

World Journal of *Diabetes*

World J Diabetes 2023 October 15; 14(10): 1450-1584



OPINION REVIEW

- 1450 Multifaceted relationship between diabetes and kidney diseases: Beyond diabetes
Esposito P, Picciotto D, Cappadona F, Costigliolo F, Russo E, Macciò L, Viazzi F

REVIEW

- 1463 Partners in diabetes epidemic: A global perspective
Wang H, Akbari-Alavijeh S, Parhar RS, Gaugler R, Hashmi S
- 1478 Role of glycolysis in diabetic atherosclerosis
Liu QJ, Yuan W, Yang P, Shao C

MINIREVIEWS

- 1493 Accessibility and utilization of healthcare services among diabetic patients: Is diabetes a poor man's ailment?
Eseadi C, Amedu AN, Ilechukwu LC, Ngwu MO, Ossai OV
- 1502 Gut microbiome supplementation as therapy for metabolic syndrome
Antony MA, Chowdhury A, Edem D, Raj R, Nain P, Joglekar M, Verma V, Kant R

ORIGINAL ARTICLE**Retrospective Study**

- 1514 Effects of vitamin D supplementation on glucose and lipid metabolism in patients with type 2 diabetes mellitus and risk factors for insulin resistance
Sun LJ, Lu JX, Li XY, Zheng TS, Zhan XR
- 1524 Effect of individualized nutrition interventions on clinical outcomes of pregnant women with gestational diabetes mellitus
Luo JY, Chen LG, Yan M, Mei YJ, Cui YQ, Jiang M
- 1532 Effects of insulin aspart and metformin on gestational diabetes mellitus and inflammatory markers
Wang Y, Song M, Qi BR
- 1541 Establishment and evaluation of a risk prediction model for gestational diabetes mellitus
Lin Q, Fang ZJ
- 1551 Analysis of influencing factors and interaction of body weight and disease outcome in patients with prediabetes
Li YY, Tong LP, Wu XD, Lin D, Lin Y, Lin XY

Observational Study

- 1562** Characteristics of glucose change in diabetes mellitus generalized through continuous wavelet transform processing: A preliminary study
Nakamura Y, Furukawa S

META-ANALYSIS

- 1573** Indirect comparison of efficacy and safety of chiglitazar and thiazolidinedione in patients with type 2 diabetes: A meta-analysis
Lin C, Li ZL, Cai XL, Hu SY, Lv F, Yang WJ, Ji LN

ABOUT COVER

Editorial Board Member of *World Journal of Diabetes*, Pradeep Kumar Dabla, MD, PGDHM, PGDBA, MAMS, FACBI, FIMSA, Professor, Department of Biochemistry, G B Pant Institute of Postgraduate Medical Education and Research, GNCTD, Delhi 110002, India. pradeep_dabla@yahoo.com

AIMS AND SCOPE

The primary aim of *World Journal of Diabetes (WJD, World J Diabetes)* is to provide scholars and readers from various fields of diabetes with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

WJD mainly publishes articles reporting research results and findings obtained in the field of diabetes and covering a wide range of topics including risk factors for diabetes, diabetes complications, experimental diabetes mellitus, type 1 diabetes mellitus, type 2 diabetes mellitus, gestational diabetes, diabetic angiopathies, diabetic cardiomyopathies, diabetic coma, diabetic ketoacidosis, diabetic nephropathies, diabetic neuropathies, Donohue syndrome, fetal macrosomia, and prediabetic state.

INDEXING/ABSTRACTING

The *WJD* is now abstracted and indexed in Science Citation Index Expanded (SCIE, also known as SciSearch®), Current Contents/Clinical Medicine, Journal Citation Reports/Science Edition, PubMed, PubMed Central, Reference Citation Analysis, China National Knowledge Infrastructure, China Science and Technology Journal Database, and Superstar Journals Database. The 2023 Edition of Journal Citation Reports® cites the 2022 impact factor (IF) for *WJD* as 4.2; IF without journal self cites: 4.1; 5-year IF: 4.5; Journal Citation Indicator: 0.69; Ranking: 51 among 145 journals in endocrinology and metabolism; and Quartile category: Q2.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Yu-Xi Chen*; Production Department Director: *Xu Guo*; Editorial Office Director: *Jia-Ru Fan*.

NAME OF JOURNAL

World Journal of Diabetes

ISSN

ISSN 1948-9358 (online)

LAUNCH DATE

June 15, 2010

FREQUENCY

Monthly

EDITORS-IN-CHIEF

Lu Cai, Md. Shahidul Islam, Michael Horowitz

EDITORIAL BOARD MEMBERS

<https://www.wjgnet.com/1948-9358/editorialboard.htm>

PUBLICATION DATE

October 15, 2023

COPYRIGHT

© 2023 Baishideng Publishing Group Inc

INSTRUCTIONS TO AUTHORS

<https://www.wjgnet.com/bpg/gerinfo/204>

GUIDELINES FOR ETHICS DOCUMENTS

<https://www.wjgnet.com/bpg/GerInfo/287>

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH

<https://www.wjgnet.com/bpg/gerinfo/240>

PUBLICATION ETHICS

<https://www.wjgnet.com/bpg/GerInfo/288>

PUBLICATION MISCONDUCT

<https://www.wjgnet.com/bpg/gerinfo/208>

ARTICLE PROCESSING CHARGE

<https://www.wjgnet.com/bpg/gerinfo/242>

STEPS FOR SUBMITTING MANUSCRIPTS

<https://www.wjgnet.com/bpg/GerInfo/239>

ONLINE SUBMISSION

<https://www.f6publishing.com>

Retrospective Study

Establishment and evaluation of a risk prediction model for gestational diabetes mellitus

Qing Lin, Zhuan-Ji Fang

Specialty type: Endocrinology and metabolism**Provenance and peer review:**

Unsolicited article; Externally peer reviewed.

