

# Screening for gestational diabetes mellitus in women with low BMI in an urban, Chinese, population: two retrospective cohorts study

Fei Guo

Shanghai Jiao Tong University

Shuai Yang

Shanghai Jiao Tong University

Yong Zhang

Shanghai Jiao Tong University

Xi Yang

Shanghai Jiao Tong University

Chen Zhang

Shanghai Jiao Tong University

Jianxia Fan (✉ [fanjianxia122@126.com](mailto:fanjianxia122@126.com))

Shanghai Jiao Tong University <https://orcid.org/0000-0003-3338-877X>

---

## Research article

**Keywords:** BMI, screening model, pregnancy, GDM

**Posted Date:** July 28th, 2019

**DOI:** <https://doi.org/10.21203/rs.2.12011/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

## Abstract

Background We design this study to assess the utility of baseline characteristics in first trimester in the prediction of gestational diabetes mellitus (GDM) in non-obese, women in an urban, Chinese, antenatal population. Methods Multiple, logistic regression analysis was used to develop a predictive model for GDM, in a retrospective cohort of 3956 women (mean age 30.61, years, mean pre-pregnant BMI= 21.45 kg/m<sup>2</sup>) during 2015 in Shanghai. We refined the predictive model with t-distributed stochastic neighbor embedding (t-SNE) to attempt to distinguish GDM from non-GDM. Results GDM prevalence in the retrospective cohort was 16.7%. Advanced age, prepregnant obesity, high first-trimester, fasting, plasma glucose, and, a family history of diabetes were positively-related to the development of GDM. The simplified point scoring system yielded an AUC of 0.69. When the model was applied to a prospective cohort of 6572 women recruited in 2016 (mean age 30.80 years, mean prepregnant BMI = 21.18 kg/m<sup>2</sup>, incidence of GDM 10.8%), the AUC was 0.70. A score of ≥3 points was 91.3% specific and 25.4% sensitive for GDM. In both cohorts, the distinct regions of GDM and non-GDM were ideally separated in the t-SNE, generating transitional boundaries in the image found by different color. Conclusions Our low BMI, antenatal population now over-consumes a Western, diabetogenic diet during pregnancy though we attribute the declining incidence (2015, 16.7% vs 2016, 10.8%) to an improved antenatal education program. We propose a score system to screen for GDM in first trimester, and also, we should concentrate our efforts on prevention through antenatal education.

## Background

Gestational diabetes mellitus (GDM) was perceived as abnormal glucose intolerance that is primarily detected during pregnancy. Rapid societal transition from traditional famine lifestyle famine to an obesogenic environment puts Chinese people at high prevalence of GDM, especially in urban area, ranging from 8.1 to 19.7% <sup>1-3</sup>. GDM results from varying degrees of insulin resistance to placenta-derived hormones, and, results in incremental deposition of maternal adipose tissue <sup>4</sup>. Typically, GDM is diagnosed with an abnormal OGTT at 24-28<sup>th</sup> week gestation. Recently, it is controversial that fasting glucose greater than 5.1 mmol/l in the first prenatal visit is appropriate for GDM diagnosis <sup>5</sup>. Clearly, identifying women at risk of GDM as early as possible in pregnancy enables interventions that reduce the risk of adverse fetal outcomes including those associated with fetal macrosomia, and, later maternal consequences including the metabolic syndrome and cardiovascular morbidity <sup>6, 7</sup>. Many groups have tried to develop predictive models based on risk factors identified in the first trimester, to identify an abnormal OGTT at 24-28 weeks. Scoring systems, biochemical assays of glucose, glycosylated haemoglobin levels, etc have all been used in different populations with varying degrees of success <sup>8-11</sup>. Established risk factors for GDM include advanced maternal age, excessive weight gain during pregnancy, overweight or obesity, diabetes in first degree relatives, giving birth to an infant with a

macrosomia, etc<sup>12</sup>. Clinical trials have shown that GDM can be prevented to some extent by intensive lifestyle modification implemented before 20 weeks of gestation<sup>13, 14</sup>.

Different populations have different risk profiles for GDM. In Western countries GDM tends to occur in overweight women ( $\text{BMI} > 30 \text{ kg/m}^2$ ) or women with increasing gestational weight gain<sup>15, 16</sup> whereas in some developing countries it affects populations of low BMI women to varying degrees. The reasons for GDM in these circumstances are not entirely clear. In the present study, our aim is to create a simple, implementable strategy to identify GDM in low BMI women in a Chinese urban population.

