Development and validation of a Non-INvaSive Pregnancy RIsk ScoRE (INSPIRE) for the screening of high-risk pregnant women for gestational diabetes mellitus in Pakistan

Sabahat Naz [ORCID] ,[1] Samreen Jamal,[1] Ali Jaffar,[2] Iqbal Azam,[1] Subhash Chandir,[1] Rahat Qureshi,[3] Neelofur Babar,[4] Aisha Syed Wali,[5] Romaina Iqbal [ORCID] [1]

[1]Department of Community Health Sciences, The Aga Khan University, Karachi, Pakistan
[2]National Institute of Cardiovascular Diseases, Karachi, Pakistan
[3]Department of Obstetrics and Gynecology, The Aga Khan University, Karachi, Pakistan
[4]The Aga Khan Hospital Karimabad, Karimabad, Pakistan
[5]Aga Khan Hospital Garden, Karachi, Pakistan

Correspondence to
Dr Romaina Iqbal;
romaina.iqbal@aku.edu

ABSTRACT

Introduction The prevalence of gestational diabetes mellitus (GDM) is on the rise in low-income and middle-income countries, such as Pakistan. Therefore, the development of a risk score that is simple, affordable and easy to administer is needed. Our study aimed to develop a Non-INvaSive Pregnancy RIsk ScoRE (INSPIRE) for GDM screening in Pakistani pregnant women based on risk factors reported in the literature.

Methods Using a cross-sectional study design, we enrolled 500 pregnant women who attended antenatal clinics at one tertiary and two secondary care hospitals in Karachi between the 28th and 32nd weeks of gestation. We randomly divided data into derivation (n=404; 80%) and validation datasets (n=96; 20%). We conducted interviews to collect information on sociodemographic factors and family history of diabetes, measured mid-upper arm circumference (MUAC) and reviewed the medical records of women for obstetric history and oral glucose tolerance test (OGTT) results. We performed a multivariable logistic regression analysis to obtain coefficients of selected predictors for GDM in the derivation dataset. Calibration was estimated using Pearson's $\chi 2$ goodness of fit test while discrimination was checked using the area under the curve (AUC) in the validation dataset.

Results Overall, the GDM prevalence was 26% (n=130). INSPIRE was based on six predictors: maternal age, MUAC, family history of diabetes, a history of GDM, previous bad obstetrical outcome and a history of macrosomia. INSPIRE achieved a good calibration (Pearson's $\chi 2$=29.55, p=0.08) and acceptable discrimination with an AUC of 0.721 (95% CI 0.61 to 0.83) with a sensitivity of 74.1% and specificity of 59.4% in the validation dataset.

Conclusion We developed and validated an INSPIRE that efficiently differentiates Pakistani pregnant women at high risk of GDM from those at low risk, thus reducing the unnecessary burden of the OGTT test.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The prevalence of gestational diabetes mellitus (GDM) is on the rise in low-income and middle-income countries, such as Pakistan, where the diagnoses for GDM are both financially burdensome and logistically challenging.
⇒ The development of a risk score that is simple, affordable and easy to administer is needed.

WHAT THIS STUDY ADDS

⇒ Non-INvaSive Pregnancy RIsk ScoRE (INSPIRE) achieved a good calibration and acceptable discrimination, with a sensitivity of 74.1% and specificity of 59.4% in the validation dataset.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ INSPIRE efficiently differentiates Pakistani pregnant women at high risk of GDM from those at low risk, thus reducing the unnecessary burden of the oral glucose tolerance test.
⇒ INSPIRE also offers the potential for early GDM screening for timely intervention among pregnant women in low-resource settings, such as Pakistan.

INTRODUCTION

Over the last few decades, Pakistan has witnessed an escalating trend in the prevalence of gestational diabetes mellitus (GDM), from 6.3% in 2003 to 19% in 2018.[1] GDM is an abnormal glucose tolerance with the onset or first recognition during pregnancy or subsequent pregnancies.[2] Women with GDM are at a greater risk of developing many short-term and long-term issues.[3] Some short-term complications involve caesarean section, pregnancy-induced hypertension, premature rupture of membrane, antepartum and postpartum haemorrhage.[4] Approximately 17%–33% of South Asian women with GDM progress to T2DM within 5–10 years after their index pregnancy.[5] Therefore, early screening of high-risk women is crucial in

providing timely intervention and preventing the development of GDM and many associated short-term and long-term complications.

The current protocol for diagnosing GDM involves a one-step 75 g oral glucose tolerance test (OGTT) conducted between the 24th and 28th weeks of gestation.[6] The OGTT involves administering a 75 g glucose load and evaluating glucose levels after 1, 2 and often 3 hours. A diagnosis of GDM is established if one or more glucose values are equal to or exceed the specified glucose thresholds.[7] However, administering an OGTT poses various challenges, including non-compliance issues due to its invasive nature, as women feel discomfort due to associated nausea and vomiting from ingesting a fixed amount of glucose, along with undergoing multiple blood draws.[8] In addition, the process requires significant time commitment as the whole process takes approximately 4–5 hours to complete, and women, particularly those employed in daily wage labour, are required to take a day off from work and manage household responsibilities, which is another major concern for homemakers. Moreover, the cost associated with the test is another major hurdle for women from low-income and middle-income families and those residing in remote settings.