Peer-review model: Single blind**Peer-review report's scientific quality classification**Grade A (Excellent): 0
Grade B (Very good): B, B
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0**P-Reviewer:** Abdelsalam MM, Egypt; Bruning JC, Germany; Lu C, United States**Received:** August 1, 2023**Peer-review started:** August 1, 2023**First decision:** August 16, 2023**Revised:** August 21, 2023**Accepted:** September 14, 2023**Article in press:** September 14, 2023**Published online:** October 15, 2023**Qing Lin, Zhuan-Ji Fang**, Department of Obstetrics, Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics and Gynecology and Pediatrics, Fujian Medical University, Fuzhou 350001, Fujian Province, China**Corresponding author:** Zhuan-Ji Fang, MM, Associate Chief Physician, Department of Obstetrics, Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics and Gynecology and Pediatrics, Fujian Medical University, No. 18 Daoshan Rd., Gulou Dist, Fuzhou 350001, Fujian Province, China. fz1qtg@163.com**Abstract****BACKGROUND**

Gestational diabetes mellitus (GDM) is a condition characterized by high blood sugar levels during pregnancy. The prevalence of GDM is on the rise globally, and this trend is particularly evident in China, which has emerged as a significant issue impacting the well-being of expectant mothers and their fetuses. Identifying and addressing GDM in a timely manner is crucial for maintaining the health of both expectant mothers and their developing fetuses. Therefore, this study aims to establish a risk prediction model for GDM and explore the effects of serum ferritin, blood glucose, and body mass index (BMI) on the occurrence of GDM.

AIM

To develop a risk prediction model to analyze factors leading to GDM, and evaluate its efficiency for early prevention.

METHODS

The clinical data of 406 pregnant women who underwent routine prenatal examination in Fujian Maternity and Child Health Hospital from April 2020 to December 2022 were retrospectively analyzed. According to whether GDM occurred, they were divided into two groups to analyze the related factors affecting GDM. Then, according to the weight of the relevant risk factors, the training set and the verification set were divided at a ratio of 7:3. Subsequently, a risk prediction model was established using logistic regression and random forest models, and the model was evaluated and verified.

RESULTS

Pre-pregnancy BMI, previous history of GDM or macrosomia, hypertension, hemoglobin (Hb) level, triglyceride level, family history of diabetes, serum ferritin, and fasting blood glucose levels during early pregnancy were de-

terminated. These factors were found to have a significant impact on the development of GDM ($P < 0.05$). According to the nomogram model's prediction of GDM in pregnancy, the area under the curve (AUC) was determined to be 0.883 [95% confidence interval (CI): 0.846-0.921], and the sensitivity and specificity were 74.1% and 87.6%, respectively. The top five variables in the random forest model for predicting the occurrence of GDM were serum ferritin, fasting blood glucose in early pregnancy, pre-pregnancy BMI, Hb level and triglyceride level. The random forest model achieved an AUC of 0.950 (95%CI: 0.927-0.973), the sensitivity was 84.8%, and the specificity was 91.4%. The Delong test showed that the AUC value of the random forest model was higher than that of the decision tree model ($P < 0.05$).

CONCLUSION

The random forest model is superior to the nomogram model in predicting the risk of GDM. This method is helpful for early diagnosis and appropriate intervention of GDM.

Key Words: Gestational diabetes mellitus; Prediction model; Model evaluation; Random forest model; Nomograms; Risk factor

©The Author(s) 2023. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: Gestational diabetes mellitus (GDM) is a common pregnancy complication, which has an important impact on maternal and child health. Early prediction of GDM can result in timely interventions in patients and improve pregnancy outcomes. This study examined various risk factors associated with GDM and established and compared two prediction models: The nomogram model and the random forest model. The random forest model has good predictive ability, which can effectively predict the risk of GDM and provide accurate references for early prevention and management of GDM.

Citation: Lin Q, Fang ZJ. Establishment and evaluation of a risk prediction model for gestational diabetes mellitus. *World J Diabetes* 2023; 14(10): 1541-1550

URL: <https://www.wjgnet.com/1948-9358/full/v14/i10/1541.htm>

DOI: <https://dx.doi.org/10.4239/wjd.v14.i10.1541>

INTRODUCTION

Gestational diabetes mellitus (GDM) is a metabolic disease that occurs or is first discovered during pregnancy[1,2] and is a risk factor for many adverse pregnancy outcomes. International data show that by 2021, the proportion of pregnant women with GDM worldwide has reached 16.7% and continues to grow[3]. Preventing GDM has become an important challenge for global health. At present, numerous studies have been conducted worldwide to predict the likelihood of GDM[4,5], but these studies are only applicable to foreign populations, and their applicability to domestic populations is not ideal. There are relatively few studies on the risk prediction of GDM in China, which needs to be further strengthened. Therefore, the objective of this investigation is to establish a predictive model for GDM risk. By comparing the predictive efficacy of the nomogram model and the random forest model, this will provide clinicians with a more scientific and accurate risk prediction tool for GDM, promote early diagnosis and intervention of GDM, and provide pregnant women with corresponding intervention measures and health education.

MATERIALS AND METHODS

General information

A retrospective analysis of 406 pregnant women aged 22-43 years, with an average age of (31.17 ± 4.02) years, who underwent a routine prenatal examination in our hospital was conducted from April 2020 to December 2022. According to whether GDM occurred, they were divided into two groups, including the GDM group ($n = 197$) and the non-GDM group ($n = 209$).

Inclusion criteria were: (1) Normal pregnant women; and (2) natural pregnancy. Exclusion criteria were: (1) Patients with diabetes who had been diagnosed or were receiving treatment before pregnancy; (2) women who could not participate in the survey and follow-up; (3) adolescent pregnant women (< 18 years old); (4) those suffering from other chronic diseases, such as cardiovascular disease, liver disease, renal dysfunction, or malignant neoplasms; and (5) pregnant women who used hormones and immunosuppressants.