## Materials And Methods

### ***Study design***

A retrospective cohort study was conducted from January 2015 to December 2015 (N=3956). Eligible subjects were recruited who were undergoing their first antenatal visit at the International Peace Maternity and Child Care Health Hospital (IPMCH). The validation cohort was recruited in 2016 (N=6572). Demographic data of the subjects were obtained by a face-to-face interview questionnaire at the first prenatal visit, which included last menstrual period, maternal age, gestational age, prepregnant weight, current height, personal history, educational levels, parity, diabetes in first degree relatives. BMI was calculated by dividing the weight (kg) by the height squared ( $\text{m}^2$ ). Eligible subjects gave blood samples for analysis of fasting plasma glucose and glycated hemoglobin (HbA1c) at the first visit to antenatal clinic.

GDM was diagnosed at 24-28 weeks of gestation according to the American Diabetes Association (ADA) criteria using abnormal plasma glucose values during the 2 hours, 75-gOGTT. Abnormal values were defined according to thresholds established by the ADA: a fasting level greater than 5.1 mmol/l, a 1-hour value greater than 10.0 mmol/l, and, a 2- hour value greater than 8.5 mmol/l.

### ***Exploratory cohort***

This cohort included 4205 women who were recruited at the IPMCH in the first trimester of pregnancy. Women with pre-existing diabetes ( $FPG \geq 7\text{mmol/L}$  and (or)  $HbA1c \geq 6.5\%$  checked in the first antenatal care) ( $n=34$ ) and multiple pregnancies ( $n=215$ ) were excluded, leaving 3956 pregnancies available for analysis.

### ***Validation cohort***

This cohort included 7022 women during 2016, recruited at the IPMCH during the first trimester pregnancy. Women with pre-existing diabetes ( $FPG \geq 7\text{mmol/L}$  and (or)  $HbA1c \geq 6.5\%$  checked in the first antenatal visit) ( $n=292$ ) and multiple pregnancies ( $n=158$ ) were excluded, leaving 6572 pregnancies available for analysis.

The key difference between the two cohorts was the incidence of GDM (16.7% v 10.8%). Many factors may contribute to variations in incidence though we believe that this was largely attributable to an antenatal education program to teach antenatal women to manage their diets and weight during pregnancy. In China it is traditional for women to increase their calorie consumption during pregnancy though this is clearly not helpful if they are consuming a diabetogenic “Western” diet.

### ***Risk scoring system***

Risk factors for GDM include advanced age, high pre-pregnancy BMI, diabetes in first degree relatives and high, FPG in the first trimester. In the retrospective cohort, multiple logistic regression analysis was used to estimate the coefficients of each risk factor and mutually-adjusted odds ratios (ORs) assigned for GDM. The continuous predictor variables age, BMI and FPG were found to be linear with log odds of the outcomes. To translate the regression coefficients into scores, we used the increasing risk of every 0.5mmol/L increase in FPG as 1 standard unit and gave it an incremental score of 1. Receiver operating characteristic (ROC) of continuous variables such as maternal age were identified to calculate the appropriate cut-off value for diagnosing GDM. The following risk factors were stratified: maternal age (<30 years and  $\geq 30$  years), diabetes in first degree relatives (no versus yes), the cut- off of FPG in the first trimester referred to FIGO guidelines ( $<5.1\text{ mmol/L}$  and  $\geq 5.1\text{ mmol/L}$ ), prepregnancy BMI was classified based on the Chinese criteria (normal,  $<24.0\text{ kg/m}^2$ ;

overweight,  $24.0 \text{ kg/m}^2 \leq \text{BMI} < 28.0 \text{ kg/m}^2$ ; obese,  $\text{BMI} \geq 28.0 \text{ kg/m}^2$ ). The lowest category of each variable was given a score of 0. The combination of all risk factors for a specific total score was computed to obtain GDM risk values. To quantify the accuracy of the prediction models to discriminate subjects with GDM from subjects without, a ROC curve was plotted, and the AUC was calculated. The best fit of the model was evaluated with the Hosmer and Lemeshow test statistics. Then, sensitivity and specificity were calculated at different risk thresholds. A scoring system was used to assign each woman to one of two risk subgroups in the prospective cohort. Spatially mapped t-distributed stochastic neighbor embedding (t-SNE) was used to distinguish non-GDM from GDM, that is, the ability to screen for a true negative group.

