Early Prediction Model for Preterm Birth Combining Demographic Characteristics and Clinical Characteristics

Mengfan Chen

Zhejiang University School of Medicine Women's Hospital <https://orcid.org/0000-0001-6924-6787>

Ni Xie

Lishui Maternal and Child Health Care Hospital

Zhaoxia Liang

Zhejiang University School of Medicine Women's Hospital

Tingting Qian

Zhejiang University School of Medicine Women's Hospital

Danqing Chen (✉ chendq@zju.edu.cn)

Zhejiang University, School of Medicine, Women's Hospital, Hangzhou310000, China

Research article

Keywords: Preterm birth, prediction, model, demographic characteristics, Clinical Characteristics

Posted Date: December 2nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-117862/v1>

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

[Read Full License](#)

Abstract

Objective. To create early prediction models for preterm birth (PTB) based on the Chinese population, combining demographic characteristics and clinical characteristics.

Methods. A retrospective study on 15197 pregnant women who were recruited in Obstetrics and Gynecology Hospital of Zhejiang University from January 1, 2017 to December 31, 2017. Demographic characteristics and clinical characteristics were collected and were randomly divided into the observation group (80%) and the validation group (20%). Multivariable Logistics regression analysis was performed to develop a risk prediction model in the observation group and the validation group. It was evaluated by the value of area under the curve (AUC) of receiver operating characteristic (ROC). Finally, we got a simple scoring system to present the preterm birth risk.

Results. There were 1082 pregnant women (8.9%) developed PTB in the observation group and 316 pregnant women (10.3%) in the validation group. Gravidity, educational level, residence, previous history of PTB, twin pregnancy, pre-gestational diabetes mellitus (type I or II), chronic hypertension, placenta previa, gestational hypertension were significant predictors of future PTB. These factors were all included in the model, the AUC was 0.746 with sensitivity of 61.4% (95%CI: 61.4-66.7%) and specificity of 86.6% (95%CI: 85.2-87.9%) at the threshold score of 8.

Conclusion. PTB can be predicted by demographic characteristics and clinical characteristics pre-pregnancy or during pregnancy, thus predicting and preventing PTB as early as possible.

Background

WHO defines preterm birth (PTB) as all births before 37 weeks of gestation, or fewer than 259 days from the first date of a woman's last menstrual period¹, but in China, PTB means delivery after 28 gestational weeks but before 37 gestational weeks. PTB leads to perinatal mortality and morbidity, short-term or long-term complications like respiratory distress syndrome, bronchopulmonary dysplasia, necrotising enterocolitis, intraventricular hemorrhage, infections, feeding difficulties, visual or hearing problems and so on². However prevention of PTB, improvement of the intrauterine condition before delivery, heightening the survival rate of newborns has become a challenging public issue.

A study published by Lacent found the incidence of PTB was 7.3 per 100 births [95% confidence interval (CI): 7.0–7.6%] or 6.7 per 100 live births (95%CI: 6.4–7.0%) during 2015 to 2016 in China³. The rising PTB rate has been reported by recent researches⁴, therefore how to predict and prevent PTB is particularly important. Most studies of PTB prediction involved ultrasound cervical length examination or biomarkers like fetal fibronectin respectively or jointly⁵. Previous studies suggested that PTB was related with maternal demographic characteristics like age, social-economic factors, and clinical characteristics like hypertensive disorders in during or before pregnancy, sexually transmitted disease, history of miscarriage, stillbirth or PTB, multiple pregnancy, placenta abruption, placenta previa, ect³⁶⁷. An American study once

established a prediction model of PTB in 2012⁶, but they served for one or multiple population as most studies did, and few studies provided risk assessment for an individual.

For the first time, this paper combines demographic and clinical characteristics to establish a PTB prediction model and validate it based on a large population of Chinese. We established a simple PTB risk scoring system, a personalized PTB risk assessment, and a theoretical basis to prevent PTB.