Research methods

The clinical data of early pregnancy (6-13 wk) were collected, including height, weight, pre-pregnancy body mass index (BMI), family history, hemoglobin (Hb) level, fasting blood glucose, and other indicators. Two persons were responsible for data entry and verification.

GDM diagnostic criteria

Pregnant women at the gestational age of 24 to 28 wk, underwent an oral glucose tolerance test. Glucose water (75 g) was consumed after 8 h fasting on an empty stomach and then blood glucose was measured 3 times within 2 h. A diagnosis of GDM was made if the blood glucose level measured ≥ 5.1 mmol/L, 10.0 mmol/L, or 8.5 mmol/L during the fasting, 1-h, or 2-h tests, respectively[6].

Statistical analysis

Statistical software SPSS 21.0 was utilized for data analysis. The measurement data were represented as the mean and standard deviation, and group comparisons were conducted using the *t*-test. The enumeration data were represented as number (percentage), and the comparison between groups was conducted using the χ^2 test or Fisher's exact test. A multivariate logistic regression analysis was utilized, and statistical significance was determined at the $P < 0.05$ level. Based on the machine learning method, the nomogram prediction model was established by R language, and the random forest model was established using the Random Forest package. The model's application performance was assessed using sensitivity, specificity, and the area under the receiver operating characteristic curve (ROC AUC). The AUC was compared using the Delong test.

RESULTS

Sample characteristics

The comparison results of the general data in the two groups showed that there were significant differences in BMI, family history of diabetes, GDM history, macrosomia, hypertension, Hb level, triglyceride level, serum ferritin, and fasting blood glucose in the first trimester of pregnancy between the two groups ($P < 0.05$). These results are shown in Table 1.

Multivariate analysis of factors for GDM

Whether GDM occurred or not was used as the dependent variable, and the statistically significant variables in the univariate results were included in the multivariate logistic regression analysis as the independent variables, and the assignment criteria of each variable are shown in Table 2. The multivariate results showed that preconception BMI, family history of diabetes, GDM history, macrosomia, hypertension, Hb level, triglyceride level, serum ferritin, and fasting blood glucose in early pregnancy were the influencing factors of GDM as shown in Table 3 ($P < 0.05$).

Development of a first-trimester risk prediction model for GDM

Nomogram model construction: The results of multivariate logistic regression analysis were plotted into a nomogram model using R language and are shown in Figure 1. The total score was derived by assigning scores to each risk factor in the nomogram, and the corresponding probability of GDM occurring was determined using the total score and its associated probability value.

Random forest prediction model construction: Nine statistically significant indicators from univariate analysis were included in the random forest model, and the values are shown in Table 2. The results showed a fixed tree value, and when $mtry = 10$, the false positive rate of the model was the smallest. Based on $mtry = 10$, when $ntree = 500$, the model error was based on stability. Therefore, based on the $mtry = 10$ and $ntree = 500$ parameters, the top 5 variables in predicting the occurrence of GDM by the random forest model were serum ferritin, fasting blood glucose in the first trimester, BMI before pregnancy, Hb level and triglyceride level, as shown in Figure 2.

Comparison of the performance of the two predictive models: The nomogram model's ability to discriminate was assessed by the ROC AUC (Table 4 and Figure 3). The AUC of the random forest model was higher than that of the nomogram model ($Z = -6.104$, $P < 0.001$).

DISCUSSION

GDM is a condition that affects glucose metabolism during pregnancy. Typically, it occurs after the 27th week of gestation, although some women may develop preexisting diabetes prior to conception. The pathogenesis of GDM is complex, and its etiology is undefined[7]. In this study, after comparing the basic characteristics between pregnant women in the group with GDM and the group without GDM, the factors affecting the occurrence of GDM were obtained by multivariate logistic regression analysis, including preconception BMI, family history of diabetes, GDM history, macrosomia, hypertension, Hb and triglyceride levels, serum ferritin, and fasting blood glucose in the first trimester. These findings are essentially congruent with those of Li *et al*[8] and Tong *et al*[9].

This study revealed that pregnant women with a positive family history of diabetes exhibited a greater likelihood of GDM occurrence in comparison to their counterparts lacking such a familial history. Diabetes has a genetic predisposition, and can be passed on genetically to the next generation. Pregnant individuals who have a familial history of diabetes may possess a genetic predisposition that elevates the likelihood of the onset of GDM. A positive family history of diabetes mellitus has been established as one of the risk factors for GDM based on various national and international

Table 1 Comparison of general data between the two groups of patients, *n* (%)

Items	GDM group (<i>n</i> = 197)	Non-GDM group (<i>n</i> = 209)	Statistics	<i>P</i> value
Age (yr)	31.57 ± 3.94	30.794 ± 4.05	1.965	0.051
Body height (cm)	159.67 ± 5.91	160.38 ± 4.95	1.351	0.189
Occupation			2.246	0.134
Regular work	49 (24.87)	66 (31.58)		
No regular work	148 (75.13)	143 (68.42)		
Education level			4.790	0.091
Junior high school and below	22 (11.17)	18 (8.61)		
High school or technical secondary school	19 (9.64)	35 (16.75)		
College undergraduate and above	156 (79.19)	156 (74.64)		
Marital status			2.884	0.092
Primary marriage	169 (85.79)	166 (79.43)		
Remarriage	28 (14.21)	43 (20.57)		
Monthly income (yuan/month)			0.878	0.349
< 3000	36 (18.27)	46 (22.01)		
≥ 3000	161 (81.73)	163 (77.99)		
Family history of DM			22.357	< 0.001
Yes	49 (24.87) ^b	16 (7.66)		
No	148 (75.13) ^b	193 (92.34)		
Family history of hypertension			0.105	0.746
Yes	38 (19.29)	43 (20.57)		
No	159 (80.71)	166 (79.43)		
GDM			28.400	< 0.001
Yes	50 (25.38) ^b	13 (6.22)		
No	147 (74.62) ^b	196 (93.78)		
Parity			1.049	0.306
Plurality	91 (46.19)	86 (41.15)		
Primiparity	106 (53.81)	123 (58.85)		
PCOS			0.398	0.528
Yes	9 (4.57)	7 (3.35)		
No	188 (95.43)	202 (96.65)		
Giant infants			19.015	< 0.001
Yes	35 (17.77) ^b	9 (4.31)		
No	162 (82.23) ^b	200 (95.69)		
Pre-pregnancy weight (kg)	57.11 ± 9.58 ^a	54.68 ± 7.26	2.898	0.004
Pre-pregnancy BMI (kg/m ²)	24.11 ± 4.08 ^b	21.25 ± 2.63	8.439	< 0.001
History of abortion			1.106	0.293
Yes	70 (35.53)	64 (30.62)		
No	127 (64.47)	145 (69.38)		
Cesarean section history			0.878	0.349
Yes	35 (17.77)	30 (15.23)		
No	162 (82.23)	179 (84.77)		