### ***Statistical analysis***

Analysis were performed using SPSS version 23.0 (SPSS, Inc., Chicago, IL) and R statistical software version 3.6.0 (packages rms and Rtsne). Continuous variables were presented as the mean value with standard deviation and categorical data was expressed as numbers and percentages. Levene's test was used to determine the homogeneity of the variances; Kolmogorov-Smirnov was used to assess the normal distribution. Summary statistics between both groups were compared using either unpaired Student's t-test or Mann-Whitney tests for continuous data, and chi-squared tests for categorical data. A p-value of  $<0.05$  was considered to indicate statistical significance.

## **Results**

### ***Baseline characteristics for two cohorts***

In the retrospective cohort of 3956 women during 2015, 662 developed GDM, giving an incidence of 16.7%, mean age was 30.61 years, mean prepregnancy BMI was  $21.45 \text{ kg/m}^2$ , mean FPG was  $4.45 \text{ mmol/L}$ , positive finding of diabetes in first degree relatives was 27.4%. In the validation cohort of 6572 women in 2016, 739 developed GDM, giving an incidence of 11.24%, mean age was 30.84 years, mean prepregnancy BMI was  $21.20 \text{ kg/m}^2$ , mean FPG was  $4.40 \text{ mmol/L}$ , positive finding of diabetes in first degree relatives was 16.5%. Comparison of baseline characteristics between GDM and non-GDM in two cohorts is shown in Table 1.

**Table 1.** General characteristics of women with normal glucose tolerance and those who developed GDM

Variables	Exploratory cohort(n=3956)			Validation cohort(n=6572)		
	GDM (n=662)	Non-GDM (n=3294)	P	GDM (n=739)	Non-GDM (n=5833)	P
Age, years	31.73±3.74	30.39±3.52	<0.001	32.24±3.95	30.66±3.69	<0.001
Pre-BMI, kg/m <sup>2</sup>	22.63±3.43	21.21±2.71	<0.001	22.45±3.37	21.04±2.69	<0.001
FH, n (%)			<0.001			<0.001
No	403(60.90)	2468(74.90)		530(71.70)	4960(85.00)	
Yes	259(39.10)	826(25.10)		209(28.30)	873(15.00)	
FPG, mmol/L	4.61±0.45	4.42±0.38	<0.001	4.59±0.52	4.38±0.38	<0.001

Data are n (%), mean±SD or median (interquartile range)

p values for differences between two groups were obtained by ANOVA or χ<sup>2</sup> test

GDM: gestational diabetes mellitus; pre-BMI: prepregnant body mass index; FH: family history; FPG: fasting plasma glucose

### ***Logistic regression results***

Table 2 showed the incidence of risk factors for GDM with OR and corresponding scores from a multiple logistic regression model in the retrospective cohort. Increasing age, prepregnancy BMI, FPG and a family history of diabetes were confirmed as increased risks for GDM. Table 3 depicted the incidence gradient for GDM according to cumulative risk score in the exploratory and validation cohorts.

The projected GDM risk was estimated with the equation:

$$P=1 / [1+\exp (-10.84+0.078 * \text{age}+0.119 * \text{BMI}+0.893 * \text{FPG}+0.491 * \text{positive family history of diabetes})]$$

**Table 2.** GDM risk factors according to disease status and Odds Ratio (OR) in the exploratory cohort

Predictors	Total (n)	Incidence of GDM n (%)	Adjusted OR (95%CI)	P	score
Age, years					
<30	1665	207 (12.40)	ref		0
≥30	2291	455 (19.50)	1.61 (1.34-1.94)	<0.001	1
Prepregnant BMI, kg/m <sup>2</sup>					
<24					
24-27.9	3361	483 (14.40)	ref		0
≥28	476	128 (26.90)	2.01 (1.59-2.55)	<0.001	2
	119	51 (42.90)	3.58 (2.38-536)	<0.001	3
Family history No					
Yes	2871	403(14.00)	ref		0
	1085	259 (23.90)	1.71(1.42-2.06)	<0.001	1
FPG, mmol/L					
<5.1	3707	560 (15.10)	ref		0
≥5.1	249	102 (41.00)	2.81 (2.10-3.76)	<0.001	2

GDM: gestational diabetes mellitus; pre-BMI: prepregnant body mass index; FPG: fasting plasma glucose

**Table 3.** Incidence gradient for GDM according to cumulative risk score in the exploratory and validation groups

Total score	Exploratory cohort		Validation cohort	
	No. in GDM	Incidence of GDM (%)	No. in GDM	Incidence of GDM (%)
0	94	8.9	101	5.0
1	225	13.7	295	9.9
2	130	20.4	121	17.0
3	94	27.6	107	19.3
4	69	38.5	66	30.0
5	23	48.9	21	42.9
6	20	52.6	16	51.6
7	7	77.8	9	75.0

GDM: gestational diabetes mellitus.

