

A Machine Learning Model for Predicting Recurrent Pregnancy Loss: Retrospective Integration of Routine Serum IL-33, C-Reactive Protein, and Lymphocyte Subset Counts

Qiuxia Liu¹, Liyun Dong²

¹Department of Obstetrics and Gynecology, XiDian Group Hospital, Xi'an, Shaanxi, People's Republic of China; ²Department of Gynecology, Xi'an Traditional Chinese Medicine Hospital, Xi'an, Shaanxi, People's Republic of China

Correspondence: Liyun Dong, Email Dongliyun512@outlook.com

Objective: Recurrent pregnancy loss (RPL), defined as two or more consecutive spontaneous miscarriages before 20 weeks of gestation, affects 2–5% of reproductive-age women globally, and current clinical predictors for it lack sufficient accuracy. This study aimed to construct a machine learning (ML) model for RPL prediction by integrating serum IL-33, C-reactive protein (CRP), and lymphocyte subset counts, and validate its performance in a retrospective cohort.

Methods: A total of 340 reproductive-age women from XiDian Group Hospital and Xi'an Traditional Chinese Medicine Hospital (January 2020–December 2024) were enrolled. Baseline clinical characteristics, IL-33, CRP levels, and lymphocyte subset counts were collected as predictors, with RPL as the primary outcome. The dataset was split into a training set (70%) and a validation set (30%). Logistic regression, random forest, and XGBoost were trained with hyperparameter optimization via grid search, and model performance was evaluated by AUC, accuracy, sensitivity, specificity, PPV, and NPV.

Results: Of the 340 participants, 85 (25.0%) had RPL and 255 (75.0%) did not. The RPL group had significantly lower IL-33 and CD4+/CD8+ ratio, higher CRP and NK cell proportions (all $p < 0.001$). XGBoost outperformed the other two models, with an AUC of 0.89 (95% CI: 0.82–0.96) in the training set and 0.85 (95% CI: 0.76–0.94) in the validation set; its validation set accuracy, sensitivity, specificity, PPV and NPV were 88.1%, 82.4%, 88.7%, 28.6% and 98.7%, respectively.

Conclusion: The ML model integrating IL-33, CRP, and lymphocyte subset counts shows good discriminatory ability for RPL, providing a preliminary reference for identifying high-risk women in clinical practice.

Keywords: recurrent pregnancy loss, machine learning, XGBoost, interleukin-33, C-reactive protein, lymphocyte subsets

Introduction

Recurrent pregnancy loss (RPL) is a common and challenging reproductive disorder, defined by the European Society of Human Reproduction and Embryology (ESHRE) as two or more consecutive spontaneous miscarriages prior to 20 weeks of gestation.^{1,2} With a global incidence of approximately 2–5% among reproductive-age women, RPL not only causes physical harm to patients but also leads to severe psychological distress, including anxiety, depression, and reduced quality of life.^{3–6} Despite extensive research, the etiology of RPL remains unclear in up to 50% of cases, with known contributors including chromosomal abnormalities, uterine anatomical defects, endocrine disorders, immune dysfunction, and environmental factors.⁷

Immunological imbalance is recognized as a key pathological mechanism underlying idiopathic RPL.⁸ The maternal immune system must establish a state of immune tolerance to the semi-allogeneic fetus to maintain normal pregnancy.^{9,10} Disruption of this tolerance can trigger fetal rejection and miscarriage. Interleukin-33 (IL-33), a member of the IL-1 cytokine family, is involved in regulating type 2 immune responses and promoting the differentiation of regulatory T cells

(Tregs), which play a critical role in maternal-fetal immune tolerance.^{11–13} Reduced serum IL-33 levels have been associated with impaired immune tolerance and increased risk of RPL in preliminary studies. C-reactive protein (CRP), a classic acute-phase inflammatory marker, reflects systemic inflammation.^{14,15} Elevated CRP levels indicate chronic low-grade inflammation, which can disrupt endometrial receptivity and placental development, thereby increasing miscarriage risk.^{16,17} Lymphocyte subsets, including T cells (CD3+, CD4+, CD8+), NK cells, and the CD4+/CD8+ ratio, are core indicators of cellular immune function. Abnormal lymphocyte subset distribution such as elevated NK cell proportion and reduced CD4+/CD8+ ratio is closely linked to RPL.^{18,19}

Traditional statistical methods for RPL prediction often fail to capture complex non-linear relationships between multiple biomarkers and clinical outcomes. Machine learning (ML), a subset of artificial intelligence, can handle high-dimensional data and identify hidden patterns, making it increasingly popular in clinical predictive modeling.^{20,21} However, few studies have integrated IL-33, CRP, and lymphocyte subset counts to construct an ML-based RPL prediction model.

This retrospective study aimed to compare the differences in serum IL-33, CRP, and lymphocyte subset counts between RPL and non-RPL patients, develop an ML model for RPL prediction using these biomarkers and baseline clinical characteristics, and validate the performance of the model to provide a clinical tool for early risk stratification of RPL.

Materials and Methods

Study Population

This retrospective cohort study enrolled 340 reproductive-age women (18–45 years old) who attended the Department of Obstetrics and Gynecology at XiDian Group Hospital and the Department of Gynecology at Xi'an Traditional Chinese Medicine Hospital from January 2020 to December 2024. The inclusion criteria were women planning pregnancy or in the first trimester (≤ 12 weeks) of pregnancy, complete clinical records including baseline characteristics and laboratory test results of IL-33, CRP, and lymphocyte subsets, and no history of assisted reproductive technology (ART) treatment.²² The exclusion criteria were ectopic pregnancy or molar pregnancy, uterine anatomical abnormalities such as uterine septum and submucosal fibroids, chromosomal abnormalities in the couple, acute or chronic infections such as toxoplasmosis, cytomegalovirus, and syphilis, autoimmune diseases such as systemic lupus erythematosus and antiphospholipid syndrome, severe endocrine disorders such as uncontrolled diabetes mellitus and hyperthyroidism, alcohol or drug abuse, and incomplete clinical or laboratory data. Patient enrollment and the construction of the predictive model are detailed in [Figure 1](#).

This study complied with the Declaration of Helsinki (2024 revision). The protocol was approved by the Ethics Committee of XiDian Group Hospital (XDJGH-2019-IRB-037) and Xi'an Traditional Chinese Medicine Hospital (XATCMH-2019-IRB-012). Informed consent was waived due to the retrospective nature of the study and the use of de-identified data.