History of preterm birth			1.872	0.171
Yes	51 (25.89)	67 (32.06)		
No	146 (74.11)	142 (67.94)		
History of stillbirth			0.212	0.645
Yes	4 (2.03)	3 (1.44)		
No	193 (97.97)	206 (98.56)		
Hypertension			14.325	< 0.001
Yes	35 (17.77) ^b	12 (5.74)		
No	162 (82.23) ^b	197 (94.26)		
Erythrocytes ($\times 10^{12}/L$)	4.31 \pm 0.41	4.27 \pm 0.39	0.994	0.321
Hb (g/L)	128.77 \pm 9.03 ^b	121.34 \pm 10.34	7.687	< 0.001
Urea (mmol/L)	2.94 \pm 0.66	2.87 \pm 0.65	1.079	0.281
Creatinine (μ mol/L)	44.71 \pm 6.81	45.38 \pm 5.82	-1.070	0.285
Uric acid (μ mol/L)	244.18 \pm 54.53	234.62 \pm 44.62	1.938	0.053
Triglyceride (mmol/L)	1.69 \pm 0.56 ^b	1.43 \pm 0.69	4.058	< 0.001
Serum ferritin (ng/mL)	73.96 \pm 18.36 ^b	53.29 \pm 15.30	12.350	< 0.001
First-trimester fasting hyperglycemia (mmol/L)	5.06 \pm 0.47 ^b	4.40 \pm 0.71	10.971	< 0.001

^a $P < 0.05$ vs non-gestational diabetes mellitus (GDM) group.

^b $P < 0.001$ vs non-GDM group. BMI: Body mass index; DM: Diabetes mellitus; GDM: Gestational diabetes mellitus; Hb: Hemoglobin; PCOS: Polycystic ovary syndrome.

Table 2 Variable assignment

Variables	Assignment
Whether GDM occurred (dependent variable)	Yes = 1, no = 0
GDM	Yes = 1, no = 0
Hypertension	Yes = 1, no = 0
Giant infant	Yes = 1, no = 0
Family history of DM	Yes = 1, no = 0
Hb	Original value input
Pre-pregnancy BMI	Original value input
Triglyceride	Original value input
Serum ferritin	Original value input
First-trimester fasting hyperglycemia	Original value input

BMI: Body mass index; DM: Diabetes mellitus; GDM: Gestational diabetes mellitus; Hb: Hemoglobin.

studies[10-12]. If a pregnant woman is diagnosed with GDM in a previous pregnancy, she is also more likely to have GDM in subsequent pregnancies, as confirmed by studies[13]. Therefore, for pregnant women with a familial predisposition to diabetes and GDM, it is recommended that doctors pay close attention to their health during pregnancy.

This study found that hypertension plays an essential role in the progress of GDM, and studies have confirmed that hypertension is one of the factors that pose an independent risk for GDM[14]. Hypertension may lead to the onset and progression of GDM by affecting placental blood flow and insulin sensitivity, causing islet cytopenia and dysfunction. In addition, this study also found that excess preconception BMI is one of the factors that pose an independent risk for GDM. This is because overweight and obesity affect insulin metabolism and production, increasing the body's need for insulin, and thus increasing the risk of GDM[15]. Therefore, weight control before pregnancy and maintaining a normal BMI can reduce the risk of GDM. For patients with hypertension during pregnancy, surveillance and intervention should be strengthened to reduce the risk of GDM.

Table 3 Multivariate logistic regression analysis results of gestational diabetes mellitus

Items	β	SE	Wald	P value ^a	OR (95%CI)
Family history of DM	1.340	0.472	8.061	0.005	3.818 (1.514-9.626)
Hypertension	1.674	0.643	6.772	0.009	5.335 (1.512-18.825)
GDM	1.201	0.519	5.353	0.021	3.323 (1.201-9.192)
Giant infant	2.269	0.647	12.312	< 0.001	2.653 (1.284-5.483)
Pre-pregnancy BMI (kg/m ²)	0.233	0.055	18.059	< 0.001	1.263 (1.134-1.406)
Hb (g/L)	0.071	0.018	14.867	< 0.001	1.073 (1.035-1.112)
Triglyceride (mmol/L)	0.792	0.249	10114	0.001	2.207 (1.355-3.596)
Serum ferritin (ng/mL)	0.070	0.010	44.508	< 0.001	1.072 (1.050-1.094)
First-trimester fasting hyperglycemia (mmol/L)	1.887	0.350	29.110	< 0.001	6.602 (3.326-13.105)

^aP < 0.05 was considered significant. BMI: Body mass index; DM: Diabetes mellitus; GDM: Gestational diabetes mellitus; Hb: Hemoglobin; SE: Standard error; OR: Odd ratio; CI: Confidence interval.