### ROC results

We depicted ROC curves with 3956 subjects from the retrospective cohort, showing a good discrimination for the GDM prediction model with AUC of 0.69 [95%CI 0.67-0.72] (Fig. 1a). Applying the retrospective set estimates to the prospective set yielded a similar of AUC 0.70 [95% CI 0.68-0.72] (Fig. 1b).

### ***Calibration results***

Figure 2 shows the *P*-value of the Hosmer and Lemeshow tests for comparing the mean probability estimates in each decile of probability in both cohorts. The observed disease incidence were 0.51 (Fig. 2a) and 0.98 (Fig. 2b), respectively.

### ***Performance of the scoring system***

Table 3 displays the cumulative risk score and related risk of developing GDM. The total score values range from 0 to 7. For example, a women aged 35 years (risk score =1), BMI 19 kg/m<sup>2</sup> (risk score =0), first degree relatives has diabetes (risk score =1), with FPG concentration of 5.9 mmol/L (risk score =2) would have a cumulative risk score of 4, with a projected GDM risk of 30.7%. When the cumulative score was set at 3, the specificity for screening GDM was 91.3% with a sensitivity of 25.4%. A nomogram is illustrated to show the probability of developing GDM (Fig. 3).

### ***Screening ability***

To assess whether the structure revealed by t-SNE could be linked to clinical outcome, and thereby discriminate subpopulations, we converted the t-SNE space to a\*b color space. In the resulting t-SNE image obtained by density-based analysis, each pixel was colored according to its property (orange for GDM and blue for non-GDM). In both cohorts (Fig. 4a; Fig. 4b), the distinct regions of GDM and non-GDM found by t-SNE were separated, generating transitional boundaries that could be highlighted by two different colors, indicating the ability to distinguish GDM from non-GDM women.

## Discussion

Our aim was to create a simple and easily accessible strategy based on clinical history to identify GDM in low BMI (ca.21 kg/m<sup>2</sup>) women in a Chinese urban population. A 7-point risk-scoring system was developed and validated with these clinical factors for GDM. Although our study demonstrates a reasonable discrimination and calibration, its sensitivity as a first trimester screening test is not high enough. The reasons for the low sensitivity may lie in recent dietary changes in Shanghai where the antenatal population now consumes a high sugar, diabetogenic, Western diet combined with a traditional propensity to increase calorie consumption during pregnancy. The effects of this combination of risks are not detectable in early pregnancy though may become apparent as pregnancy advances. It is clearly important that our antenatal population avoids the risk of over-eating a diabetogenic diet, and, we believe that changes to our antenatal education program in 2015-16 may have contributed to the sharp decline in the incidence of GDM (2015, 16.7% vs 2016, 10.8%).

Various first-trimester prediction models for GDM in Caucasian populations have been proposed. A recent study in Sydney (n=985)<sup>8</sup> combined previous GDM, family history of diabetes, age, ethnicity, BMI before pregnancy and parity together as predictors, and the model had an AUC value of 0.88 with a detection rate of 70.2% at a 10% fixed false positive rate (FPR). Similarly, a prospective cohort study (n=995) in Netherlands<sup>17</sup> reported a risk prediction model for GDM based on ethnicity, family history of DM, prepregnancy BMI, history of GDM, yielding an AUC of 0.77 and 75% identified GDM. Further, Caliskan et al.<sup>18</sup> developed a scoring system to differentiate between women at low and high risk for GDM. They identified the following risk factors for GDM from a retrospective case-control study in a Turkish population: maternal age  $\geq$ 25 years, prepregnant BMI  $\geq$ 25 kg/m<sup>2</sup>, first-degree relative with diabetes mellitus, previous macrosomic fetus and a previous adverse pregnancy outcome, and the result was accurate with diagnosing 85% of cases. In Canada, Naylor et al<sup>19</sup> developed a clinical scoring system based on age, BMI and ethnicity has the ability to detect 81.2 to 82.6 percent of the women with GDM. It was a clinical prediction model of GDM tailored to nulliparous women within specific racial/ethnic groups<sup>17</sup>. The prediction model included five risk factors: race, age at delivery, prepregnancy BMI, family history of diabetes, and pre-existing hypertension. Asian and Filipino women had increased risk of GDM at lower BMI's compared with Caucasian and African American women. The biological mechanisms behind this

observation remain unclear. The AUC for different races had moderate predictive performance. When using risk stratification strategy, 23.1% Asian were correctly classified and Hispanic and Hawaiian/Pacific Islander were 70.3% and 46.9%, respectively.