Selection Criteria for Variables

The primary outcome was the diagnosis of RPL, defined as two or more consecutive spontaneous miscarriages confirmed by transvaginal ultrasound and clinical follow-up. The follow-up duration was up to 20 weeks of gestation. Predictor variables were divided into two categories: baseline clinical characteristics and laboratory biomarkers. Baseline clinical characteristics included age (years), body mass index (BMI, kg/m^2), gravidity (number of pregnancies), parity (number of live births), number of previous miscarriages, smoking status (yes/no), alcohol consumption (yes/no), history of thyroid disease (yes/no), and history of polycystic ovary syndrome (PCOS, yes/no).

Laboratory biomarkers included serum IL-33 (pg/mL , measured by enzyme-linked immunosorbent assay [ELISA], R&D Systems, USA), serum CRP (mg/L , measured by immunoturbidimetry, Beckman Coulter, USA), and lymphocyte subset counts (measured by flow cytometry, BD Biosciences, USA) including CD3+ T cells (%), CD4+ T cells (%), CD8+ T cells (%), NK cells (%), and CD4+/CD8+ ratio.

Construction of the Predictive Model

Missing values accounting for $<1\%$ of the total data were imputed using the mean value of the corresponding variable.²³ Continuous variables were normalized to a standard normal distribution with mean = 0 and standard deviation = 1 to

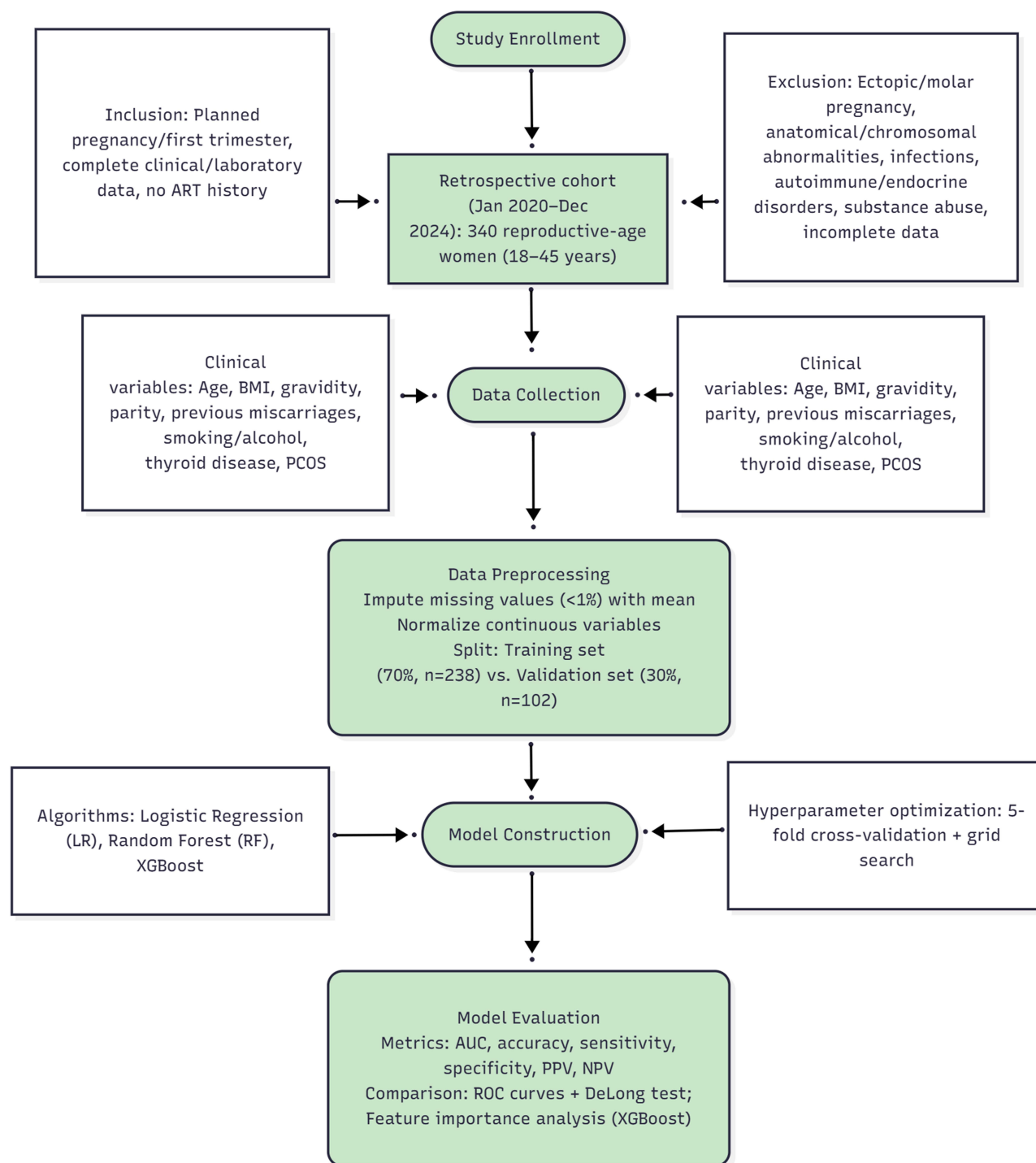


Figure 1 Workflow of a Machine Learning-Based Retrospective Cohort Study for Recurrent Pregnancy Loss Prediction.

eliminate the influence of different measurement scales. Data standardization was only performed on the training set after stratified random sampling of the dataset, and the mean and standard deviation parameters obtained from the training set were applied to the validation set for consistent standardization, which avoided the risk of statistical data leakage. The dataset was randomly split into a training set (70%, $n = 238$) and a validation set (30%, $n = 102$) using a stratified random sampling method to maintain the proportion of RPL cases in both sets.

Algorithm Selection and Hyperparameter Optimization

Three commonly used ML algorithms were selected for model construction. Logistic regression (LR) is a linear model for binary classification and was used as a reference model.²⁴ Random forest (RF) is an ensemble learning algorithm based on decision trees, which reduces overfitting by integrating multiple decision trees.²⁵ Extreme gradient boosting (XGBoost) is a gradient boosting algorithm with regularized terms, which optimizes model performance by iteratively correcting prediction errors.²⁶

Hyperparameter optimization was performed using 5-fold cross-validation on the training set with grid search.²⁷ Key hyperparameters for RF included the number of decision trees (ntree) ranging from 50 to 500 with a step of 50 and the number of features selected for each split (mtry) ranging from 3 to 10 with a step of 1. Key hyperparameters for XGBoost included learning rate ranging from 0.01 to 0.3 with a step of 0.01, maximum depth of trees ranging from 3 to 10 with a step of 1, and number of estimators ranging from 50 to 500 with a step of 50. The optimal hyperparameters were determined based on the highest AUC value in cross-validation.