Table 4 Prediction performance evaluation results of the nomogram model and random forest model (%)

Prediction model		Sensitivity	Specificity	AUC (95%CI)
Nomogram model	Training set	74.1	87.6	0.883 (0.846-0.921)
	Validation set	81.0	81.2	0.850 (0.782-0.918)
Random forest model	Training set	91.4	84.8	0.950 (0.927-0.973)
	Validation set	89.7	89.7	0.918 (0.868-0.967)

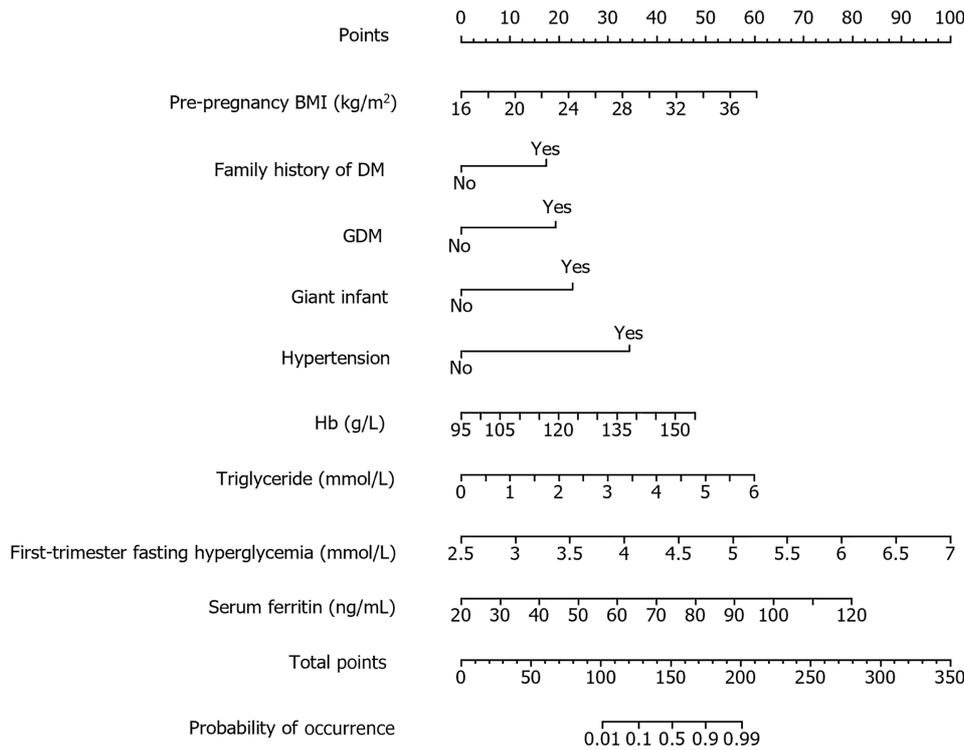
AUC: Area under the curve; CI: Confidence interval.

Sissala *et al*[16] found that Hb level is a risk factor for GDM, this finding is in alignment with the outcomes of the present investigation. The reason for this is that the level of Hb may affect the diastolic blood pressure of pregnant women, thereby increasing maternal peripheral vascular resistance. This condition may reduce the stiffness of the large arteries and lead to the formation of insulin resistance, thereby increasing the risk of GDM[17,18]. Serum ferritin is a major form of intracellular iron storage, and the body’s iron stores are positively correlated with Hb levels. Research has indicated that pregnant women diagnosed with GDM exhibit elevated serum ferritin levels in comparison to their non-GDM counterparts; therefore, regular measurement of Hb levels and serum ferritin levels during pregnancy can help pregnant women detect problems in a timely manner and take corresponding treatment measures. Studies have demonstrated that lipid and lipoprotein abnormalities, including elevated triglycerides, are associated with insulin resistance and type 2 diabetes, hence leading to significantly higher levels of triglycerides in GDM compared to non-GDM patients[19,20]. Therefore, monitoring blood lipid levels during pregnancy is of great clinical significance to effectively predict the onset of GDM.

In this investigation, the nomogram model and random forest model were established by applying preconception BMI, family history of diabetes, GDM history, macrosomia, hypertension, Hb and triglyceride levels, serum ferritin, and fasting blood glucose levels in the first trimester, and compared the prediction effect of the model. It was found that the AUC of GDM exhibited a value of 0.950 (95% confidence interval: 0.927-0.973), with a sensitivity rate of 91.4% and specificity rate of 84.80%. Compared with the nomogram model, it had better calibration and prediction accuracy. The reason for this may be that compared with the logistic regression model, the random forest model is not easy to overfit, has more advantages in processing high-dimensional data, and does not require feature selection.