Since the prevalence of GDM is on the rise in low-income and middle-income countries (LMICs), such as Pakistan, where diagnoses for GDM are both financially burdensome and logistically challenging, the development of a risk score that is simple, affordable and easy to administer is needed.[9] A risk score objectively estimates the probability of the presence or future development of an adverse health condition based on a combination of risk factors.[10] Risk scores are generally developed using an epidemiological approach that links risk factors (eg, weight, a history of GDM and family history of diabetes) with the outcomes (eg, GDM) and should be validated among the target population.[11]

Many countries have developed risk scores to identify and offer the OGTT test only to women at high risk due to cost and discomfort.[12–14] However, those risk scores are not applicable to South Asian women due to the differences in the risk attributes, such as ethnicity, sociodemographic factors, body composition, and other obstetrical factors.[15 16] In addition, the risk factors used in the risk scores are complex to administer to every pregnant woman, especially in LMICs. For example, Gao et al developed a risk score to predict GDM in Chinese pregnant women, in which one of the factors is alanine transaminase (ALT), which is not a routine laboratory test for pregnant women, especially in LMICs, and hence limit its applicability.[12] Kumar et al developed a risk score using easily measurable predictors, including arterial blood pressure, maternal age, a history of GDM and ethnicity for Singaporean women. However, this risk score relies on an artificial intelligence prediction model, necessitating trained personnel to input women's information and generate GDM risk predictions.[17] A locally relevant risk score would enable identifying high-risk pregnant women

for further referrals for OGTT, reducing the unnecessary burden on low-risk pregnant women. Therefore, this study aimed to develop a Non-INvaSive Pregnancy RIsk ScoRE (INSPIRE) for GDM screening in Pakistani pregnant women based on risk factors reported in the literature and validate it with the 2-hour 75 g OGTT.

MATERIALS AND METHODS

Study design, setting and duration

Using a cross-sectional study design, we developed and validated an INSPIRE for screening pregnant women at risk of GDM. The validation process employed criterion validity, involving validating an INSPIRE against the gold standard, that is, 2-hours, 75 g OGTT.[10] The study was conducted at the Aga Khan University Hospital (AKUH) main Stadium Road Campus and its two secondary care hospitals in Karimabad and Garden from February to May 2016.

Recruitment and data collection

Pregnant women who visited the AKUH main campus and its two secondary care hospitals in Karimabad and Garden antenatal clinics were approached between the 28th and 32nd weeks of gestation. Women were purposively selected based on the predefined criteria, such as women with singleton pregnancy, aged 18–45, who already had the OGTT between the 24th and 28th weeks of gestation, and their OGTT results were available in the hospital medical record. Those with known diabetes, cardiac disease, renal failure, taking medications that influenced glucose metabolism and incomplete medical records were not invited to participate. Detailed information on the study's aims and processes was provided to the participants. Written informed consent was obtained for those who met eligibility and agreed to participate. Data collectors were trained to review the medical records of women to extract information on obstetric history (a history of GDM, history of abortion, miscarriages, stillbirth, intrauterine death, macrosomia) and the results of OGTT. They were also trained to conduct face-to-face interviews with the participants using a structured questionnaire to obtain sociodemographic information (age, education, occupation, language and household income) and family history of diabetes. The data collectors were also trained to perform anthropometry measurements, such as height, weight and mid-upper arm circumference (MUAC).

Sample size

A total of 402 pregnant women were required to achieve 80% power, considering an anticipated prevalence of GDM 19%,[1] a precision of 0.10 between the area under the curve (AUC) under the null hypothesis of 0.80 and 5% level of significance using a two-sided z-test.[12] The sample size was calculated using PASS V.11.

Statistical analysis

We enrolled 500 pregnant women and randomly split them into two subsets: the derivation dataset, which

Table 1 Characteristics of participants based on the GDM status (N=500)

Characteristics	GDM (n=130) n (%)	Non-GDM (n=370) n (%)	P value
Sociodemographic			
Age (years); mean±SD	30.3±5.1	27.4±4.6	<0.001
<25	16 (12.3)	98 (26.5)	
25–29	48 (36.9)	151 (40.8)	
30 and above	66 (50.8)	121 (32.7)	<0.001
Education			0.74
No formal education	12 (9.2)	44 (11.9)	
Matriculation (class 10)	15 (11.5)	32 (8.6)	
Intermediate (class 12)	25 (19.2)	80 (21.6)	
Graduate (class 14)	40 (30.8)	116 (31.4)	
Postgraduate (class>14)	38 (29.2)	98 (26.5)	
Occupation			0.33
Working	19 (14.6)	68 (18.4)	
Homemaker	111 (85.4)	302 (81.6)	
Language			0.37
Urdu	93 (71.5)	249 (67.3)	
Others	37 (28.5)	121 (32.7)	
Household monthly income (Pkr); median (IQR)	50 000 (65 000)	50 000 (65 000)	0.88
Health-related			
Height (cm); mean±SD	156.2±6.0	156.3±7.9	0.85
Weight (kg); mean±SD	77.2±17.5	69.0±12.4	<0.001
MUAC (cm); mean±SD	31.0±3.9	29.4±3.1	<0.001
>32	42 (32.3)	57 (15.4)	
≤32	88 (67.7)	313 (82.6)	<0.001
Family history of diabetes			<0.001
Yes	107 (82.3)	166 (44.9)	
Number of pregnancies			0.07
Primiparous	42 (32.3)	152 (41.1)	
Multiparous	88 (67.7)	218 (58.9)	
A history of GDM			<0.001
Yes	43 (51.8)	5 (2.3)	
Adverse obstetrical history*			<0.001
Yes	19 (21.6)	9 (4.1)	
History of macrosomia (baby birth weight >4 kg)			<0.001
Yes	15 (18.1)	7 (3.3)	