Predictive Model Evaluation

Model performance was evaluated using multiple metrics. Area under the receiver operating characteristic curve (AUC) reflects the overall discriminatory ability of the model, with AUC > 0.8 indicating good performance. Accuracy is the proportion of correctly predicted cases among all cases. Sensitivity (recall) is the proportion of true RPL cases correctly identified by the model. Specificity is the proportion of true non-RPL cases correctly identified by the model. Positive predictive value (PPV) is the proportion of predicted RPL cases that are true RPL cases. Negative predictive value (NPV) is the proportion of predicted non-RPL cases that are true non-RPL cases.

ROC curves were plotted for each model in the training and validation sets, and the DeLong test was used to compare AUC values between different models.

Statistical Methods

All statistical analyses and ML model construction were performed using R software. Continuous variables with normal distribution were presented as mean \pm standard deviation (SD), and comparisons between the RPL and non-RPL groups were performed using independent samples t-tests. Categorical variables were presented as frequencies, and comparisons between groups were performed using the chi-square test or Fisher's exact test when the expected frequency < 5. A two-tailed p value < 0.05 was considered statistically significant.

Results

Baseline Characteristics of the Study Population

As shown in Table 1, a total of 340 participants were included in the final analysis, of whom 85 (25.0%) were diagnosed with recurrent pregnancy loss (RPL, RPL group) and 255 (75.0%) were non-RPL cases (non-RPL group). The median age of the RPL group was significantly higher than that of the non-RPL group [33.0 (30.0–36.0) years vs 30.0 (27.0–33.0) years, $p < 0.001$]. Additionally, the RPL group exhibited a significantly greater median number of previous miscarriages [2 (1–3) vs 1 (0–1), $p < 0.001$] and a higher prevalence of thyroid disease (23.5% vs 6.2%, $p < 0.001$) compared with the non-RPL group. No statistically significant differences were noted between the two groups in terms of body mass index (BMI), gravidity, parity, smoking status, alcohol consumption, or history of polycystic ovary syndrome (PCOS) (all $p > 0.05$). Collectively, these results suggest that advanced age, an increased number of prior miscarriages, and a history of thyroid disease are associated with RPL in this study cohort. In contrast, common metabolic factors (eg, BMI) and lifestyle-related variables (eg, smoking, alcohol consumption) did not differ significantly between women with and without RPL.

Distribution of Serum Biomarkers and Lymphocyte Subsets

As shown in Table 2 and Figure 2, the RPL group had significantly lower serum IL-33 levels [15.80 (11.83, 18.00) pg/mL vs 28.00 (24.15, 32.05) pg/mL, $p < 0.001$] and significantly higher CRP levels [8.90 (7.05, 10.30) mg/L vs 3.10 (2.40,

Table 1 Baseline Characteristics of Participants Stratified by Recurrent Pregnancy Loss Status

Variables	Training Cohort			P-value	Testing Cohort			P-value
	Overall (n=237)	RPL (n=62)	Non-RPL (n=175)		Overall (n=103)	RPL (n=23)	Non-RPL (n=80)	
Age (years) (median [IQR])	31.00 [28.00, 34.00]	35.00 [30.00, 37.00]	30.00 [27.50, 32.00]	<0.001	31.00 [28.00, 33.00]	33.00 [31.00, 36.50]	30.00 [27.00, 32.00]	<0.001
BMI (kg/m ²) (median [IQR])	24.10 [22.50, 26.00]	24.85 [22.75, 26.00]	23.90 [22.45, 25.90]	0.362	24.20 [22.20, 25.95]	23.90 [22.50, 25.15]	24.35 [21.77, 26.40]	0.299
Gravidity Count (%)								
1	70 (29.5)	19 (30.6)	51 (29.1)	0.114	40 (38.8)	9 (39.1)	31 (38.8)	0.829
2	85 (35.9)	23 (37.1)	62 (35.4)		39 (37.9)	10 (43.5)	29 (36.2)	
3	65 (27.4)	12 (19.4)	53 (30.3)		20 (19.4)	3 (13.0)	17 (21.2)	
4	16 (6.8)	7 (11.3)	9 (5.1)		4 (3.9)	1 (4.3)	3 (3.8)	
5	1 (0.4)	1 (1.6)	0 (0.0)		0 (0.0)	0 (0.0)	0 (0.0)	
Parity Count (%)								
0	129 (54.4)	39 (62.9)	90 (51.4)	0.21	44 (42.7)	10 (43.5)	34 (42.5)	0.991
1	91 (38.4)	18 (29.0)	73 (41.7)		46 (44.7)	10 (43.5)	36 (45.0)	
2	17 (7.2)	5 (8.1)	12 (6.9)		13 (12.6)	3 (13.0)	10 (12.5)	
Previous Miscarriages (%)								
0	50 (21.1)	1 (1.6)	49 (28.0)	<0.001	26 (25.2)	1 (4.3)	25 (31.2)	<0.001
1	120 (50.6)	6 (9.7)	114 (65.1)		51 (49.5)	3 (13.0)	48 (60.0)	
2	49 (20.7)	37 (59.7)	12 (6.9)		18 (17.5)	11 (47.8)	7 (8.8)	
3	17 (7.2)	17 (27.4)	0 (0.0)		8 (7.8)	8 (34.8)	0 (0.0)	
4	1 (0.4)	1 (1.6)	0 (0.0)		0 (0.0)	0 (0.0)	0 (0.0)	
Smoking Status (%)								
Non-smoker	218 (92.0)	58 (93.5)	160 (91.4)	0.798	91 (88.3)	22 (95.7)	69 (86.2)	0.384
Smoker	19 (8.0)	4 (6.5)	15 (8.6)		12 (11.7)	1 (4.3)	11 (13.8)	
Alcohol Status (%)								
Non-drinker	224 (94.5)	59 (95.2)	165 (94.3)	1	101 (98.1)	23 (100.0)	78 (97.5)	1
Drinker	13 (5.5)	3 (4.8)	10 (5.7)		2 (1.9)	0 (0.0)	2 (2.5)	
Thyroid Disease (%)								
No thyroid disease	210 (88.6)	47 (75.8)	163 (93.1)	0.001	95 (92.2)	19 (82.6)	76 (95.0)	0.13
With thyroid disease	27 (11.4)	15 (24.2)	12 (6.9)		8 (7.8)	4 (17.4)	4 (5.0)	
PCOS Status (%)								
No PCOS	218 (92.0)	57 (91.9)	161 (92.0)	1	95 (92.2)	21 (91.3)	74 (92.5)	1
With PCOS	19 (8.0)	5 (8.1)	14 (8.0)		8 (7.8)	2 (8.7)	6 (7.5)	
Serum IL-33 (pg/mL) (median [IQR])	25.60 [18.80, 30.60]	15.80 [11.83, 18.00]	28.00 [24.15, 32.05]	<0.001	27.80 [20.60, 33.15]	14.90 [13.80, 17.65]	30.00 [26.08, 34.42]	<0.001
Serum CRP (mg/L) (median [IQR])	3.50 [2.60, 5.40]	8.90 [7.05, 10.30]	3.10 [2.40, 3.70]	<0.001	3.70 [2.75, 5.15]	8.50 [7.45, 10.90]	3.20 [2.48, 4.23]	<0.001
CD3+ T cells (%) (median [IQR])	72.00 [68.60, 75.20]	67.85 [64.43, 71.68]	73.00 [70.20, 76.20]	<0.001	71.50 [68.90, 74.65]	67.40 [63.95, 70.95]	72.45 [70.20, 76.35]	<0.001
CD4+ T cells (%) (median [IQR])	37.00 [33.10, 40.30]	31.70 [29.83, 34.15]	38.60 [35.35, 42.50]	<0.001	37.90 [34.00, 40.55]	31.50 [29.00, 36.25]	38.60 [35.62, 41.32]	<0.001
CD8+ T cells (%) (median [IQR])	22.70 [19.30, 26.70]	28.05 [26.60, 30.15]	21.20 [18.15, 24.35]	<0.001	22.20 [20.05, 25.65]	28.00 [26.15, 31.35]	21.15 [19.00, 23.55]	<0.001
CD4+/CD8+ ratio (median [IQR])	1.60 [1.33, 1.90]	1.23 [1.08, 1.40]	1.75 [1.49, 2.00]	<0.001	1.69 [1.29, 1.97]	1.21 [0.98, 1.38]	1.87 [1.54, 2.01]	<0.001
NK cells (%) (median [IQR])	12.70 [10.30, 16.80]	18.50 [16.80, 20.17]	11.70 [9.55, 13.70]	<0.001	12.30 [9.55, 15.30]	18.10 [15.60, 22.85]	11.50 [9.10, 13.50]	<0.001