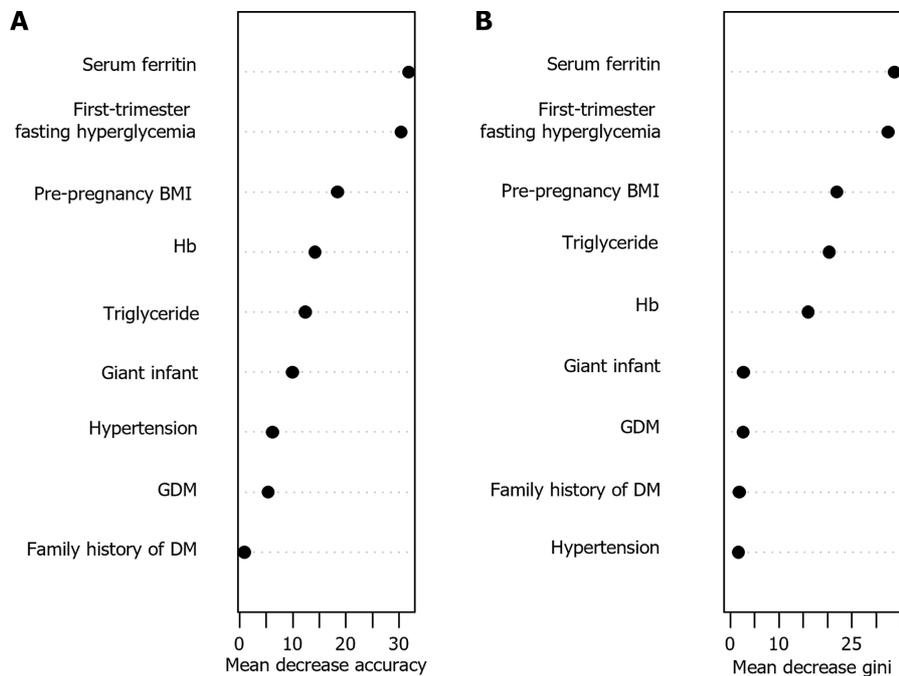
CONCLUSION

In summary, nine indicators, including preconception BMI, family history of diabetes, GDM history, macrosomia, hypertension, Hb and triglyceride levels, serum ferritin, and fasting blood glucose level in early pregnancy, effectively predicted the incidence of GDM. In this study, the predictive model for risk assessment of GDM based on the results of multivariate analysis had a better predictive effect, and the random forest model had higher efficiency in predicting the risk of GDM, which can effectively anticipate the likelihood of developing diabetes. In pregnant women, this has important guiding significance for the prevention and treatment of GDM. However, this study only collected data in one



DOI: 10.4239/wjd.v14.i10.1541 Copyright ©The Author(s) 2023.

Figure 1 Risk prediction nomogram model of gestational diabetes mellitus. BMI: Body mass index; DM: Diabetes mellitus; GDM: Gestational diabetes mellitus; Hb: Hemoglobin.



DOI: 10.4239/wjd.v14.i10.1541 Copyright ©The Author(s) 2023.

Figure 2 Variable importance analysis of random forest model. A: The diagram shows that the value of each variable was changed into a random number, and the random forest also measured the degree of reduction in accuracy; B: The importance of each variable was compared by calculating the heterogeneous influence of each variable on the observations on each node of the classification tree. The larger the value, the greater the importance of the variable. BMI: Body mass index; DM: Diabetes mellitus; GDM: Gestational diabetes mellitus; Hb: Hemoglobin.

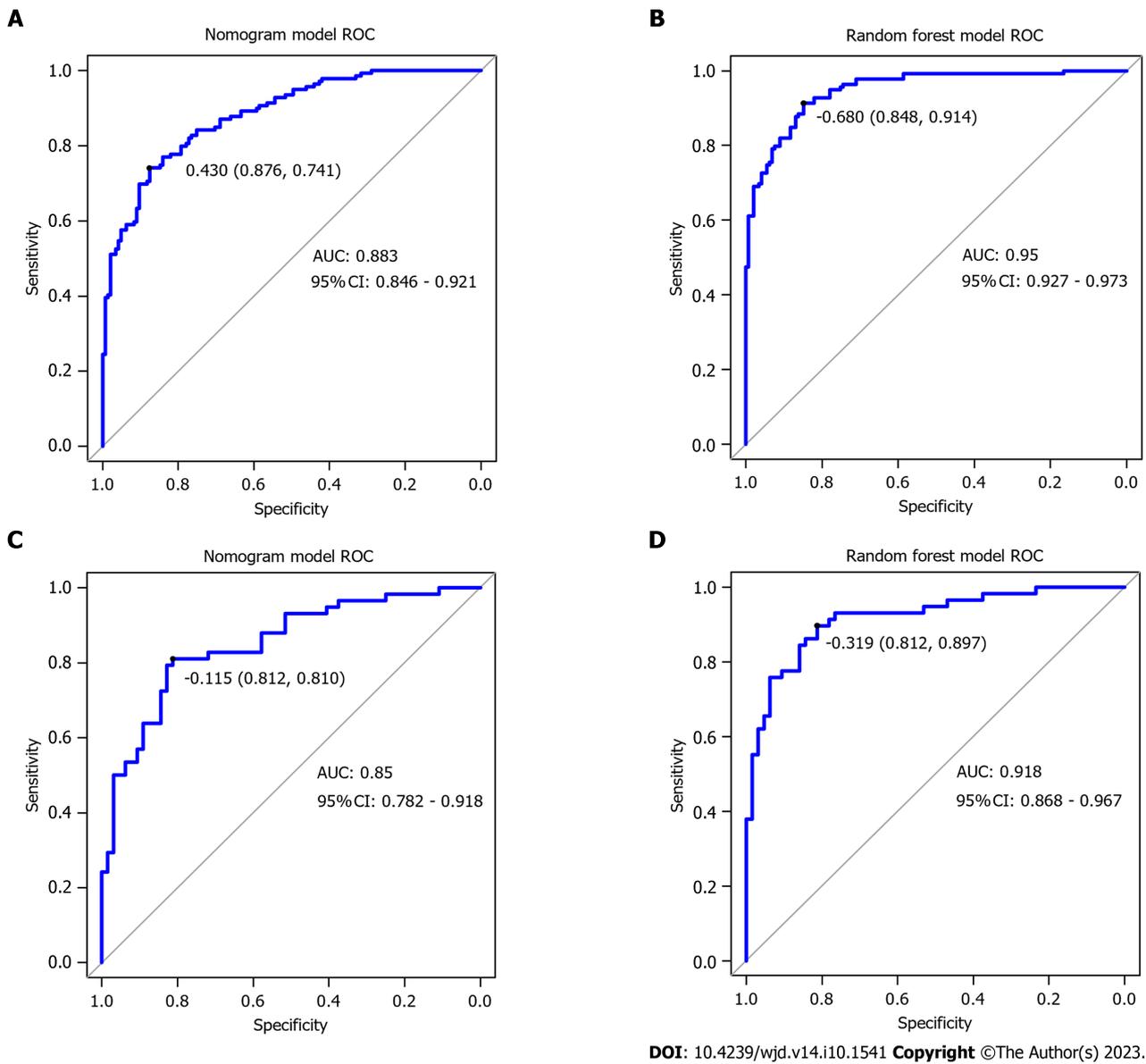


Figure 3 Receiver operating characteristic curve of the two models for predicting gestational diabetes mellitus. A: The figure shows the receiver operating characteristic curve (ROC) curve of the training set nomogram; B: The figure shows the ROC curve of the training set random forest model; C: The figure shows the ROC curve of the validation set line graph; D: The figure shows the ROC curve of the validation set random forest model. ROC: Receiver operating characteristic curve; AUC: Area under the curve.

hospital, and the sample size was small, which had certain limitations, and it is necessary to include a larger sample size for large-scale model verification in the future to provide a reference for clinical prediction of the incidence of GDM.