*Adverse obstetrical history includes miscarriages, abortions, stillbirths and intrauterine death.
GDM, gestational diabetes mellitus; MUAC, mid-upper arm circumference.

included 80% of the sample (n=404) and the validation dataset, which included the remaining 20% (n=96).

The characteristics of the study participants were presented as mean±SD for normally distributed continuous variables, the median and IQR for skewed continuous variables, and frequencies with percentages for categorical variables. The baseline characteristics were compared based on the GDM status and the derivation and validation datasets. The two datasets were compared using the independent t-test for continuous variables and the $\chi 2$ test for categorical variables. The derivation dataset was used to develop the risk score, and the validation dataset was used to validate its performance. Data were analysed by using Stata (V.17, StataCorp).

Development of INSPIRE

An extensive literature review was conducted to identify risk factors associated with the development of GDM. Based on the risk factors found in the literature, a list of variables was selected to develop the risk scores (online supplemental table 1). The variables included were sociodemographic factors,[18–20] such as maternal age (in years at the time of pregnancy), education and occupation of women and household monthly income, anthropometry, that is, MUAC was measured in cm. Information on familial risk factors,[21] such as family history of diabetes, as well as obstetric risk factors,[22] including parity (number of times a woman has given birth to a fetus with a gestational age of 24 weeks or more, regardless of whether the child was born alive or was stillborn), a history of GDM, previous adverse obstetrical history (miscarriages, abortions, stillbirths and intrauterine death), macrosomia (previous baby with the birth weight >4 kg) was recorded.

We developed INSPIRE from the derivation dataset (n=404) using the multivariable logistic regression model. First, the univariate logistic regression analysis was performed to obtain ORs and 95% CIs. The outcome variable was the development of GDM, and the independent variables were the characteristics of participants in terms of sociodemographic, anthropometry, familial and obstetric risk factors. In the univariate analysis, the association between each independent variable (maternal age, MUAC, family history of diabetes, a history of GDM, parity, education, occupation, household monthly income, adverse obstetrical history, and history of macrosomia) and the development of GDM was assessed using a significance level of p<0.25. We categorised age into a binary variable: <25 years and ≥25 years, based on the evidence suggesting that Asian women aged 25 years and above are at a higher risk of developing GDM.[23] Women with MUAC>32.0 cm were considered obese as it aligned with a prepregnancy body mass index (BMI) >30 kg/m². [24] Using the stepwise forward selection approach in the multivariable logistic regression analysis, only those variables found to be statistically significant (p<0.05) or judged to be clinically important were retained. The regression coefficient (β) was obtained and rounded to

Table 2 Parameter estimates of INSPIRE for the screening of GDM in the derivation dataset

Variables used in the development of INSPIRE	Univariate analysis	Multivariable analysis		
	OR (95% CI)	β (SE)	OR (95% CI)	Score (0–10)
Maternal age (years)				
<25	REF	REF	REF	
≥25	3.59 (1.73 to 7.46)	0.86 (0.44)	2.36 (1.03 to 5.41)	1
MUAC (in cm)				
≤32	REF	REF	REF	
>32	2.48 (1.48 to 4.16)	0.70 (0.33)	2.02 (1.07 to 3.82)	1
Family history of diabetes				
No	REF	REF	REF	
Yes	5.89 (3.34 to 10.39)	1.38 (0.33)	3.99 (2.09 to 7.60)	1
A history of GDM				
No	REF	REF	REF	
Yes	39.89 (13.73 to 115.89)	3.31 (0.58)	27.27 (8.80 to 84.47)	3
History of adverse obstetric outcome				
No	REF	REF	REF	
Yes	18.27 (5.20 to 64.15)	2.05 (0.73)	7.78 (1.88 to 32.18)	2
History of macrosomia				
No	REF	REF	REF	
Yes	8.55 (2.97 to 24.63)	1.50 (0.67)	4.35 (1.19 to 15.96)	2
Formal education				
Yes	REF			
No	0.60 (0.27 to 1.34)	NS		
Occupation				
Working	REF			
Homemaker	1.27 (0.69 to 2.33)	NS		
Household monthly income (pkr)				
>50 000	REF			
≤50 000	0.86 (0.55 to 1.35)	NS		
Parity				
Primiparous	REF			
Multiparous	1.40 (0.88 to 2.25)	NS		

GDM, gestational diabetes mellitus; INSPIRE, Non-INvaSive Pregnancy RIsk ScoRE; MUAC, mid-upper arm circumference; NS, not significant.

the nearest integer to assign a score to each variable in the final multivariable logistic regression model.