Abbreviations: BMI, Body mass index; CRP, C-reactive protein; IQR, Interquartile range; PCOS, Polycystic ovary syndrome; RPL, Recurrent pregnancy loss.

Table 2 Distribution of Serum Biomarkers and Lymphocyte Subsets

Variables	Training Cohort			P-value	Testing Cohort			P-value
	Overall (n=237)	RPL (n=62)	Non-RPL (n=175)		Overall (n=103)	RPL (n=23)	Non-RPL (n=80)	
Serum IL-33 (pg/mL) (median [IQR])	25.60 [18.80, 30.60]	15.80 [11.83, 18.00]	28.00 [24.15, 32.05]	<0.001	27.80 [20.60, 33.15]	14.90 [13.80, 17.65]	30.00 [26.08, 34.42]	<0.001
Serum CRP (mg/L) (median [IQR])	3.50 [2.60, 5.40]	8.90 [7.05, 10.30]	3.10 [2.40, 3.70]	<0.001	3.70 [2.75, 5.15]	8.50 [7.45, 10.90]	3.20 [2.48, 4.23]	<0.001
CD3+ T cells (%) (median [IQR])	72.00 [68.60, 75.20]	67.85 [64.43, 71.68]	73.00 [70.20, 76.20]	<0.001	71.50 [68.90, 74.65]	67.40 [63.95, 70.95]	72.45 [70.20, 76.35]	<0.001
CD4+ T cells (%) (median [IQR])	37.00 [33.10, 40.30]	31.70 [29.83, 34.15]	38.60 [35.35, 42.50]	<0.001	37.90 [34.00, 40.55]	31.50 [29.00, 36.25]	38.60 [35.62, 41.32]	<0.001
CD8+ T cells (%) (median [IQR])	22.70 [19.30, 26.70]	28.05 [26.60, 30.15]	21.20 [18.15, 24.35]	<0.001	22.20 [20.05, 25.65]	28.00 [26.15, 31.35]	21.15 [19.00, 23.55]	<0.001
CD4+/CD8+ ratio (median [IQR])	1.60 [1.33, 1.90]	1.23 [1.08, 1.40]	1.75 [1.49, 2.00]	<0.001	1.69 [1.29, 1.97]	1.21 [0.98, 1.38]	1.87 [1.54, 2.01]	<0.001
NK cells (%) (median [IQR])	12.70 [10.30, 16.80]	18.50 [16.80, 20.17]	11.70 [9.55, 13.70]	<0.001	12.30 [9.55, 15.30]	18.10 [15.60, 22.85]	11.50 [9.10, 13.50]	<0.001

Abbreviations: IQR, Interquartile range; IL-33, Interleukin-33; NK, Natural killer; RPL, Recurrent pregnancy loss.