Methods

Participants and data

This large retrospective study was conducted among 18000 pregnant women who did prenatal care and finally delivered at Women's Hospital School of Medicine Zhejiang University during January 1, 2017 to December 31, 2017. The exclusion criteria were the following: (1) Therapeutic induced labor because of stillbirth or birth defects (n = 246); (2) Not Han nationality (n = 34); (3) Women with incomplete information (n = 2523). Finally, a total of 15197 women remained in our study. QMmedical electronic medical records system and Union electronic medical records query system were utilized to obtain demographic and clinical characteristics information and telephone follow-up survey was conducted for patients with uncompleted obstetrics history. All women were randomly divided into the observation group and the validation group in our study, the observation group contained about 80% of all (n = 12131), whereas the validation group contained about 20% (n = 3066).

Clinical risk factors

According to previous studies of PTB, we gathered data including age, parity, gravidity, height, weight (before pregnancy), social-economic factors (career, educational level, residence), pre-pregnancy BMI, family history of diabetes mellitus (defined as a first-degree relative with diabetes mellitus type I or II), previous obstetrics history (PTB, spontaneous abortion, macrosomia, therapeutic induced labor because of stillbirth or birth defects), the complications of pregnancy (twin pregnancy, chronic hypertension, hypertensive disorders of pregnancy, pre-gestational diabetes mellitus (type I or II), gestational diabetes mellitus (GDM), placenta previa, scar uterus, uterine myoma) and other clinical data. Pre-pregnancy BMI was calculated as $\text{weight (kg)} / (\text{height (m)})^2$. The pre-pregnancy BMI consisted of four groups by Chinese BMI standard: lean group ($< 18.5 \text{ kg/m}^2$), normal group ($18.5\text{--}23.9 \text{ kg/m}^2$), overweight group ($24\text{--}27.9 \text{ kg/m}^2$) and obese group ($> 28 \text{ kg/m}^2$). According to Chinese physical labor strength classification criteria (GB3869-1997), career was divided into "light labor" (such as staffs, civil servants etc.), "medium labor" (such as teachers, students, individual operators), "heavy labor" (such as workers, farmers, medical workers, police and soldiers etc.) and "others". Educational level was divided into low (lower than college), middle (college), and high (higher than college). Residence was divided into rural and urban areas. The clinical features were classified as positive and negative according to the presence or absence of the corresponding history.

Statistical analysis

For statistical analysis of the data SPSS statistical package for social sciences. Release 21.0 was used. Univariate logistic regression analysis determined significant risk factors for PTB and reported an odds ratio (OR) for each risk factor (unadjusted); multivariate logistic regression identified the adjusted odds ratio (AOR) for all related factors in the scoring system. Bilateral $P < 0.05$ was considered statistically significant.

Clinical scoring system

Data from the observation group were used to develop a simple prediction scoring system. If women's clinical characteristics were significantly associated with the prevalence of PTB ($P < 0.05$) in univariate analysis, they would enter multivariate logistic regression. Clinical score for each risk factor was derived from rounded log adjusted odds ratio score following multivariate logistic regression, which ultimately create a clinical scoring system for independent PTB predictors.

Scoring system performance was examined at different score thresholds with intervals of 0.5 points in the validation group to estimate the optimal scoring threshold. Each score equal to or greater than the threshold score was assumed to be PTB, and then numbers of the assumed and real PTB conditions under each threshold score were separately counted, respectively. Sensitivity, specificity, positive and negative predictive values were calculated for each score, and the Wilson method estimated the 95% CI. Performance of the clinical scoring system would be assessed by AUC.

Results

Demographic and clinical risk factors

Of 12131 women in the observation group, 1082 women (8.9%) developed PTB. The correlation between each risk factor included in logistic regression and PTB at univariate and multivariate levels were shown in Table 1. And we can observe that gravidity, age, pre-pregnancy BMI, educational levels, residence, family history of diabetes mellitus, previous obstetrical history like PTB, spontaneous abortion and complications in pregnancy like twin pregnancy, chronic hypertension, pre-gestational diabetes mellitus (type I or II), placenta previa, scar uterus, gestational hypertension associated with PTB independently ($P < 0.05$). However, parity ($P = 0.329$), previous history of macrosomia ($P = 0.181$), fetal abnormality ($P = 0.111$), GDM in this pregnancy ($P = 0.067$), uterine myoma ($P = 0.549$), career ($P > 0.05$, except "light labor" or "others" groups) were not significantly correlated with PTB.

Table 1

Prevalence and odds ratios of all demographic and clinical characteristics in the observation group.