Our study tested the model both internally and externally to assess its predictive performance. Including both first and second degree relatives with diabetes in the retrospective cohort may have over-estimated the contribution of family history with some measurement bias. So, in the prospective cohort we only considered the first degree relatives with diabetes. There are some limitations to be addressed. Firstly, our study was based on the hospital cohort, and the homogeneity of the cohort increased the internal validity but weakened the confidence of the conclusion. Secondly, we did not incorporate excessive gestational weight gain (GWG) though it was a key factor for developing GDM. Finally, the sample size of the retrospective cohort was smaller than that in the prospective cohort, because some women returned to their home towns to deliver, and, others refused to participate with the questionnaire or OGTT test in 2015. The final sample size (n=3956) was sufficient for establishing the prediction model, suggesting that these attritional features did not result in a type I error.

In conclusion, we hypothesize that the recent change to a Western, diabetogenic diet together with a traditional propensity to over-consume during pregnancy provides an explanation for the high incidence of GDM in our low BMI, population, and, why our validated, first trimester model has such low sensitivity for GDM in the first trimester of pregnancy. Low sensitivity of a prediction model will cause high false negative rates, which will delay the patient's visit, affect the course of disease and prognosis. Similarly, low specificity of a prediction model will result in high false positive rate, which will waste medical resources and increase patients' anxiety. Our simple model with high specificity can be used as a screening method to identify those who won't develop GDM. Continued monitoring after the first trimester with timely intervention are necessary for those women who are at continuing risk of GDM. Our results suggest that first trimester modeling of GDM risk may not be a practical proposition because the condition largely develops after the first trimester of pregnancy, and, our efforts may be better directed at antenatal, dietary education in the first trimester to attempt to prevent GDM in the first instance.

# Declarations

## Ethics approval and consent to participate

The study was approved by the Ethics Committee of the International Peace Maternity and Child Health Hospital, School of Medicine, Shanghai Jiaotong University. The data analysis procedures conformed to the guidelines in the Declaration of Helsinki. All participants provided written informed consent.

## Consent for publication

Not applicable.

## Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

## Competing interests

The authors declare that they have no conflict of interest.

## Funding

This work was supported by the Shanghai Municipal Commissions of Health and Family Planning Program [grant number 15GWZK0701].

## Author contributions

All the authors contributed to the work and approved the final version of this manuscript. This study was designed by Jianxia Fan, Fei Guo and Shuai Yang. Chen Zhang and Fei Guo performed the statistical analysis and wrote the manuscript. Yong Zhang and Xi Yang compiled data. Jianxia Fan reviewed and edited manuscript.

## Acknowledgments

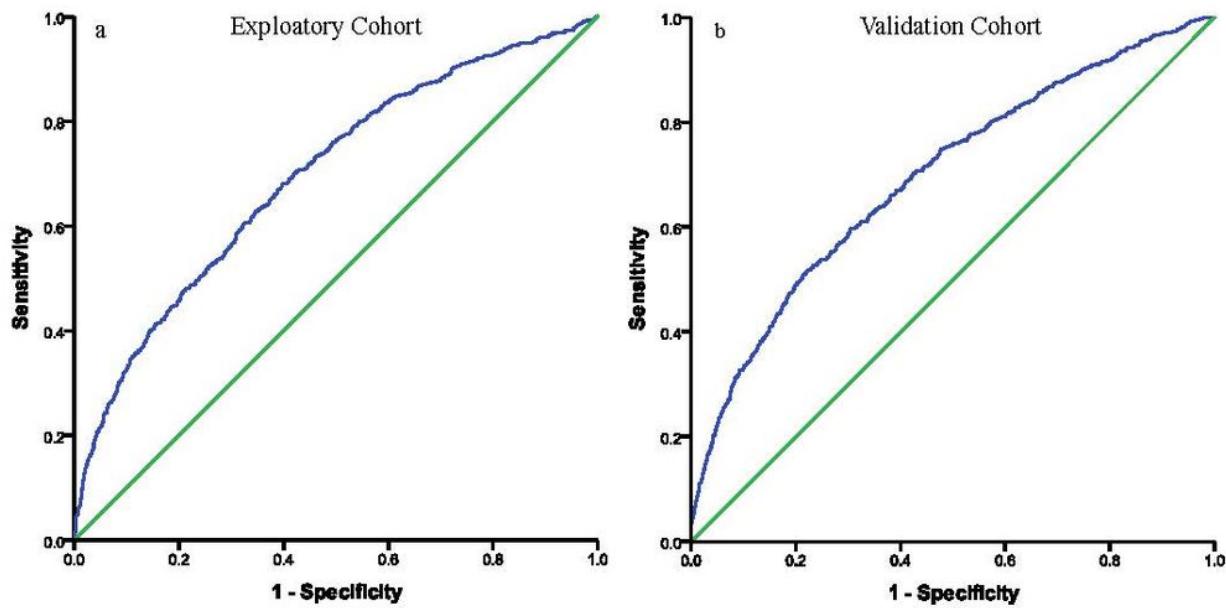
The authors thank the staff at the International Peace Maternity and Child Health Hospital for their technic supports. The contribution of the children and parents is highly appreciated.