Validation of INSPIRE
To assess the performance of INSPIRE on the validation dataset, the derived scores were applied to the validation dataset and divided into deciles according to their predicted probability of GDM. The observed and expected probabilities of GDM in the deciles were compared. Pearson's χ^2 goodness of fit test was employed to check the calibration of INSPIRE. A p value of more than 0.05 was considered an acceptable calibration. The discrimination of INSPIRE was assessed based on the probability of GDM derived from the logistic regression equation and the simplified risk score. Discrimination was measured by assessing the AUC in a receiver operating characteristic by plotting the sensitivity on the y-axis versus the false positives (1−specificity) on the x-axis. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and Youden index at different cut-off points of the risk score were calculated, and a cut-off point was identified to distinguish high-risk women from the low risk to use the risk score in antenatal care (ANC) settings.

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

RESULTS

Characteristics of the study participants

Among the 500 women in the study, 130 (26%) developed GDM. While comparing the characteristics of women with GDM to those without GDM in table 1, women with GDM were more likely to be older and had higher body weight and MUAC. In addition, more women with GDM had a positive family history of diabetes, a history of GDM, along with previous bad obstetrical outcomes and a history of macrosomia.

We also compared the baseline characteristics of women based on the derivation and validation datasets in online supplemental table 2. The characteristics of women were comparable concerning all variables in the derivation and validation datasets, except for household monthly income (p=0.03), indicating that the random allocation of women in the derivation and validation datasets worked well.

Development of INSPIRE

The derivation dataset (n=404) had 103 (25.5%) GDM cases. The selected predictors, their ORs with 95% CI in the univariate analysis and the regression coefficients (β) with SEs and ORs with 95% CI in the multivariable analysis are presented in table 2.

Among the potential predictors, education, occupation, household monthly income and parity were no longer significant and thus not included in the multivariable analysis. Consequently, the INSPIRE was based on six predictors: maternal age, MUAC, family history of diabetes, a history of GDM, previous bad obstetrical outcome and a history of macrosomia.

Based on the final model, women 25 years and above OR 2.36 (95% CI 1.03 to 5.41), MUAC>32 cm OR 2.02 (95% CI 1.07 to 3.82), family history of diabetes OR 3.99 (95% CI 2.09 to 7.60), a history of GDM OR 27.27 (95% CI 8.80 to 84.47), history of bad obstetrical outcome OR 7.78 (95% CI 1.88 to 32.18), and a history of macrosomia OR 4.35 (95% CI 1.19 to 15.96) was significantly associated with the development of GDM with an overall X^2 value of 143.86.

Based on the final model's regression coefficients (β), the score was calculated for each variable where we rounded off coefficients (β) to the nearest integer as described in table 2.

Validation of INSPIRE

Based on the probability of GDM derived from the logistic regression equation

The validation dataset (n=96) had 27 (28.1%) GDM cases. We estimated the probability of GDM in the validation dataset. We derived an equation from the logistic regression model and used it to calculate the probability of GDM.

$$\text{Probability of GDM} = \frac{e}{1+e}$$

$$\text{Probability of GDM} = \frac{\begin{aligned} &-3.41 + 0.86 \times \text{maternal age} \\ &(1 \text{ if } \geq 25 \text{ years, } 0 \text{ if } < 25 \text{ years}) \\ &+0.70 \times \text{MUAC } (1 \text{ if } > 32 \text{ cm, } 0 \text{ if } \leq 30 \text{ kg/m}^2) \\ &+1.38 \times \text{family history of diabetes } (1 \text{ if yes, } 0 \text{ if no}) \\ &+3.31 \times \text{previous history of GDM } (1 \text{ if yes, } 0 \text{ if no}) \\ &+2.05 \times \text{history of bad obstetric outcome } (1 \text{ if yes, } 0 \text{ if no}) \\ &+1.47 \times \text{history of macrosomia } (1 \text{ if yes, } 0 \text{ if no}) \end{aligned}}{\begin{aligned} &1 + (-3.41 + 0.86 \times \text{maternal age} \\ &(1 \text{ if } \geq 25 \text{ years, } 0 \text{ if } < 25 \text{ years}) \\ &+0.70 \times \text{MUAC } (1 \text{ if } > 32 \text{ cm, } 0 \text{ if } \leq 30 \text{ kg/m}^2) \\ &+1.38 \times \text{family history of diabetes } (1 \text{ if yes, } 0 \text{ if no}) \\ &+3.31 \times \text{previous history of GDM } (1 \text{ if yes, } 0 \text{ if no}) \\ &+2.05 \times \text{history of bad obstetric outcome } (1 \text{ if yes, } 0 \text{ if no}) \\ &+1.47 \times \text{history of macrosomia} \times (1 \text{ if yes, } 0 \text{ if no}) \end{aligned}}$$

INSPIRE had a good calibration, with the predicted probabilities of GDM being similar to the observed probabilities (Pearson's $\chi 2$=29.55, p=0.08) (figure 1). INSPIRE achieved an AUC of 0.721 (95% CI 0.61 to 0.83) with a sensitivity of 74.1% and specificity of 59.4% in the validation dataset (figure 2).

