; 15: e0241319. <https://doi.org/10.1371/journal.pone.0241319>.
- [18] Miletić T, Stoini E, Mikulandra F, Tadin I, Roje D, Milić N. Effect of parental anthropometric parameters on neonatal birth weight and birth length. *Collegium Antropologicum*. 2007; 31: 993–997.
- [19] Rochow N, AlSamnan M, So HY, Olbertz D, Pelc A, Däbritz J, *et al.* Maternal body height is a stronger predictor of birth weight than ethnicity: analysis of birth weight percentile charts. *Journal of Perinatal Medicine*. 2018; 47: 22–29. <https://doi.org/10.1515/jpm-2017-0349>.
- [20] Cho HY, Jung I, Kim SJ. The association between maternal hyperglycemia and perinatal outcomes in gestational diabetes mellitus patients: A retrospective cohort study. *Medicine*. 2016; 95: e4712. <https://doi.org/10.1097/MD.0000000000004712>.
- [21] Malaza N, Masete M, Adam S, Dias S, Nyawo T, Pheiffer C. A Systematic Review to Compare Adverse Pregnancy Outcomes in Women with Pregestational Diabetes and Gestational Diabetes. *International Journal of Environmental Research and Public Health*. 2022; 19: 10846. <https://doi.org/10.3390/ijerph191710846>.
- [22] Bhavadharini B, Anjana RM, Deepa M, Pradeepa R, Uma R, Saravanan P, *et al.* Association between number of abnormal glucose values and severity of fasting plasma glucose in IADPSG criteria and maternal outcomes in women with gestational diabetes mellitus. *Acta Diabetologica*. 2022; 59: 349–357. <https://doi.org/10.1007/s00592-021-01815-6>.
- [23] Herrera E, Desoye G. Maternal and fetal lipid metabolism under normal and gestational diabetic conditions. *Hormone Molecular*

- Biology and Clinical Investigation. 2016; 26: 109–127. <https://doi.org/10.1515/hmbci-2015-0025>.
- [24] Hu J, Gillies CL, Lin S, Stewart ZA, Melford SE, Abrams KR, *et al.* Association of maternal lipid profile and gestational diabetes mellitus: A systematic review and meta-analysis of 292 studies and 97,880 women. *EClinicalMedicine*. 2021; 34: 100830. <https://doi.org/10.1016/j.eclinm.2021.100830>.
- [25] Ji J, Wu P, Li G, He Z, Wang S, Yu W, *et al.* The associations of ferritin, serum lipid and plasma glucose levels across pregnancy in women with gestational diabetes mellitus and newborn birth weight. *BMC Pregnancy and Childbirth*. 2023; 23: 478. <https://doi.org/10.1186/s12884-023-05806-z>.
- [26] Ryckman KK, Spracklen CN, Smith CJ, Robinson JG, Saftlas AF. Maternal lipid levels during pregnancy and gestational diabetes: a systematic review and meta-analysis. *BJOG: an International Journal of Obstetrics and Gynaecology*. 2015; 122: 643–651. <https://doi.org/10.1111/1471-0528.13261>.
- [27] Shi P, Tang J, Yin X. Association between second- and third-trimester maternal lipid profiles and adverse perinatal outcomes among women with GDM and non-GDM: a retrospective cohort study. *BMC Pregnancy and Childbirth*. 2023; 23: 318. <https://doi.org/10.1186/s12884-023-05630-5>.
- [28] Wang Z, Peng Y, Mao S, Zhang L, Guo Y. The correlation between blood-lipid ratio in the first trimester and large-for-gestational-age infants. *Lipids in Health and Disease*. 2023; 22: 18. <https://doi.org/10.1186/s12944-023-01781-8>.
- [29] Adank MC, Benschop L, Kors AW, Peterbroers KR, Smak Gre-goer AM, Mulder MT, *et al.* Maternal lipid profile in early pregnancy is associated with foetal growth and the risk of a child born large-for-gestational age: a population-based prospective cohort study: Maternal lipid profile in early pregnancy and foetal growth. *BMC Medicine*. 2020; 18: 276. <https://doi.org/10.1186/s12916-020-01730-7>.
- [30] Kaneko K, Ito Y, Ebara T, Kato S, Matsuki T, Tamada H, *et al.* Association of Maternal Total Cholesterol With SGA or LGA Birth at Term: the Japan Environment and Children's Study. *The Journal of Clinical Endocrinology and Metabolism*. 2022; 107: e118–e129. <https://doi.org/10.1210/clinem/dgab618>.
- [31] Rafaqat S, Sattar A, Khalid A, Rafaqat S. Role of liver pa-rameters in diabetes mellitus - a narrative review. *Endocrine Regulations*. 2023; 57: 200–220. <https://doi.org/10.2478/enr-2023-0024>.
- [32] Don BR, Kaysen G. POOR NUTRITIONAL STATUS AND IN-FLAMMATION: Serum Albumin: Relationship to Inflammation and Nutrition. *Seminars in Dialysis*. 2004; 17: 432–437. <https://doi.org/10.1111/j.0894-0959.2004.17603.x>.
- [33] Olmos-Ortiz A, Flores-Espinosa P, Díaz L, Velázquez P, Ramírez-Isarraraz C, Zaga-Clavellina V. Immunoendocrine Dysregulation during Gestational Diabetes Mellitus: The Central Role of the Placenta. *International Journal of Molecular Sciences*. 2021; 22: 8087. <https://doi.org/10.3390/ijms22158087>.
- [34] Watanabe K, Kinoshita H, Okamoto T, Sugiura K, Kawashima S, Kimura T. Antioxidant Properties of Albumin and Diseases Related to Obstetrics and Gynecology. *Antioxidants (Basel)*. 2025; 14: 55. <https://doi.org/10.3390/antiox14010055>.
- [35] Liu W, Huang Z, Tang S, Zhang Z, Yu Q, He J. Changes of Serum Sex Hormone-Binding Globulin, Homocysteine, and Hypersensitive CRP Levels during Pregnancy and Their Relationship with Gestational Diabetes Mellitus. *Gynecologic and Obstetric Investigation*. 2021; 86: 193–199. <https://doi.org/10.1159/000515085>.
- [36] Radzicka S, Ziolkowska K, Zaborowski MP, Brazert J, Pietryga M. Serum homocysteine and vitamin B12 levels in women with gestational diabetes mellitus. *Ginekologia Polska*. 2019; 90: 381–387. <https://doi.org/10.5603/GP.2019.0066>.
- [37] Chaudhry SH, Taljaard M, MacFarlane AJ, Gaudet LM, Smith GN, Rodger M, *et al.* The role of maternal homocysteine concentration in placenta-mediated complications: findings from the Ottawa and Kingston birth cohort. *BMC Pregnancy and Childbirth*. 2019; 19: 75. <https://doi.org/10.1186/s12884-019-2219-5>.
- [38] Chatterjee B, Fatima F, Seth S, Sinha Roy S. Moderate Elevation of Homocysteine Induces Endothelial Dysfunction through Adaptive UPR Activation and Metabolic Rewiring. *Cells*. 2024; 13: 214. <https://doi.org/10.3390/cells13030214>.
- [39] Feng P, Wang G, Yu Q, Zhu W, Zhong C. First-trimester blood urea nitrogen and risk of gestational diabetes mellitus. *Journal of Cellular and Molecular Medicine*. 2020; 24: 2416–2422. <https://doi.org/10.1111/jcmm.14924>.