Characteristics	Total n = 12131	PTB n = 1082	PTB Prevalence	P	OR	95%CI
Demographic Characteristics						
Parity						
0	7096	648	9.1%	-	1.0	-
≥ 1	5035	434	8.6%	0.329	0.9	0.8–1.1
Gravidity						
1 & 2	4773	403	8.4%	-	1.0	-
> 2	7358	679	9.2%	0.000	1.3	1.2–1.5
Maternal age						
< 35 years	9905	857	8.7%	-	1.0	-
≥ 35 years	2226	225	10.1%	0.03	1.2	1.0–1.4
Pre-pregnancy BMI (kg/m ²)						
< 18.5	2002	181	9.0%	0.303	1.1	0.9–1.3
18.5–23.9	8405	700	8.3%	-	1.0	-
24-27.9	1433	163	11.4%	0.000	1.4	1.2–1.7
≥ 28	291	38	13.1%	0.005	1.7	1.2–2.3
Career						
Light labor	9994	876	8.8%	0.000	1.0	-
Medium labor	559	54	9.7%	0.468	1.1	0.8–1.5
Heavy labor	1411	120	8.5%	0.745	1.0	0.8–1.2
Others	167	32	19.2%	0.000	2.5	1.7–3.7

Characteristics	Total n = 12131	PTB n = 1082	PTB Prevalence	P	OR	95%CI
Educational level						
Low	1469	226	15.4%	0.008	0.6	0.5–0.7
Middle	8856	742	8.4%	-	1.0	-
High	1806	74	4.1%	0.000	1.9	1.6–2.2
Residence						
Urban area	9693	739	7.6%	-	1.0	-
Rural area	2438	343	14.1%	0.000	2.0	1.7–2.3
Clinical Characteristics						
Family history of diabetes mellitus						
No	11659	1024	8.8%	-	1.0	-
Yes	472	58	12.3%	0.009	1.5	1.1–1.9
PTB history						
No	11872	1021	8.6%	-	1.0	-
Yes	259	61	23.6%	0.000	3.3	2.4–4.4
Twin pregnancy						
No	11672	834	7.1%	-	1.0	-
Yes	459	248	54.1%	0.000	15.3	12.5–18.6
Placenta previa						
No	11710	934	8.0%	-	1.0	-
Yes	421	148	35.2%	0.000	6.3	5.1–7.7
Gestational hypertension						
No	11550	939	8.1%	-	1.0	-
Yes	581	143	24.6%	0.000	3.7	3.0–4.5

Characteristics		Total n = 12131	PTB n = 1082	PTB Prevalence	P	OR	95%CI
Chronic hypertension							
No		11992	1060	8.8%	-	1.0	-
Yes		139	22	15.8%	0.005	1.9	1.2–3.1
Pre-gestational diabetes mellitus							
No		12105	1074	8.9%	-	1.0	-
Yes		26	8	30.8%	0.023	4.6	2.0–10.5
Stillbirth or birth defects history							
No	11999	1065	8.9%	-	1.0		-
Yes	132	17	12.9%	0.111	1.5		0.9–2.5
Scar uterus							
No	9351	807	8.6%	-	1.0		-
Yes	2780	275	9.9%	0.041	1.2		1.0–1.3
Uterine myoma							
No	11505	1022	8.9%	-	1.0		-
Yes	626	60	9.6%	0.549	1.1		0.8–1.4

Twin pregnancy was the strongest risk factor for subsequent PTB, and it was 16.1 times more likely to have PTB than singleton pregnancy (95% CI: 13.1–19.8, $P < 0.001$). Placenta previa was also a significant risk factor and the prevalence of PTB was 6.6 times higher than women without placenta previa (95% CI: 5.2–8.3, $P < 0.001$). Compared with women who had the first or second pregnancy, more than two pregnancies was 1.2 times (95%CI: 1.0–1.4, $P = 0.018$) to induce PTB, respectively. Women with previous history of PTB were 2.9 times (95%CI: 2.1–4.0, $P < 0.001$) more likely to induce PTB than those without. If this pregnancy was complicated with chronic hypertension, DM or gestational hypertension, the risk of PTB rose to 2.1 times (95%CI: 1.3–3.4, $P = 0.005$), 3.0 times (95%CI: 1.2–7.8, $P = 0.023$), 3.3 times (95% CI: 2.6–4.1, $P < 0.001$), respectively. Interestingly, women who had a higher educational level (university or above) or lived in cities had less incidence of PTB ($P < 0.001$).