The sensitivity, specificity, PPV and NPV of INSPIRE at different cut-off points are summarised in table 3. We selected the cut-off of 0.208 to screen high-risk women for GDM, with a sensitivity of 74.1%, specificity of 59.4%, PPV of 41.7% and NPV of 85.4%. The selected cut-off score achieved a Youden index value of 0.34.

Based on the simplified risk scores

Table 4 presents the INSPIRE risk score. The risk factors are shown in the first column, followed by the specific questions related to the risk factors in the middle column. Only one option should be selected for each question, and the points associated with the chosen option should be written in the last column. All questions should be answered to estimate the accurate risk score for GDM risk.

Based on the risk score, INSPIRE achieved an AUC of 0.703 (95% CI 0.59 to 0.82) with a sensitivity of 74.1% and specificity of 56.5% in the validation dataset (figure 2).

The sensitivity, specificity, PPV and NPV of INSPIRE at different cut-off points are summarised in table 5. We selected the cut-off of 2 to screen high-risk women for GDM, with a sensitivity of 74.1%, specificity of 56.5%, PPV of 40.0% and NPV of 84.8%. The selected cut-off score achieved a Youden index value of 0.31.

DISCUSSION

The prevalence of GDM continues to increase among South-Asian women, including Pakistan. An INSPIRE that is easy to administer, with adequate performance,

Pearson's χ2 =29.55, p=0.08
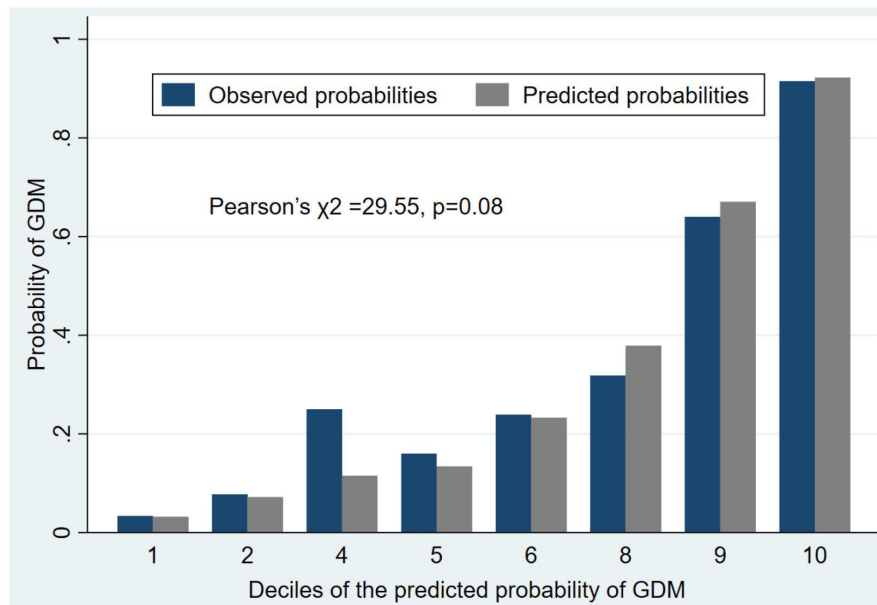
Figure 1 The predicted and observed probability of GDM. GDM, gestational diabetes mellitus.

could serve as an initial screening step, distinguishing high-risk pregnant women for GDM from those at low risk for further referral to the diagnostic test, such as OGTT. This approach aims to mitigate challenges related to financial burden and logistical issues associated with OGTT, specifically in low-resource settings. Our study developed and validated INSPIRE for screening high-risk pregnant women for GDM. We found six predictors, such as maternal age, MUAC, a history of GDM, family history of diabetes, previous bad obstetrical outcome and a history of macrosomia associated with the risk of developing GDM. INSPIRE had a good calibration with the p value of Pearson's χ2 0.08, with adequate discrimination

showing an AUC of 0.721 (95% CI 0.61 to 0.83) with a sensitivity of 74.1% and specificity of 59.4% for the validation dataset.

Naylor *et al* developed and validated the first clinical scoring system for GDM prediction on different ethnic groups in the West.[25] The risk score was based on age, race and prepregnancy BMI. Since the scores were derived from Europeans, Americans and Asians living in Canada, the applicability of the risk score to the Pakistani population is limited. However, recognising the significance of higher BMI and age as a risk factor for GDM,[21] we incorporated it in INSPIRE. Literature supported a correlation between maternal MUAC during pregnancy
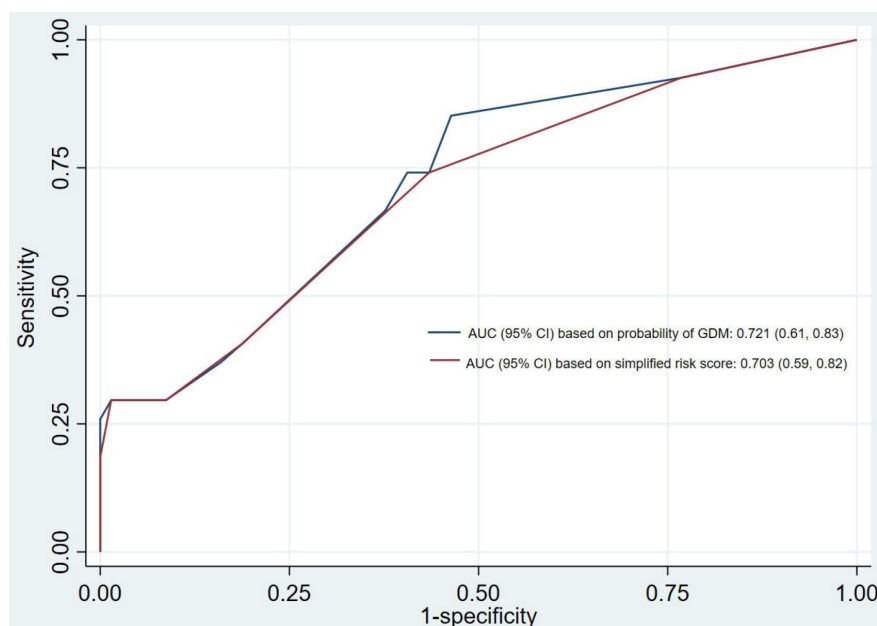
Figure 2 ROC curve of INSPIRE based on probability of GDM and simplified risk scores in the validation dataset. AUC, area under the curve; GDM, gestational diabetes mellitus; INSPIRE, Non-INvaSive Pregnancy RIsk ScoRE; ROC, receiver operating characteristic curve.