ARTICLE HIGHLIGHTS

Research background

Gestational diabetes mellitus (GDM) is a common metabolic disease during pregnancy, which has adverse effects on maternal and child health. The establishment and evaluation of risk prediction models can help to identify high-risk groups early and take corresponding intervention measures to reduce the risk in pregnant women and newborns. At present, research in this field mainly focuses on the screening of predictors and the construction of models and explores their reliability and practicability. These studies provide a theoretical basis and method support for the prevention and management of gestational diabetes.

Research motivation

The purpose of this study is to establish a reliable risk prediction model for gestational diabetes to help doctors detect and treat patients with GDM. The key issues to be solved in this study include determining the best predictors and

establishing effective models. Solving these problems is of great significance for improving the diagnostic rate of early diabetes and reducing the risk of complications in pregnant women and fetuses. It will also have a positive effect on future research in this field.

Research objectives

The main objective of this study is to establish a reliable risk prediction model for GDM. The achieved goals include obtaining the risk factors of GDM, establishing a risk factor prediction model, and evaluating the model. The random forest model has a good prediction effect, which can effectively predict the risk of diabetes in pregnant women and indicate the direction for future research in this field.

Research methods

In this study, a retrospective case analysis method was adopted, and the study subjects were stratified into two groups: Those with GDM and those without GDM. According to whether GDM occurred, the general data of the two groups of pregnant women were investigated and analyzed, and we established a risk prediction model for GDM during the trimester using both the logistic regression and random forest models, and the two models were evaluated and validated. The peculiarity and novelty of the research methods lie in the adoption of machine learning methods, which greatly improve the accuracy and reliability of the model.

Research results

This study successfully established a risk prediction model for early gestational diabetes in pregnant women (random forest and nomogram model). After analyzing and screening a number of clinical factors, the random forest model had high prediction accuracy and judgment ability. This study provides strong support for early prevention and intervention of gestational diabetes in pregnant women and provides a reference value for further research in this field. In the future, it is necessary to further expand the sample size, improve the considered factors and verify the stability and applicability of the model.

Research conclusions

This study proposed a model for predicting the likelihood of developing gestational diabetes during the early stages of pregnancy and compared the predictive effects of the random forest and nomogram models. The results suggested that the random forest model can more accurately predict the risk of gestational diabetes during early pregnancy.

Research perspectives

Future research should focus on improving the risk prediction model of gestational diabetes in pregnant women and improve the accuracy and stability of the model to meet clinical needs. We should also explore new predictors, explore pathological mechanisms, and identify intervention strategies to reduce the risk of diabetes and its complications in pregnant women and improve maternal health.

FOOTNOTES

Author contributions: Lin Q designed and performed the research and wrote the paper; Fang ZJ designed the research and supervised the report.

Institutional review board statement: This study was reviewed and approved by the Ethics Committee of the Fujian Maternity and Child Health Hospital.

Informed consent statement: As the study used anonymous and pre-existing data, the requirement for the informed consent from patients was waived.

Conflict-of-interest statement: We have no financial relationships to disclose.

Data sharing statement: No additional data are available.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>

Country/Territory of origin: China

ORCID number: Qing Lin 0009-0003-4682-7249; Zhuan-Ji Fang 0000-0001-6637-4556.