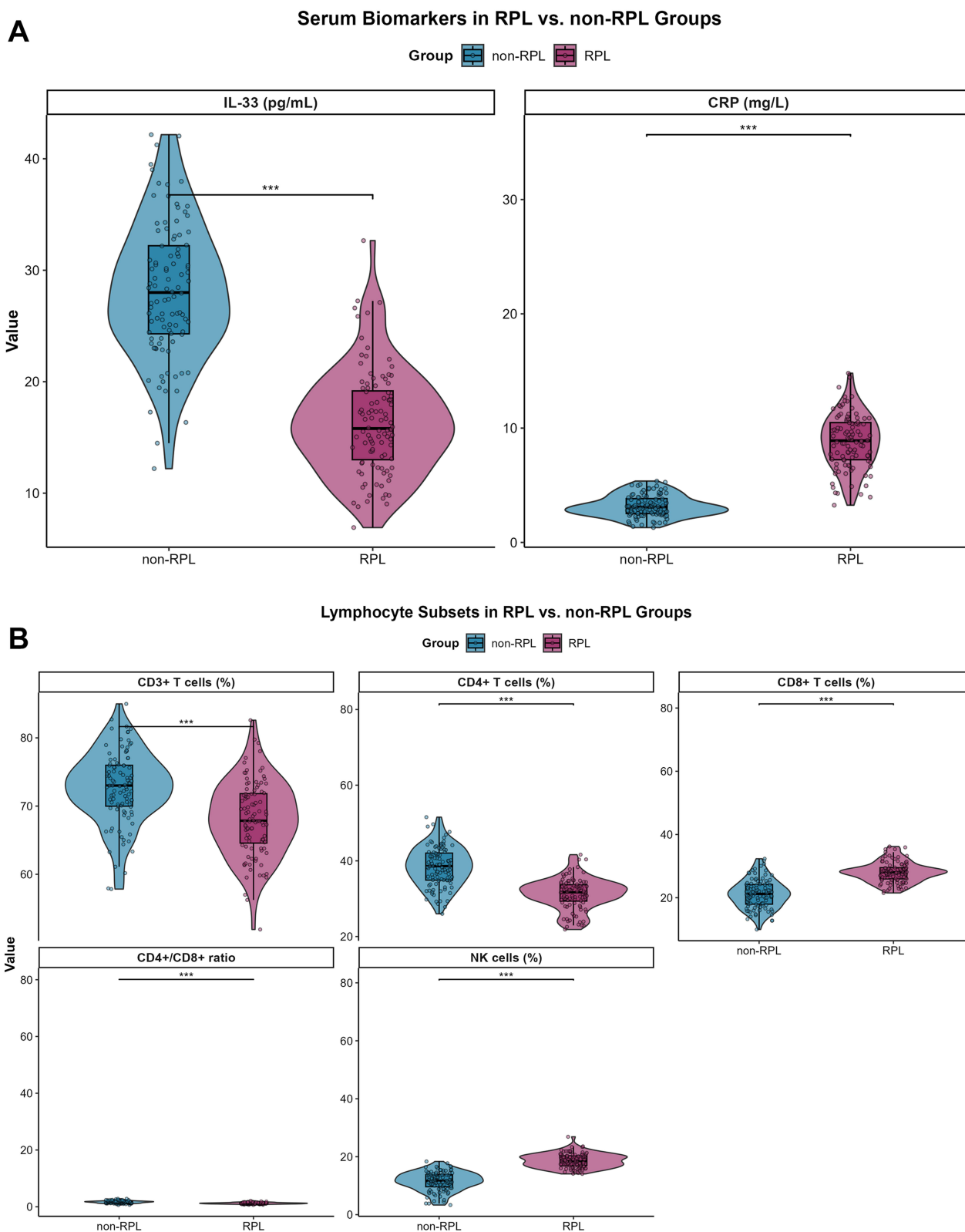


Figure 2 Alterations in Serum Biomarker Levels and Lymphocyte Subset Distributions in Women with Recurrent Pregnancy Loss (RPL) Versus Non-RPL Controls.

Notes: **(A)** This panel visualizes the distribution of two serum biomarkers (IL-33 and CRP) between the RPL and non-RPL groups, using violin-box plots (violins indicate data density; boxes represent median/quartiles; points denote individual samples). Results show that the RPL group has significantly lower serum IL-33 levels and significantly higher C-reactive protein (CRP) levels compared to the non-RPL group (all comparisons: $p < 0.001$, marked by "***"). **(B)** This panel displays the distribution of peripheral blood lymphocyte subsets (CD3+ T cells, CD4+ T cells, CD8+ T cells, CD4+/CD8+ ratio, and NK cells) in the two groups. Relative to the non-RPL group, the RPL group exhibits significantly reduced proportions of CD3+ T cells and CD4+ T cells, a decreased CD4+/CD8+ ratio, and elevated proportions of CD8+ T cells and natural killer (NK) cells (all comparisons: $p < 0.001$, marked by "***").

3.70) mg/L, $p < 0.001$] than the non-RPL group. For lymphocyte subsets, the RPL group exhibited significantly lower CD3+ T cell proportions [67.85 (64.43, 71.68)% vs 73.00 (70.20, 76.20)%, $p < 0.001$], lower CD4+ T cell proportions [31.70 (29.83, 34.15)% vs 38.60 (35.35, 42.50)%, $p < 0.001$], higher CD8+ T cell proportions [28.05 (26.60, 30.15)% vs 21.20 (18.15, 24.35)%, $p < 0.001$], lower CD4+/CD8+ ratios [1.23 (1.08, 1.40) vs 1.75 (1.49, 2.00), $p < 0.001$], and higher NK cell proportions [18.50 (16.80, 20.17)% vs 11.70 (9.55, 13.70)%, $p < 0.001$] compared with the non-RPL group. Collectively, these results demonstrate profound alterations in inflammatory biomarkers and lymphocyte subset distribution in women with RPL, suggesting a potential role of immune dysregulation and enhanced systemic inflammation in the pathogenesis of RPL.

Performance of the Machine Learning Model in Training and Validation Sets

In the training set, the XGBoost model achieved the highest AUC (0.89, 95% CI: 0.82–0.96), followed by RF (0.85, 95% CI: 0.77–0.93) and LR (0.78, 95% CI: 0.69–0.87). The AUC of XGBoost was significantly higher than that of LR but not significantly different from that of RF (Figure 3A). In the validation set, the XGBoost model still had the highest AUC (0.85, 95% CI: 0.76–0.94), with an accuracy of 88.1%, sensitivity of 82.4%, specificity of 88.7%, PPV of 28.6%, and NPV of 98.7%. The RF model had an AUC of 0.80 (95% CI: 0.70–0.90) in the validation set, while the LR model had the lowest AUC (0.75, 95% CI: 0.64–0.86). The XGBoost curve in the ROC curves of the three models in the validation set was above RF and LR curves, indicating better discriminatory ability (Figure 3B and C). These data confirm that the XGBoost model outperforms RF and LR models in predicting RPL, with robust performance in both training and independent validation sets.