Table 3 Sensitivity, specificity and predictive values at different cut-offs of INSPIRE in the validation dataset

GDM risk score cut-off points	Sensitivity %	Specificity %	PPV %	NPV %	Youden index
0.072	92.6	23.2	32.1	88.9	0.16
0.115	85.2	53.6	41.8	90.2	0.39
0.125	74.1	56.5	40	84.8	0.31
0.135	74.1	57.9	40.8	85.1	0.32
0.208*	74.1	59.4	41.7	85.4	0.34
0.235	66.7	62.3	40.9	82.7	0.29
0.375	40.7	81.2	45.8	77.8	0.22
0.382	37.0	84.1	47.6	77.3	0.21
0.679	29.6	91.3	61.5	77.1	0.21
0.705	29.6	92.6	61.5	77.1	0.22
0.729	29.6	98.5	88.9	78.2	0.28
0.809	25.9	100	100	77.5	0.26
0.894	22.2	100	100	76.7	0.22
0.944	18.5	100	100	75.8	0.19
0.985	11.1	100	100	74.2	0.11
0.987	3.70	100	100	72.6	0.04

*The suggested cut-off point.
GDM, gestational diabetes mellitus; INSPIRE, Non-INvaSive Pregnancy RIsk ScoRE; NPV, negative predictive value; PPV, positive predictive value.

Table 4 The INSPIRE risk score

Risk factors	Question (select only one option)	Point for answer	Point for each section
Maternal age	Is a woman aged <25 years? OR Is a woman aged ≥25 years?	0 1	Points:
MUAC	Is a woman's MUAC≤32 cm? OR Is a woman's MUAC>32 cm?	0 1	Points:
Family history of diabetes	Does the woman have a family history of diabetes?		
	No	0	
	Yes	1	Points:
A history of GDM	Has the woman experienced GDM in a previous pregnancy?		
	No	0	
	Yes	3	Points:
A history of bad obstetric outcome	Has the woman experienced a history of stillbirth, abortion, miscarriage or intrauterine death?		
	No	0	
	Yes	2	Points:
A history of macrosomia	Did the woman have a previous baby with a birth weight >4 kg?		
	No	0	
	Yes	2	Points:
		Total points:	

GDM, gestational diabetes mellitus; INSPIRE, Non-INvaSive Pregnancy RIsk ScoRE; MUAC, mid-upper arm circumference.

Table 5 Sensitivity, specificity and predictive values at different cut-offs of INSPIRE in the validation dataset

GDM risk score cut-off points	Sensitivity %	Specificity %	PPV %	NPV %	Youden index
1	92.6	23.2	32.1	89.0	0.16
2*	74.1	56.5	40.0	84.8	0.31
3	40.7	81.2	45.8	77.8	0.22
4	29.6	91.3	57.1	76.8	0.21
5	29.6	98.5	89.0	78.2	0.28
6	18.5	100	100	75.8	0.19
7	11.1	100	100	74.2	0.11
8	3.70	100	100	72.6	0.04

*The suggested cut-off point.
GDM, gestational diabetes mellitus; INSPIRE, Non-INvaSive Pregnancy RIsk ScoRE; NPV, negative predictive value; PPV, positive predictive value.

and prepregnancy BMI, irrespective of gestational age.[24] Given the unavailability of information on the women's prepregnancy weight status in our setting, MUAC serves as a valuable proxy. Since MUAC can be easily measured during ANC services, we used it as a surrogate for prepregnancy BMI and found it a significant predictor for GDM.

Many risk scores have been developed to identify high-risk women for GDM.[12–14] However, these risk scores have limited applicability for women in low-resource settings like Pakistan. For instance, Gao et al derived risk scores on Chinese pregnant women based on early pregnancy risk factors (maternal age, BMI, height, systolic BP, ALT and family history of diabetes), as well as four during pregnancy modifiable risk factors (physical activity, sitting time at home, passive smoking and weight gain from registration to Glucose Challenge Test), had adequate calibration (p value for Hosmer Lemshow test >0.25) and discrimination (AUC 0.71; 95% CI 0.68 to 0.74).[12] However, implementing such a risk score is difficult as ALT is not a routine test in ANC services in our setting. In addition, the applicability of this risk score is limited to only those women who seek proper ANC services to collect information on weight gain. Furthermore, including many continuous factors in the risk score makes it complex, requiring skilled personnel for accurate calculation. In contrast, we dichotomised all six potential risk factors in INSPIRE, facilitating its practical application in clinical settings.