## References

1. Zhu WW, Yang HX, Wang C, et al.:**High Prevalence of Gestational Diabetes Mellitus in Beijing: Effect of Maternal Birth Weight and Other Risk Factors.** Chin Med J 2017, 130:1019-25.
2. Chan JC, Zhang Y, Ning G:**Diabetes in China: a societal solution for a personal challenge.** Lancet Diabetes Endocrinol 2014, 2:969-79.
3. Leng J, Shao P, Zhang C, et al.:**Prevalence of gestational diabetes mellitus and its risk factors in Chinese pregnant women: a prospective population-based study in Tianjin, China.** PLoS One 2015, 10:e0121029.
4. McCabe CF, Perng W:**Metabolomics of Diabetes in Pregnancy.** Curr Diabetes Rep 2017, 17:12.
5. Cosson E, Carbillon L, Valensi P:**High Fasting Plasma Glucose during Early Pregnancy: A Review about Early Gestational Diabetes Mellitus.** J Diabetes Res 2017.
6. Burlina S, Dalfra MG, Chilelli NC, Lapolla A:**Gestational Diabetes Mellitus and Future Cardiovascular Risk: An Update.** Int J Endocrinol 2016, 2016:2070926.
7. Mohammadbeigi A, Farhadifar F, Soufi Zadeh N, et al.:**Fetal macrosomia: risk factors, maternal, and perinatal outcome.** Ann Med Health Sci Res 2013, 3:546-50.
8. Sweeting AN, Appelblom H, Ross GP, et al.:**First trimester prediction of gestational diabetes mellitus: A clinical model based on maternal demographic parameters.** Diabetes Res Clin Pract 2017, 127:44-50.
9. Theriault S, Giguere Y, Masse J, Girouard J, Forest J-C:**Early prediction of gestational diabetes: a practical model combining clinical and biochemical markers.** Clin Chem Lab Med 2016, 54:509-18.
10. Eleftheriades M, Papastefanou I, Lambrinoudaki I, et al.:**Elevated placental growth factor concentrations at 11-14 weeks of gestation to predict gestational diabetes mellitus.** Metabolism 2014, 63:1419-25.
11. van Leeuwen M, Opmeer BC, Zweers EJK, et al.:**External validation of a clinical scoring system for the risk of gestational diabetes mellitus.** Diabetes Research and Clinical Practice 2009, 85:96-101.