Clinical scoring system

After multiple logistic regression, the score for each clinical risk factor in the observation group was derived from the natural logarithm of AOR, multiplied by 10 for convenience (Table 2). For example, more than two deliveries would receive 1.8 points on “gravidity”, while a twin pregnancy would receive 27.7 points on “twin pregnancy”, and so on. A pregnant woman, would receive a total score after counting 9 items like “gravidity”, “educational level”, “residence”, “previous history of PTB”, “twin pregnancy”, “pre-gestational diabetes mellitus (type I or II)”, “chronic hypertension”, “placenta previa”, “gestational hypertension” which was a simple clinical scoring system for PTB.

Table 2
The risk prediction model of preterm birth.

Characteristics	P	AOR	95%CI	Score
Demographic Characteristics				
Gravidity				
1 & 2	-	-	-	0
> 2	0.018	1.2	1.0-1.4	1.8
Educational level				
Low	0.008	0.7	0.5–0.9	-3.5
Middle	-	-	-	0
High	0.000	1.5	1.2–1.7	3.7
Residence				
Urban area	-	-	-	0
Rural area	0.000	1.6	1.4–1.9	4.7
Clinical Characteristics				
PTB history				
No	-	-	-	0
Yes	0.000	2.9	2.1-4.0	10.6
Twin pregnancy				
No	-	-	-	0
Yes	0.000	16.1	13.1–19.8	27.7
Placenta previa				
No	-	-	-	0
Yes	0.000	6.6	5.2–8.3	18.8
Gestational hypertension				
No	-	-	-	0
Yes	0.000	3.7	2.6–4.1	11.8
Chronic hypertension				

Characteristics	P	AOR	95%CI	Score
No	-	-	-	0
Yes	0.005	2.1	1.3–3.4	7.3
Pre-gestational diabetes mellitus				
No	-	-	-	0
Yes	0.023	3.0	1.2–7.8	11.0

Validation of scoring system

Of 3066 women in the validation group, 316 women (10.3%) developed PTB. Table 3 presents performances of the derived clinical scoring system in the validation group. Overall, the most favourable combination of sensitivity and specificity was achieved at the score threshold of 8 with a sensitivity of 61.4% (95%CI: 61.4–66.7%) and a specificity of 86.6% (95%CI: 85.2–87.9%). The receiver-operating characteristic curve in Fig. 1 illustrates the trade-offs between sensitivity and specificity for the scoring system. The area under the curve is shown as 0.749 (95%CI: 0.732–0.767) (Fig. 1), indicating a 74.9% probability that a randomly selected patient with PTB will receive a higher risk score than a randomly selected patient without PTB.

Table 3
The sensitivity and specificity at every score threshold in the validation gr group.