Comparison of Different Machine Learning Algorithms

The DeLong test showed that the AUC of XGBoost was significantly higher than that of LR and marginally higher than that of RF in the validation set (Figure 4A). Feature importance analysis of the XGBoost model showed that the number of previous miscarriages, serum IL-33 levels, CD4+/CD8+ ratio, and CRP levels were the top four most important predictors of RPL. The importance score of previous miscarriage number is 0.28, IL-33 is 0.24, CD4+/CD8+ ratio is 0.21, and CRP is 0.18 (Figure 4B). These results highlight the superior discriminatory power of the XGBoost algorithm for RPL prediction and identify prior miscarriage history, IL-33, CD4+/CD8+ ratio, and CRP as key clinical and immunological predictors worthy of further investigation.

Discussion

This retrospective study constructed an ML model for RPL prediction by integrating baseline clinical characteristics, serum IL-33, CRP, and lymphocyte subset counts, and validated its performance in a cohort of 340 reproductive-age women. The key findings provide new insights into both the pathophysiological mechanisms of RPL and the clinical application of ML in reproductive medicine. Specifically, RPL patients exhibit distinct immunological and inflammatory profiles characterized by reduced IL-33, elevated CRP, and abnormal lymphocyte subset distribution. The XGBoost model incorporating these indicators outperforms traditional linear models and other ensemble algorithms, demonstrating potential as a clinical risk stratification tool. Moreover, the number of previous miscarriages, IL-33, CD4+/CD8+ ratio, and CRP emerge as the most impactful predictive factors, highlighting the central role of immune tolerance and inflammation in RPL pathogenesis.

The distinct biomarker patterns observed in RPL patients align with the well-established role of immune imbalance and chronic inflammation in disrupting maternal-fetal tolerance. IL-33, a damage-associated molecular pattern molecule expressed in endometrial stromal cells and placental trophoblasts, exerts dual effects on maternal immune regulation. It not only promotes the differentiation of naive T cells into type 2 helper T cells (Th2) but also enhances the proliferation and functional activity of Tregs, which are critical for suppressing anti-fetal immune responses.^{28,29} The significantly lower IL-33 levels in the RPL group observed in this study suggest a compromised ability to establish Th2-dominant and Treg-mediated immune tolerance. This is consistent with preclinical studies showing that IL-33 knockout mice exhibit higher rates of fetal resorption due to impaired Treg infiltration in the decidua.²⁹ Clinically, this finding supports the

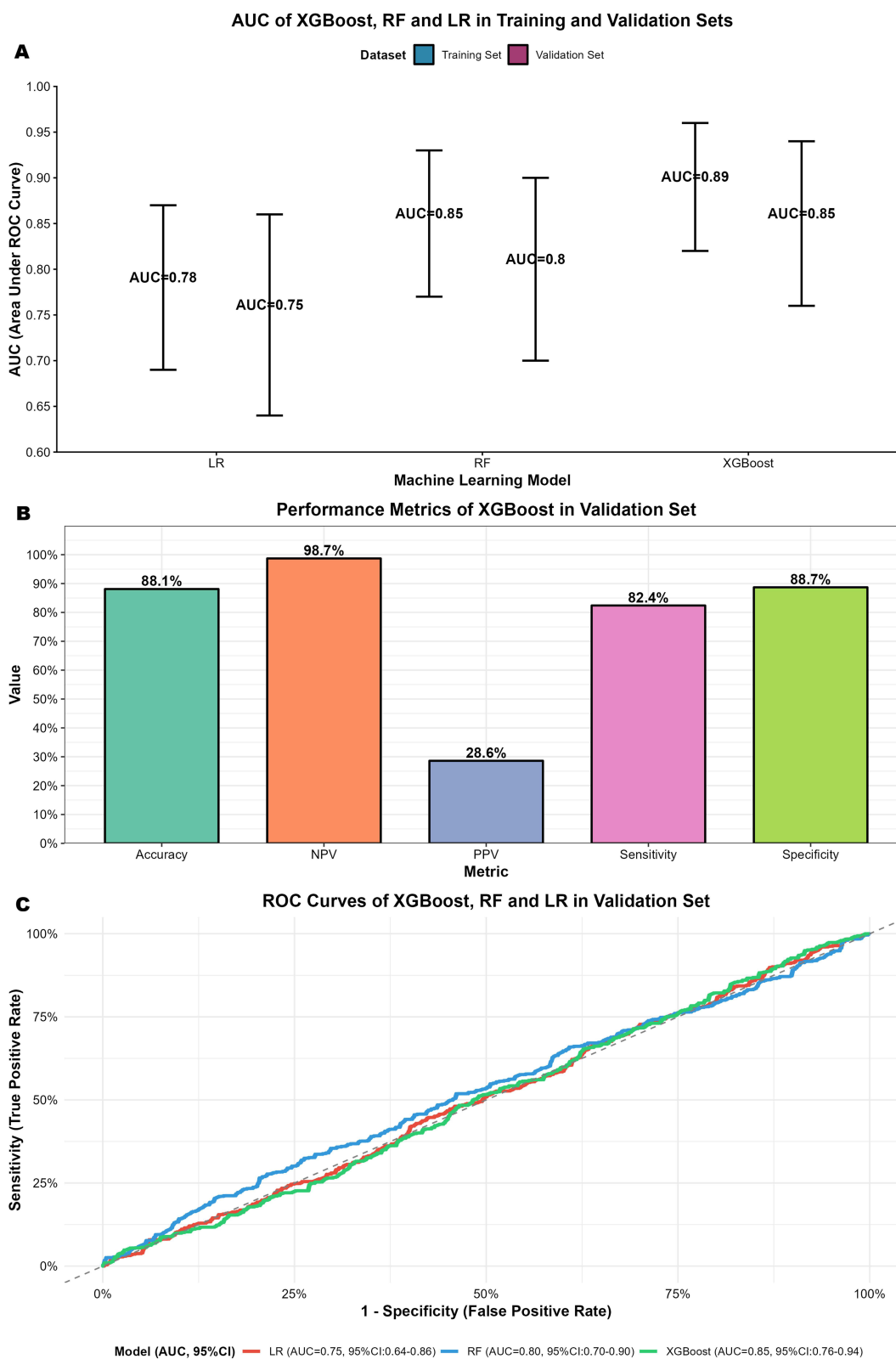


Figure 3 Performance of XGBoost, RF, and LR Models for RPL Prediction in Training and Validation Sets.

Notes: (A) Compares the area under the receiver operating characteristic curve (AUC) (with 95% confidence intervals) of XGBoost, RF, and LR in the training and validation sets; (B) Displays key performance metrics (eg, accuracy, negative predictive value [NPV]) of the XGBoost model in the validation set using a bar plot; (C) Presents the receiver operating characteristic (ROC) curves of the three models in the validation set, which reflect their discriminatory ability.