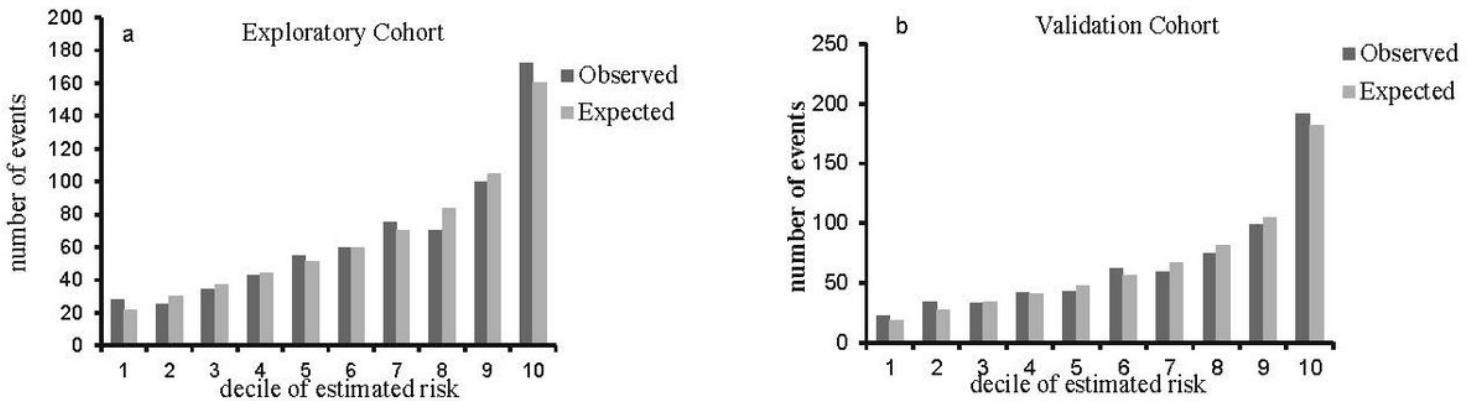
12. Kjos SL, Buchanan TA:**Current concepts: Gestational diabetes mellitus.** *N Engl J Med* 1999, 341:1749-56.
13. Shen H, Liu X, Chen Y, He B, Cheng W:**Associations of lipid levels during gestation with hypertensive disorders of pregnancy and gestational diabetes mellitus: a prospective longitudinal cohort study.** *Bmj Open* 2016, 6.
14. Leng J, Liu G, Zhang C, et al.:**Physical activity, sedentary behaviors and risk of gestational diabetes mellitus: a population-based cross-sectional study in Tianjin, China.** *Eur J Endocrinol* 2016, 174:763-73.
15. Hedderson MM, Gunderson EP, Ferrara A:**Gestational Weight Gain and Risk of Gestational Diabetes Mellitus.** *Obstet Gynecol* 2010, 115:597-604.
16. Catalano PM, McIntyre HD, Cruickshank JK, et al.:**The Hyperglycemia and Adverse Pregnancy Outcome Study Associations of GDM and obesity with pregnancy outcomes.** *Diabetes Care* 2012, 35:780-6.
17. Donovan BM, Breheny PJ, Robinson JG, et al.:**Development and validation of a clinical model for preconception and early pregnancy risk prediction of gestational diabetes mellitus in nulliparous women.** *PLoS One* 2019, 14:e0215173.
18. Caliskan E, Kayikcioglu F, Ozturk N, Koc S, Haberal A:**A population-based risk factor scoring will decrease unnecessary testing for the diagnosis of gestational diabetes mellitus.** *Acta Obstet Gynecol Scand* 2004, 83:524-30.
19. Naylor CD, Sermer M, Chen E, Farine D:**Selective screening for gestational diabetes mellitus. Toronto Trihospital Gestational Diabetes Project Investigators.** *N Engl J Med* 1997, 337:1591-6.

## Figures



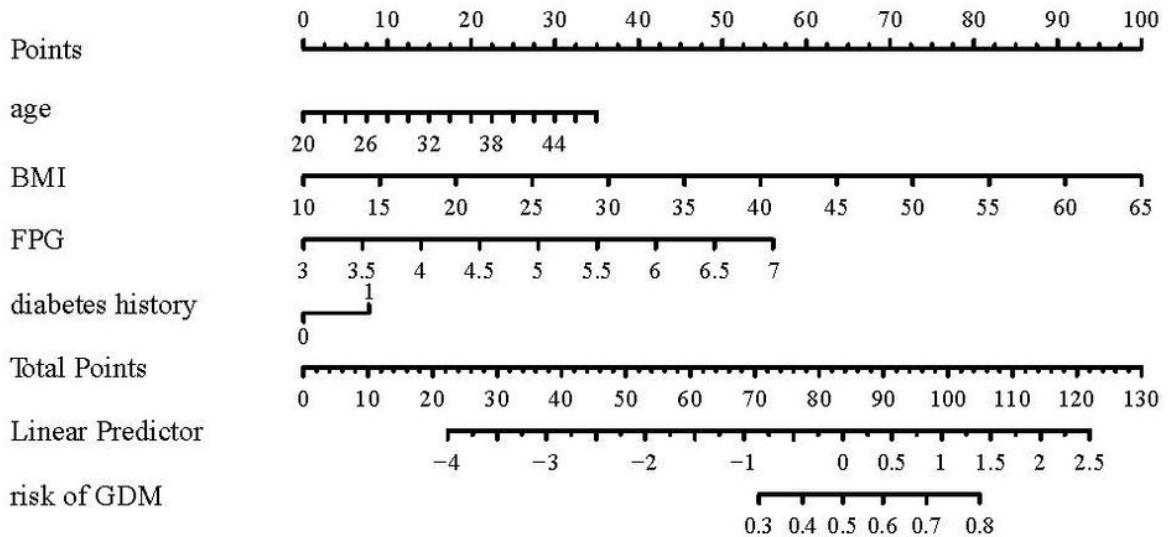
**Figure 1**

ROC curve analysis of combination of basic characteristics and glycemic markers for diagnosing GDM in both cohorts