Score threshold	Sensitivity (95%CI)	Specificity (95%CI)	Positive predictive value(95%CI)	Negative predictive value(95%CI)	□□□□
0	0.974(0.948–0.988)	0.113(0.102–0.126)	0.112(0.100–0.124)	0.975(0.949–0.988)	0.089
1	0.898(0.858–0.928)	0.335(0.318–0.353)	0.134(0.120–0.150)	0.966(0.952–0.976)	0.234
2	0.781(0.731–0.825)	0.669(0.651–0.687)	0.213(0.190–0.238)	0.963(0.954–0.971)	0.451
3	0.775(0.724–0.819)	0.676(0.658–0.693)	0.215(0.192–0.241)	0.963(0.953–0.970)	0.452
4	0.759(0.707–0.804)	0.697(0.679–0.714)	0.223(0.199–0.250)	0.961(0.952–0.969)	0.457
5	0.724(0.671–0.772)	0.741(0.724–0.758)	0.243(0.216–0.272)	0.959(0.949–0.966)	0.467
6	0.667(0.612–0.718)	0.797(0.781–0.812)	0.274(0.243–0.308)	0.954(0.944–0.962)	0.465
7	0.613(0.557–0.667)	0.865(0.851–0.877)	0.343(0.304–0.384)	0.951(0.941–0.959)	0.479
8	0.613(0.613–0.667)	0.866(0.852–0.878)	0.345(0.306–0.386)	0.951(0.941–0.959)	0.480
9	0.594(0.538–0.649)	0.878(0.865–0.890)	0.360(0.319–0.403)	0.949(0.940–0.957)	0.474
10	0.591(0.535–0.646)	0.882(0.869–0.893)	0.365(0.324–0.409)	0.949(0.940–0.957)	0.474
11	0.556(0.500–0.612)	0.909(0.897–0.919)	0.413(0.366–0.461)	0.946(0.937–0.955)	0.466
12	0.544(0.487–0.599)	0.922(0.912–0.932)	0.447(0.397–0.499)	0.946(0.936–0.954)	0.467
13	0.518(0.462–0.575)	0.932(0.922–0.941)	0.468(0.415–0.522)	0.944(0.934–0.952)	0.451
14	0.493(0.437–0.550)	0.941(0.931–0.949)	0.490(0.434–0.546)	0.941(0.932–0.950)	0.435
15	0.503(0.447–0.558)	0.943(0.933–0.951)	0.509(0.453–0.565)	0.941(0.932–0.950)	0.446
16	0.490(0.434–0.546)	0.942(0.933–0.951)	0.496(0.440–0.553)	0.941(0.931–0.949)	0.433

Discussion

This retrospective study established an early PTB prediction model by investigating the maternal demographics and clinical characteristics of our hospital in 2017. Our prediction model included 9 items like gravidity, educational level, residence, previous history of PTB, twin pregnancy, pre-gestational diabetes mellitus (type I or II), chronic hypertension, placenta previa, gestational hypertension. The area under the receiver-operating characteristic curve was reported as 0.749 (95%CI: 0.732–0.767). What's more, we found that at the score threshold of 8, the sensitivity and the specificity of this prediction model reached 61.4% (95%CI: 61.4–66.7%) and 86.6% (95%CI: 85.2–87.9%), which were the most favorable combination.

We've found two previous studies of PTB prediction models, Schaaf et al.⁶ reported a large sample of 1524058 singleton pregnancies covering 13 scoring items, and the AUC was 0.63. They also concluded that East Asians had higher PTB risk compared with Caucasians (OR = 1.1, 95%CI: 1.04–1.16). This suggested the necessity of establishing a PTB prediction model for the Chinese population. CELIK et al.⁸ used cervical length to predict PTB with an AUC of 0.903 and the AUC reached 0.919 when combined with other obstetrical history. They predicted PTB for different gestational weeks such as < 28 weeks, 28–30 weeks, 31–33 weeks, and 34–36 weeks. However, none of them verified their model. Although our model didn't reach the sample size as them, it had the AUC of 0.749 and verified it. It provided PTB risk for every single woman simply and intuitively.

Our study found that twin pregnancy was the strongest predictor of PTB, and twin pregnancy was 16.1 times more likely to have PTB than singleton pregnancy (95% CI: 13.1–19.8, $P < 0.001$). Women with previous history of PTB had an increase risk to induce PTB than those without (OR = 2.9, 95%CI: 2.1–4.0, $P < 0.001$). Meta-analysis of 6 cohort studies showed that women with a history of PTB in twin pregnancy had a significantly higher risk of PTB in subsequent singleton pregnancy (OR = 3.34; 95% CI: 1.50–7.45). The earlier delivered twin pregnancy the higher were chances of PTB in subsequent singleton pregnancy⁹. Fuchs et al.¹⁰ pointed out that bed rest, preventive oral inhibitors of uterine contraction, progesterone, preventive cervical cerclage and other methods were common prevention for PTB, but there was no evidence that they could reduce the risk of PTB. Fernandez-Macias et al.¹¹ found that 17-alpha-hydroxyprogesterone caproate (17-OHPC) got a 29%, 26%, 40% reduction in recurrent PTB at < 37, <35, and < 32 gestational weeks respectively, also 68% reduction in neonatal death (95%CI: 0.15–0.66, $P = 0.002$).