Another risk score developed for Tanzanian women aimed at identifying high-risk women for GDM included only three risk factors, such as MUAC, a history of stillbirth and family history of diabetes, with an AUC of 0.64 (95% CI 0.56 to 0.72).[13] However, this risk score has some limitations as it missed crucial risk factors, that is, maternal age and a history of GDM and had lower predictive capability. Literature supports the causal relationship between maternal age and a history of GDM with the risk

of developing GDM.[21] Considering the significance of the causal relationship, we included both risk factors in INSPIRE and found them to be significant predictors of GDM.

We found the a history of GDM to be a highly significant predictor for GDM as we have a well-distributed representation of both primiparous and multiparous women in our derivation and validation datasets. Including this predictor in our analysis addresses a limitation observed in many existing GDM risk scores that have overlooked the potential impact of the history of GDM on the development of GDM,[12 13] enhancing its applicability among women with varied parity backgrounds.

The relationship between a history of macrosomia (baby birth weight >4 kg) and the risk of GDM is well established due to the elevated maternal blood glucose levels passing through the placenta to the fetus, causing macrosomia characterised by increased fetal body fat deposition.[26] INSPIRE aligns well with the existing evidence and observed history of macrosomia as a significant predictor for the risk of GDM.

For our risk score, we opted for a minimum cut-off of 2 to screen high-risk women for GDM, with a sensitivity of 74.1%, specificity of 56.5% and an AUC of 0.703 (95% CI 0.59 to 0.82). This choice aligns with the nature of the screening tool, where higher sensitivity is preferred to minimise false negative results. Using the cut-off value of 2 as a threshold to identify high-risk women for GDM, approximately 50% of women would undergo OGTT, and more than 74% of women with GDM could be identified, with a missed diagnosis rate of less than 26%.

Although similar to other risk scores,[12–14] INSPIRE based on Pakistani women has identified many key predictors like maternal age, MUAC, family history of diabetes, a history of GDM and previous poor pregnancy outcomes. It is necessary to determine population-specific risk scores due to the differences in the risk attributes among populations in terms of ethnicity, body composition and other obstetrical factors.[15 16] However, it is important to note that INSPIRE achieved calibration and discrimination similar to or above those based on other populations.[12–14]

INSPIRE has several strengths. To the best of our knowledge, INSPIRE is the first risk score for GDM risk prediction among the Pakistani population. INSPIRE will serve as a screening tool to identify high-risk women and further refer them for diagnostic tests, thus reducing the unnecessary burden of tests for low-risk women. We performed an extensive literature search to identify potential risk factors associated with the development of GDM and included them in our model. We intentionally dichotomised all potential predictors in INSPIRE, facilitating its applicability in routine ANC services. INSPIRE was derived and validated among pregnant women in which the GDM diagnosis was made using a gold standard, that is, a 2-hour, 75 g OGTT.[10] We collected information on the obstetric history from the medical records of the hospital, thus reducing the recall bias.

INSPIRE does have some limitations. Since the study setting was based on hospitals, we found a higher prevalence of GDM among our study population. This could potentially introduce admission bias as the women seeking care might be at higher risk than those who do not seek care or receive care in a community-based setting. However, we included pregnant women from one tertiary and two secondary care hospitals providing ANC services to women from varied socioeconomic statuses, that is, women from high to very low socioeconomic backgrounds; hence, our study findings are generalisable. In addition, we observed a wide 95% CIs for certain predictors, including a history of GDM, previous bad obstetric outcomes and a history of macrosomia due to the low frequency of these outcomes among women without GDM. Nonetheless, this imbalance reflects the expected differences between women with and without GDM. Furthermore, we could not externally validate INSPIRE due to time and resource constraints. Moreover, we did not collect information on the modifiable risk factors of women; however, we believe that using six substantial risk factors in the INSPIRE, based on existing literature, is sufficient to identify high-risk pregnant women at risk of GDM. Since information on obstetrical history and OGTT results was obtained from the hospital medical records, women with incomplete information were excluded, which may introduce selection bias and require careful interpretations of the study findings.

In conclusion, we developed and validated a non-invasive, easy-to-administer risk score (INSPIRE) that enables screening high-risk pregnant women for GDM during the ANC services, thus reducing the unnecessary burden of performing OGTT on low-risk pregnant women. INSPIRE would also facilitate the identification of high-risk pregnant women earlier during the first trimester based on the established risk factors, thus providing targeted intervention and may prevent the development of GDM and, hence, many short-term and long-term complications associated with it. Further research is needed to validate the performance of INSPIRE on the external dataset so that INSPIRE could be implemented earlier during pregnancy within communities to assess risk through female health workers who provide door-to-door services in low-resource communities in Pakistan.