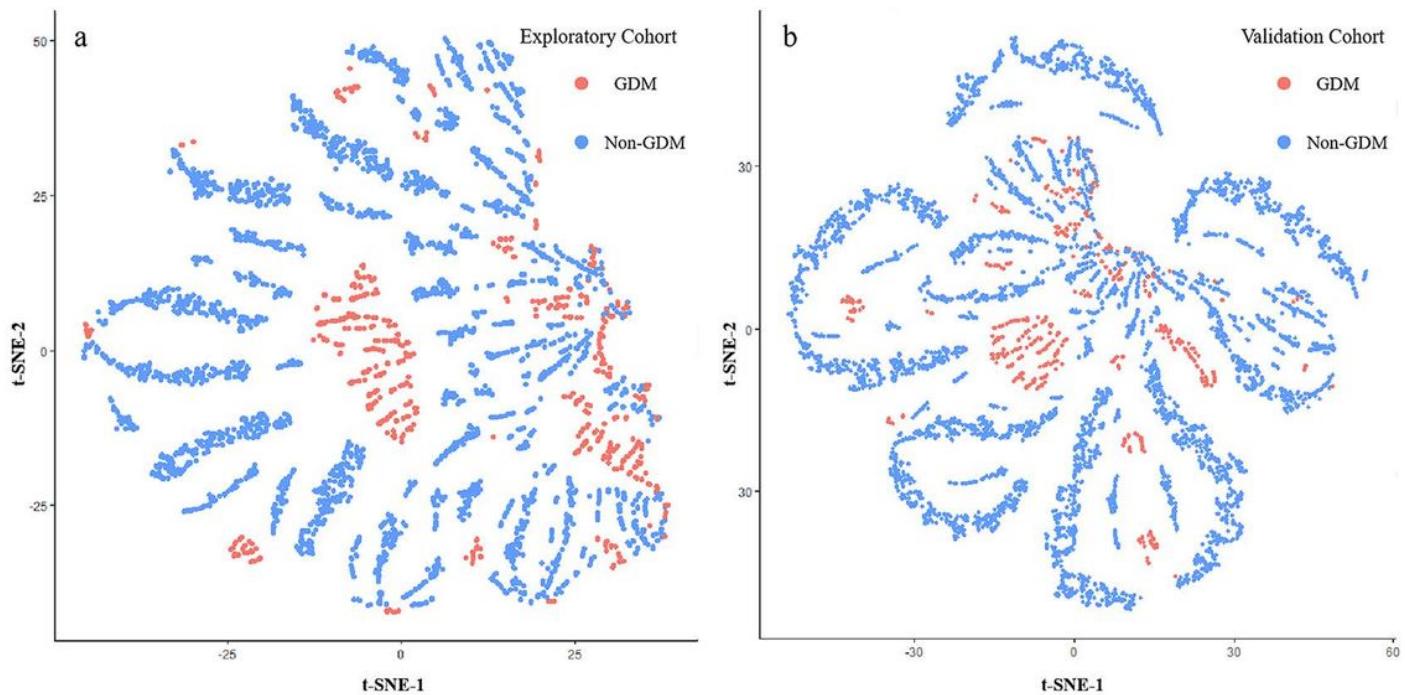
## Figure 2

Calibration results. Comparison of the mean probability estimated in each decile of probability with the observed disease incidence in both cohorts. The models predicted risk very well in both retrospective cohort ( $p=0.51$ ) (a) and prospective cohort ( $p=0.98$ ) (b).



## Figure 3

Nomogram to estimate the risk of GDM. Each predictor is assigned a score on each axis. Compute the sum of points for all predictors and denote this value as the total points. The corresponding "risk of GDM" of "total point" was converted to a predicted probability of GDM. GDM: gestational diabetes mellitus; BMI: body mass index; FPG: fasting plasma glucose.



**Figure 4**

Screening ability of this model. Orange color represented for GDM women, blue point represented for non-GDM women. t-SNE result showed the majority of non-GDM women were separated from GDM women by an obvious boundary in the retrospective cohort (a). t-SNE showed the similar result in the prospective cohort (b). GDM: gestational diabetes mellitus.