In a cohort study involved 7633 pregnant women, Shen et al.¹² observed that both gestational hypertension (AOR = 1.8) and preeclampsia (AOR = 7.1) were significantly associated with PTB ($P < 0.01$), and they also found that folate had a protective effect on early-onset preeclampsia ($P < 0.01$). Premkumar et al.¹³ discovered 8.8% women in 23425 singleton pregnancies developed PTB and 3.8% diagnosed with chronic hypertension. There were increased odds of PTB among women with chronic hypertension (AOR = 2.74, 95%CI: 2.28–3.29, $P < 0.001$). In addition to this, a recent study published in The Lancet showed low-dose aspirin administered between 6 full gestational weeks and less than 14 gestational weeks

among women from low-income and middle-income countries, would reduce the incidence of PTB before 34 weeks in pregnancy complicated with hypertensive disorders of (OR = 0.38, 95%CI:0.17–0.85, P = 0.015)¹⁴. For women with chronic hypertension or gestational hypertension, if they can identify the related high-risk factors, supply folate earlier, keep close prenatal care, control blood pressure during pregnancy and take low-dose aspirin, etc. may reduce the incidence of PTB¹⁴.

In addition, a retrospective study found that 46.8%(29/62) placenta previa cases were presented by PTB, whether it was spontaneous or iatrogenic¹⁵. Previous studies had discovered that the risk factors of placenta previa include multiparas, advanced maternal age (> 35 years old), prior cesarean delivery etc¹⁶. Our study found that increasing maternal age (P = 0.008) and gravidity (P < 0.001), history of cesarean delivery (P = 0.041) heightened the PTB rate at univariate level respectively, but they were excluded by multivariate logistic regression finally. A recent Chinese study has also shown the incidence of PTB among pregnant women after 1, 2 and multiple cesarean sections as 7.8% (64/822), 9.1% (75/822), and 29.7% (244/822), respectively¹⁷. Therefore, encouraging childbirth at child-bearing age and reducing unnecessary abortion or cesarean section can reduce the incidence of placenta previa and PTB. Most noteworthy from our study, pre-gestational diabetes mellitus (type I or II) was significantly related with PTB (OR: 4.57, 95% CI: 1.98–10.52). Wei et al.¹⁸ also reported that the risk of PTB was higher in women who had pre-gestational diabetes mellitus (OR = 1.17, 95% CI: 1.14–1.20; P < 0.001), after a retrospective study of 6447339 Chinese pregnant women who participated in prenatal care. Allison et al.¹⁹ also revealed that PTB was statistically significant with AOR 1.55 (95% CI: 1.03–2.32) in women with type-2 diabetes mellitus. They found that women who did not present for prenatal care until time of delivery had an increased risk of stillbirth and PTB which pointed out the indispensable necessity of routine prenatal care.

Interestingly, we also discovered that women who lived in rural area or had lower educational level (below college degree) were more likely to have PTB, the AOR were 1.59 (95%CI: 1.37–1.86, P < 0.001) and 1.45 (95%CI: 1.22–1.72, P < 0.005), respectively. A study from Norway pointed out that the risk of PTB was 1.5-fold increase in single mothers or mothers with lower educational level²⁰. In Michigan state, USA, every percent increase of unemployment during the first trimester of pregnancy was related to a 3% increase of PTB²¹. A meta-analysis of 12 European cohort studies also revealed that women with lower educational level had a higher PTB rate, the most Significant of which were in Netherlands (P = 0.001) and Norway (P = 0.009). The incidence of PTB was 7.0% among women with lower educational level in Netherlands and 4.9% among those well-educated. Corresponding PTB rate was 9.7% and 5.9% in Norway women respectively²². A systematic review believed that this might be related to the limited data collection and the laggard facilities in economic backward areas²³. We presume that well-educated women and urban women may pay more attention to their perinatal health, may be more likely to attend routine or extra prenatal care persistently. They have access to medical resources when needed, and most of them can get better nutrition during pregnancy. For example, our hospital supplies several free classes to pregnant women per week, which covers some general knowledge of pregnancy complications, home care for

women or newborns and so on. We also provide remote fetal heart monitoring services to women who demand, this may decrease the occurrence of PTB to a certain degree.