S-Editor: Qu XL

L-Editor: A

P-Editor: Chen YX

REFERENCES

- 1 **Ge L**, Huang P, Miao H, Yu H, Wu D, Chen F, Lin Y, Li W, Hua J. The new landscape of differentially expression proteins in placenta tissues of gestational diabetes based on iTRAQ proteomics. *Placenta* 2023; **131**: 36-48 [PMID: 36473392 DOI: 10.1016/j.placenta.2022.11.012]
- 2 **Prentice PM**, Olga L, Petry CJ, Simmons D, Murphy HR, Hughes IA, Acerini CL, Ong KK, Dunger DB. Reduced size at birth and persisting reductions in adiposity in recent, compared with earlier, cohorts of infants born to mothers with gestational diabetes mellitus. *Diabetologia* 2019; **62**: 1977-1987 [PMID: 31396660 DOI: 10.1007/s00125-019-4970-6]
- 3 **Sun H**, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, Stein C, Basit A, Chan JCN, Mbanya JC, Pavkov ME, Ramachandaran A, Wild SH, James S, Herman WH, Zhang P, Bommer C, Kuo S, Boyko EJ, Magliano DJ. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res Clin Pract* 2022; **183**: 109119 [PMID: 34879977 DOI: 10.1016/j.diabres.2021.109119]
- 4 **Lamain-de Ruiter M**, Kwee A, Naaktgeboren CA, de Groot I, Evers IM, Groenendaal F, Hering YR, Huisjes AJ, Kirpestein C, Moninx WM, Siljee JE, Van 't Zelfde A, van Oirschot CM, Vankan-Buitelaar SA, Vonk MA, Wiegers TA, Zwart JJ, Franx A, Moons KG, Koster MP. External validation of prognostic models to predict risk of gestational diabetes mellitus in one Dutch cohort: prospective multicentre cohort study. *BMJ* 2016; **354**: i4338 [PMID: 27576867 DOI: 10.1136/bmj.i4338]
- 5 **Sweeting AN**, Wong J, Appelblom H, Ross GP, Kouru H, Williams PF, Sairanen M, Hyett JA. A Novel Early Pregnancy Risk Prediction Model for Gestational Diabetes Mellitus. *Fetal Diagn Ther* 2019; **45**: 76-84 [PMID: 29898442 DOI: 10.1159/000486853]
- 6 **Luo J**, Geng X, Zhou J, Liang S, Zheng W, Li G. Characteristics of the oral glucose tolerance test in women with different pre-pregnancy body mass index and the effect of gestational diabetes mellitus on twin pregnancy outcomes. *Clinics (Sao Paulo)* 2023; **78**: 100272 [PMID: 37604047 DOI: 10.1016/j.clinsp.2023.100272]
- 7 **Li C**, Li N, Liu C, Li H. Causal effect of early life adiposity on gestational diabetes mellitus and mediating roles of lipidomic biomarkers. *Front Nutr* 2023; **10**: 1225376 [PMID: 37538923 DOI: 10.3389/fnut.2023.1225376]
- 8 **Li R**, Yuan K, Yu X, Jiang Y, Liu P, Zhang K. Construction and validation of risk prediction model for gestational diabetes based on a nomogram. *Am J Transl Res* 2023; **15**: 1223-1230 [PMID: 36915791]
- 9 **Tong JN**, Wu LL, Chen YX, Guan XN, Tian FY, Zhang HF, Liu K, Yin AQ, Wu XX, Prof JN. Fasting plasma glucose in the first trimester is related to gestational diabetes mellitus and adverse pregnancy outcomes. *Endocrine* 2022; **75**: 70-81 [PMID: 34342804 DOI: 10.1007/s12020-021-02831-w]
- 10 **Wang Y**, Luo B. [Risk factors analysis of gestational diabetes mellitus based on International Association of Diabetes Pregnancy Study Groups criteria]. *Nan Fang Yi Ke Da Xue Xue Bao* 2019; **39**: 572-578 [PMID: 31140422 DOI: 10.12122/j.issn.1673-4254.2019.05.12]
- 11 **Sapanont K**, Sunsaneevithayakul P, Boriboonhirunsarn D. Relationship between ABO blood group and gestational diabetes mellitus. *J Matern Fetal Neonatal Med* 2021; **34**: 1255-1259 [PMID: 31204532 DOI: 10.1080/14767058.2019.1633299]
- 12 **Lewandowska M**. Gestational Diabetes Mellitus (GDM) Risk for Declared Family History of Diabetes, in Combination with BMI Categories. *Int J Environ Res Public Health* 2021; **18** [PMID: 34203509 DOI: 10.3390/ijerph18136936]
- 13 **Abualhamael S**, Mosli H, Baig M, Noor AM, Alshehri FM. Prevalence and Associated Risk Factors of Gestational Diabetes Mellitus at a University Hospital in Saudi Arabia. *Pak J Med Sci* 2019; **35**: 325-329 [PMID: 31086509 DOI: 10.12669/pjms.35.2.498]
- 14 **Sun P**, Liu K, Cui X, Zhang L, Cao T. Establishment of a nomogram model to predict the risk of macrosomia in patients with gestational diabetes mellitus. *J Matern Fetal Neonatal Med* 2023; **36**: 2232072 [PMID: 37408128 DOI: 10.1080/14767058.2023.2232072]
- 15 **Benchahong S**, Sunsaneevithayakul P, Boriboonhirunsarn D. The Association Between Body Fat Index and Gestational Diabetes Mellitus: A Prospective Cohort Study. *Cureus* 2023; **15**: e39615 [PMID: 37388597 DOI: 10.7759/cureus.39615]
- 16 **Sissala N**, Mustaniemi S, Kajantie E, Väärasmäki M, Koivunen P. Higher hemoglobin levels are an independent risk factor for gestational diabetes. *Sci Rep* 2022; **12**: 1686 [PMID: 35102239 DOI: 10.1038/s41598-022-05801-y]
- 17 **Yong HY**, Mohd Shariff Z, Mohd Yusof BN, Rejali Z, Tee YYS, Bindels J, van der Beek EM. Early pregnancy hemoglobin is associated with the risk of gestational diabetes mellitus: a retrospective cohort study. *Br J Nutr* 2022; **128**: 2097-2104 [PMID: 35139935 DOI: 10.1017/S000711452100502X]
- 18 **Gao CJ**, Huang XM, Chen ZP, Sheng L, Xu J, Li Y, Li XY, Zhang R, Yu ZY, Zha BB, Wu YY, Yang M, Ding HY, Sun TG, Zhang YQ, Ma L, Liu J. [High level of hemoglobin during the first trimester of pregnancy associated with the risk of gestational diabetes mellitus]. *Zhonghua Fu Chan Ke Za Zhi* 2019; **54**: 654-659 [PMID: 31648440 DOI: 10.3760/cma.j.issn.0529-567x.2019.10.002]
- 19 **Shen L**, Wang D, Huang Y, Ye L, Zhu C, Zhang S, Cai S, Wang Z, Chen H. Longitudinal trends in lipid profiles during pregnancy: Association with gestational diabetes mellitus and longitudinal trends in insulin indices. *Front Endocrinol (Lausanne)* 2022; **13**: 1080633 [PMID: 36714591 DOI: 10.3389/fendo.2022.1080633]
- 20 **Song T**, Su G, Chi Y, Wu T, Xu Y, Chen C. Triglyceride-glucose index predicts the risk of gestational diabetes mellitus: a systematic review and meta-analysis. *Gynecol Endocrinol* 2022; **38**: 10-15 [PMID: 34184968 DOI: 10.1080/09513590.2021.1940932]



Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: <https://www.f6publishing.com/helpdesk>
<https://www.wjgnet.com>