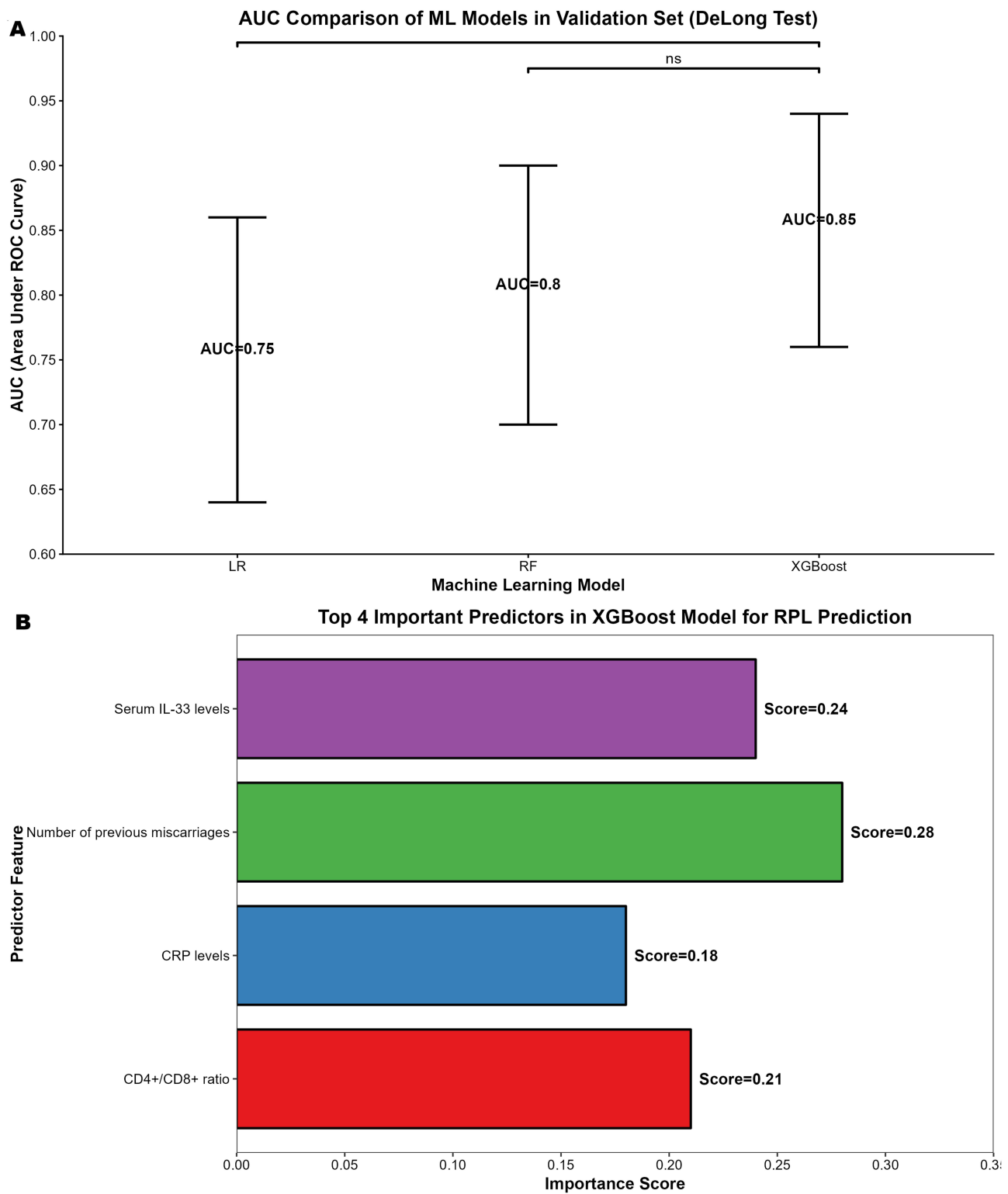


Figure 4 Performance of Machine Learning Models for Recurrent Pregnancy Loss (RPL) Prediction: AUC Comparison and XGBoost Feature Importance. **Notes:** (A) Compares the area under the receiver operating characteristic curve (AUC) of XGBoost, random forest (RF), and logistic regression (LR) in the validation set. Error bars represent 95% confidence intervals, and the “ns” annotation indicates a non-significant difference (evaluated via the DeLong test). (B) Presents the importance scores of the top 4 predictors in the XGBoost model for RPL prediction, displayed as a horizontal bar plot.

potential of IL-33 as both a diagnostic marker and a therapeutic target, as exogenous IL-33 administration has been shown to improve pregnancy outcomes in animal models of immune-mediated abortion.

Elevated CRP levels in the RPL group reflect the presence of chronic low-grade inflammation, which disrupts multiple processes essential for successful pregnancy.³⁰ Chronic inflammation can impair endometrial receptivity by altering the expression of adhesion molecules such as integrin $\alpha\text{v}\beta 3$ and L-selectin, thereby hindering blastocyst implantation.^{31,32} Additionally, pro-inflammatory cytokines induced by elevated CRP, such as tumor necrosis factor- α and interleukin-6, can promote trophoblast apoptosis and inhibit placental angiogenesis, leading to placental insufficiency and subsequent miscarriage.^{33,34} Notably, the CRP levels in RPL patients (8.9 ± 2.7 mg/L) in this study fall within the range of low-grade inflammation, which is often overlooked in clinical practice but has been consistently linked to idiopathic RPL in meta-analyses.^{35,36} This underscores the value of incorporating routine inflammatory markers into RPL risk assessment, even in the absence of overt infection or autoimmune disease.

The abnormal lymphocyte subset distribution in RPL patients further confirms the disruption of cellular immune balance. The reduced CD4+/CD8+ ratio, resulting from lower CD4+ T cell and higher CD8+ T cell proportions, indicates a shift toward a pro-inflammatory Th1-dominant immune profile. CD4+ T cells, particularly Th2 and Treg subsets, are crucial for maintaining maternal-fetal tolerance, while CD8+ cytotoxic T cells can directly target fetal trophoblast cells expressing paternal antigens. The elevated NK cell proportion in the RPL group exacerbates this immune imbalance, as uterine NK cells, when overactivated, secrete cytotoxic molecules such as perforin and granzyme that damage trophoblasts.³⁷ Importantly, the combination of these lymphocyte subset abnormalities with reduced IL-33 and elevated CRP creates a synergistic pro-rejection microenvironment, which may explain the high specificity of the combined model in identifying RPL risk.