All information our prediction model may need, can be easily obtained from women's history and demographic characteristics without blood tests or other invasive inspections (except for examinations that are necessary to assess the severity of pregnancy complications). Therefore, this makes it convenient, practical, efficient and non-invasive. This prediction model helps to identify high PTB risk group or individual during early pregnancy even before pregnancy. The risk of PTB increases according to the rising (or high), scores especially for those with scores greater than 15 points, the risk of PTB is as high as 50%. Close pregnancy monitoring or targeted intervention for women with high PTB risk, may reduce the incidence of PTB. However, there are many limitations in our research. Firstly, it's prone to memory bias being a retrospective study, information bias due to the exclusion of some samples with incomplete information. In addition, our population was concentrated in a large specialist women hospital rather than a community hospital, which may increase the number of women with higher risks. Most of the population sample lives in Hangzhou or surrounding cities in Zhejiang Province, so we cannot represent all regions or ethnicities. Although previous study thought that it might be more efficient when combining demographic and clinical characteristics with cervical length data⁸, we failed to obtain it because our hospital didn't take cervical length measurement as a routine perinatal monitoring content. Thus, future prediction model can consider to include some laboratory indicators and cervical length data before or during pregnancy, which may improve the accuracy and efficacy of the prediction model, besides that, they can verify whether the existing prevention for PTB is effective or not.

Conclusion

In summary, our prediction model predicts the possibility of PTB based on the demographic and clinical characteristics like gravidity, educational level, residence, previous history of PTB, twin pregnancy, pre-gestational diabetes mellitus (type I or II), chronic hypertension, placenta previa, gestational hypertension. We can identify high PTB risk group or individual during early trimester even before pregnancy. Reinforcing intensive perinatal care and preventive measures specific to each risk factor may help reduce PTB and improve pregnancy outcomes.

Abbreviations

Preterm Birth PTB

Area Under the Curve AUC

Receiver Operating Characteristic ROC

Confidence Interval CI

Body Mass Index BMI

Adjusted Odds Ratio AOR

Odds Ratio OR

Gestational Diabetes Mellitus GDM

17-alpha-hydroxyprogesterone caproate 17-OHPC

Declarations

Ethics approval and consent to participate

The study was approved by the Human Research Ethics committee at Zhejiang University School of Medicine, Women's Hospital.

Consent for publication

Not applicable.

Availability of data and material:

The data that support the findings of this study are available in "ScienceDB" with the DOI: 10.11922/sciencedb.00314

Competing interests

The authors declare that they have no competing interests.

Funding

This study was supported in part by grants from Development Center for Medical Science & Technology National Health Commission of the People's Republic of China, "National Key Research and Development Project"2018YFC1004600". It's a national project and does not involve any private or corporate interests.

Authors' contributions:

M.C. and N.X. conceived and designed this project; M.C. did the analysis and wrote the manuscript; L.Z. and D.C. supervised this project; M.C., N.X. and T.Q. helped to collect data and analyzing the data. All authors have contributed to read and agreed the final content of manuscript for submission.

References

1. Organization, W. H. WHO: recommended definitions, terminology and format for statistical tables related to the perinatal period and use of a new certificate for cause of perinatal deaths.