ORCID iDs
Sabahat Naz http://orcid.org/0000-0003-4448-0061
Romaina Iqbal http://orcid.org/0000-0002-5364-4366

REFERENCES

1. Sheikh A, Sheikh L. Changing prevalence of gestational diabetes mellitus during pregnancy over more than a decade. *J Pak Med Assoc* 2020;70:1477–8.
2. Ramos-Leví AM, Pérez-Ferre N, Fernández MD, *et al*. Risk factors for gestational diabetes mellitus in a large population of women living in Spain: implications for preventative strategies. *Int J Endocrinol* 2012;2012:312529.
3. Bener A, Saleh NM, Al-Hamaq A. Prevalence of gestational diabetes and associated maternal and neonatal complications in a fast-developing community: global comparisons. *Int J Womens Health* 2011;3:367–73.
4. Muche AA, Olayemi OO, Gete YK. Effects of gestational diabetes mellitus on risk of adverse maternal outcomes: a prospective cohort study in Northwest Ethiopia. *BMC Pregnancy Childbirth* 2020;20:73.
5. Gadve SS, Chavanda S, Mukherjee AD, *et al*. Risk of developing type 2 diabetes mellitus in South Asian women with history of gestational diabetes mellitus: A systematic review and meta-analysis. *Indian J Endocrinol Metab* 2021;25:176–81.
6. Panel IC. International Association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. *Diabetes Care* 2010;33:676–82.
7. Force UPST. Screening for gestational diabetes: US preventive services task force recommendation statement. *JAMA* 2021;326:531–8.
8. Lachmann EH, Fox RA, Dennison RA, *et al*. Barriers to completing oral glucose tolerance testing in women at risk of gestational diabetes. *Diabet Med* 2020;37:1482–9.
9. Handlos LN, Witte DR, Almdal TP, *et al*. Risk scores for diabetes and impaired Glycaemia in the Middle East and North Africa. *Diabet Med* 2013;30:443–51.
10. Lindstram J, Tuomilehto J. The diabetes risk score A practical tool to predict type 2 diabetes risk. *Diabetes Care* 2003;26:725–31.
11. Harrison CL, Lombard CB, East C, *et al*. Risk stratification in early pregnancy for women at increased risk of gestational diabetes. *Diabetes Res Clin Pract* 2015;107:61–8.
12. Gao S, Leng J, Liu H, *et al*. Development and validation of an early pregnancy risk score for the prediction of gestational diabetes mellitus in Chinese pregnant women. *BMJ Open Diabetes Res Care* 2020;8:e000909.
13. Nombo AP, Mwanri AW, Brouwer-Brolsma EM, *et al*. Gestational diabetes mellitus risk score: A practical tool to predict gestational diabetes mellitus risk in Tanzania. *Diabetes Res Clin Pract* 2018;145:130–7.

14 Caliskan E, Kayikcioglu F, Oztürk N, *et al*. 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.

15 Iqbal R, Rafique G, Badruddin S, *et al*. Increased body fat percentage and physical inactivity are independent predictors of gestational diabetes mellitus in South Asian women. *Eur J Clin Nutr* 2007;61:736–42.

16 Anand SS, Yusuf S, Vuksan V, *et al*. Differences in risk factors, Atherosclerosis, and cardiovascular disease between ethnic groups in Canada: the study of health assessment and risk in ethnic groups (SHARE). *Lancet* 2000;356:279–84.

17 Kumar M, Chen L, Tan K, *et al*. Population-centric risk prediction modeling for gestational diabetes mellitus: A machine learning approach. *Diabetes Res Clin Pract* 2022;185:S0168-8227(22)00049-3.

18 Khan R, Ali K, Khan Z. Sociodemographic risk factors of gestational diabetes mellitus. *Pak J Med Sci* 2013;29:843–6.

19 Ahkter J, Qureshi R, Rahim F, *et al*. Diabetes in pregnancy in Pakistani women: prevalence and complications in an indigenous South Asian community. *Diabet Med* 1996;13:189–91.

20 Teh WT, Teede HJ, Paul E, *et al*. Risk factors for gestational diabetes mellitus: implications for the application of screening guidelines. *Aust N Z J Obstet Gynaecol* 2011;51:26–30.

21 Lee KW, Ching SM, Ramachandran V, *et al*. Prevalence and risk factors of gestational diabetes mellitus in Asia: a systematic review and meta-analysis. *BMC Pregnancy Childbirth* 2018;18:494.

22 Sohail R, Bashir T, Javaid K, *et al*. Association of risk factors of GDM with outcome of GCT in Obstetrical population. *Annals KEMU* 2016;10.

23 Li Y, Ren X, He L, *et al*. Maternal age and the risk of gestational diabetes mellitus: A systematic review and meta-analysis of over 120 million participants. *Diabetes Res Clin Pract* 2020;162:S0168-8227(19)30502-9.

24 Fakier A, Petro G, Fawcus S. Mid-upper arm circumference: A Surrogate for body mass index in pregnant women. *S Afr Med J* 2017;107:606–10.

25 Naylor CD, Sermer M, Chen E, *et al*. Selective screening for gestational diabetes mellitus. *N Engl J Med* 1997;337:1591–6.

26 Kc K, Shakya S, Zhang H. Gestational diabetes mellitus and Macrosomia: A literature review. *Ann Nutr Metab* 2015;66 Suppl 2:14–20.