- Modifications recommended by FIGO as amended October 14, 1976. *Acta Obs. Gynecol Scand* **56**, 247–253 (1977).
2. Chawanpaiboon, S. *et al.* Global , regional , and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. *Lancet. Glob. Heal.* (2018). doi:10.1016/S2214-109X(18)30451-0
 3. Chen, C. *et al.* Epidemiology of preterm birth in China in 2015 and 2016: a nationwide survey. *lacent* **392** (2018). doi:10.1016/S0140-6736(18)32702-8
 4. Wadon, M., Modi, N., Wong, H. S., Thapar, A. & Donovan, M. C. O. Recent advances in the genetics of preterm birth. *Ann. Hum. Genet.* **3**, 1–9 (2019).
 5. Asli, Z., Kaplan, O. & Ozgu-erdinc, A. S. Prediction of Preterm Birth: Maternal Characteristics , Ultrasound Markers , and Biomarkers: An Updated Overview. *J. Pregnancy* 1–8 (2018).
 6. Schaaf, J. M., Ravelli, A. C. J., Willem, B., Mol, J. & Abu-hanna, A. Development of a prognostic model for predicting spontaneous singleton preterm birth. *Eur. J. Obstet. Gynecol.* **164**, 150–155 (2012).
 7. Torchin, H. & Ancel, P. Epidemiology and risk factors of preterm birth. *J. Gynecol. Obstet. Biol. la Reprod.* **45**, 1213–1230 (2016).
 8. Celik, E., To, M., Gajewska, K., Smith, G. C. S. & Nicolaides, K. H. Cervical length and obstetric history predict spontaneous preterm birth: development and validation of a model to provide individualized risk assessment. *Ultrasound Obstet. Gynecol.* **31**, 549–554 (2008).
 9. Menzies, R. *et al.* Risk of singleton preterm birth after prior twin preterm birth: a cohort study. *J. Matern. Neonatal Med.* 1–6 (2019). doi:10.1080/14767058.2019.1581166
 10. Fuchs, F. & Senat, M. Multiple gestations and preterm birth. *Semin. Fetal Neonatal Med.* **21**, 113–120 (2016).
 11. Fernandez-Macias, R., Martinez-Portilla, R. J., Cerrillos, L., Figueras, F. & Palacio, M. A systematic review and meta-analysis of randomized controlled trials comparing 17-alpha-hydroxyprogesterone caproate versus placebo for the prevention of recurrent preterm birth. *Int. J. Gynecol. Obstet.* **147**, 156–164 (2019).
 12. Shen, M. *et al.* Comparison of risk factors and outcomes of gestational hypertension and pre-eclampsia. *PLoS One* **12**, (2017).
 13. Premkumar, A., Henry, D. E., Mph, M. M., Ma, S. N. & Norton, M. E. The interaction between maternal race/ethnicity and chronic hypertension on preterm birth. *Am. J. Obstet. Gynecol.* 1–8 (2016). doi:10.1016/j.ajog.2016.08.019
 14. Davies, E. L., Bell, J. S. & Bhattacharya, S. Preeclampsia and preterm delivery: A population-based case-control study. *Hypertens. pregnancy* **35**, 510–519 (2016).
 15. Rangaswamy, M. & Govindaraju, K. Fetomaternal outcome in placenta previa - a retrospective study in teaching hospital. *Int. J. Reprod. Contraception, Obstet. Gynecol.* **5**, 3081–3084 (2016).
 16. Zlatnik, M. G., Cheng, Y. W., Norton, M. E., Thiet, M.-P. & Caughey, A. B. Placenta previa and the risk of preterm delivery. *J. Matern. neonatal Med.* **20**, 719–923 (2007).

17. Tu, S. & He, J. Clinical Analysis of Pregnancy Outcome in Pregnant Women with Multiple Cesarean Section. *Chinese J. Obstet. Gynecol.* **54**, 615–619. (2019).
18. Wei, Y. *et al.* Preconception diabetes mellitus and adverse pregnancy outcomes in over 6.4 million women: A population-based cohort study in China. *PLoS One* **16**, e1002926 (2019).
19. Allen, A. J., Snowden, J. M., Lau, B., Cheng, Y. & Aaron, B. Type-2 diabetes mellitus: does prenatal care affect outcomes? *J. Matern. neonatal Med.* **31**, 93–97 (2017).
20. Oftedal, A.-M., Busterud, K., Irgens, L. M., Haug, K. & Rasmussen, S. Socio-economic risk factors for preterm birth in Norway 1999-2009. *Scand. J. Public Health* **44**, 587–592. (2016).
21. Margerison, C. E., Luo, Z. & Li, Y. Economic conditions during pregnancy and preterm birth: A maternal fixed-effects analysis. *Paediatr. Perinat. Epidemiol.* **33**, 154–161 (2019).
22. Ruiz, M. *et al.* Mother ' s education and the risk of preterm and small for gestational age birth : a DRIVERS meta-analysis of 12 European cohorts. *J. Epidemiol. Community Heal.* **69**, 826–833 (2015).
23. Vogel, J. P. *et al.* Global, regional and national levels and trends of preterm birth rates for 1990 to 2014: protocol for development of World Health Organization estimates. *Reprod. Health* **13**, 76 (2016).

Figures

AUC	p	95%CI
0.749	0.000	0.732-0.767

Figure 1

The receiver-operating characteristic curve of the validation group

AUC	p	95%CI
0.749	0.000	0.732–0.767

Figure 1

The receiver-operating characteristic curve of the validation group