The findings of this study are consistent with previous research on individual biomarkers but advance the field by integrating multiple indicators into a unified ML framework. For example, a case-control study involving 66 Egyptian patients with recurrent pregnancy loss (RPL) and 66 matched healthy non-pregnant controls reported that serum interleukin-33 (IL-33) levels were significantly higher in RPL patients compared with the control group. Additionally, in the RPL cohort, serum IL-33 levels showed a positive correlation with both age and the number of miscarriages, while exhibiting a negative correlation with the number of deliveries, suggesting that IL-33 may serve as a predictive biomarker for RPL in early pregnancy.¹² Our study extends this by showing that IL-33, when combined with other biomarkers, contributes to a much higher overall predictive accuracy (88.1% in the validation set). Similarly, A case-control study comparing recurrent pregnancy loss (RPL) patients with healthy controls reported that serum CRP, IL-1 β , and IL-4 levels were significantly elevated in RPL patients, while TGF- β and ferritin levels were decreased ($p < 0.05$), with no significant differences in IL-10, IL-6, or TNF- α levels between the groups indicating immune dysregulation and potential oxidative stress in RPL pathogenesis.³⁸ Our model overcomes this limitation by leveraging the complementary information from immune and inflammatory markers.

In terms of ML applications in RPL prediction, existing studies have primarily focused on either clinical characteristics or a single class of biomarkers. For instance, A retrospective study of 1153 RPL patients using multivariate logistic regression to construct a nomogram (incorporating age, number of previous pregnancy losses, autoantibodies, coagulation indicators, and uterine artery hemodynamic parameters) achieved an AUC of 0.808 in the training cohort and 0.731 in the validation cohort, outperforming the model with only age and previous pregnancy loss count for predicting subsequent pregnancy loss.³⁹ Our study, by integrating clinical, inflammatory, and immune markers, achieves a higher AUC of 0.85 in the validation set, demonstrating the value of multi-dimensional data integration. This is consistent with the broader trend in medical AI, where combining heterogeneous data types improves predictive performance by capturing complex disease mechanisms that cannot be fully represented by individual data modalities.

The XGBoost model developed in this study offers several key advantages for clinical practice. First, its high sensitivity (82.4%) and NPV (98.7%) make it particularly useful for ruling out low-risk individuals. Women with negative predictions by the model can avoid unnecessary invasive examinations such as hysteroscopy, karyotyping, and autoimmune antibody testing, which not only reduce healthcare costs but also alleviate the psychological stress associated with excessive medical intervention. Second, the model uses routine clinical tests that are widely available in primary and secondary hospitals, ensuring high accessibility. IL-33, although not yet a standard RPL test, can be measured using commercial ELISA kits with good reproducibility, and lymphocyte subset analysis is a common immunological test in gynecological practice.

Third, the feature importance ranking provides actionable clinical insights. The number of previous miscarriages, as the most important predictor, reinforces current clinical guidelines that recommend more intensive evaluation for women with two or more miscarriages. The high importance of IL-33 suggests that measuring this cytokine may be particularly valuable for women with idiopathic RPL, where traditional tests fail to identify a cause. Clinicians could use the model's output to prioritize interventions: for example, women with low IL-33 levels might benefit from immunomodulatory therapies such as low-dose aspirin or corticosteroids, while those with high CRP could be advised on lifestyle modifications to reduce inflammation, such as weight loss and dietary changes.

The relatively low PPV (28.6%) of the model, while a limitation, is largely a reflection of the low incidence of RPL (5.0%) in the study population, a phenomenon known as the base rate effect. In clinical practice, this means that a positive prediction should be interpreted in conjunction with other clinical information, such as uterine ultrasound findings and endocrine profiles. However, even with this limitation, the model's ability to identify high-risk individuals who would otherwise be missed by conventional screening remains valuable, as early intervention in these cases can significantly improve pregnancy outcomes.

This study has key limitations that require attention in future research. First, this study has a relatively small positive sample size (85 RPL cases), which may affect the stability and generalizability of the predictive model. The small sample size may lead to insufficient training of the model on rare feature combinations, and the model performance may be overestimated in the internal validation. Future multicenter prospective studies with a larger number of RPL cases are needed to further validate the model. This is illustrated by the 23.5% incidence of thyroid disease in the RPL group, which exceeds the 15–20% reported in multi-center studies. Future research should adopt large-scale multi-center prospective cohorts with diverse populations, expand the number of RPL cases for rigorous validation and hyperparameter optimization, and incorporate additional validated RPL-related variables such as couple HLA compatibility and microbiome data. Second, missing data on therapeutic interventions during follow-up (eg, progesterone, immunomodulatory drugs) may have affected pregnancy outcomes and model accuracy. Additionally, the 20-week gestation follow-up period means the model's predictive value for adverse outcomes like late miscarriage (12–20 weeks) or preeclampsia remains unclear. Subsequent studies should include treatment variables as confounding factors and extend follow-up to cover the entire pregnancy and postpartum period. Third, the current R-based model is not easily accessible to clinicians without programming experience. Developing user-friendly web or mobile applications based on this model will facilitate its clinical popularization.

Conclusion

In conclusion, the XGBoost model integrating baseline clinical characteristics, serum IL-33, CRP, and lymphocyte subset counts shows good discriminatory ability in predicting recurrent pregnancy loss, with high sensitivity and negative predictive value. The number of previous miscarriages, serum IL-33 levels, CD4+/CD8+ ratio, and CRP levels are the key predictive factors of RPL, which suggests the potential association of immune tolerance and chronic inflammation with RPL pathogenesis. This model is an exploratory research result, and its clinical application is limited by low positive predictive value, small positive sample size and single-center retrospective design. Future multicenter prospective studies with larger sample sizes and extended follow-up are needed to validate and refine the model.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Yu N, Kwak-Kim J, Bao S. Unexplained recurrent pregnancy loss: novel causes and advanced treatment. *J Reprod Immunol.* 2023;155:103785. doi:10.1016/j.jri.2022.103785
2. Dimitriadis E, Menkhorst E, Saito S, Kutteh WH, Brosens JJ. Recurrent pregnancy loss. *Nat Rev Dis Primers.* 2020;6(1):98. doi:10.1038/s41572-020-00228-z
3. Quenby S, Gallos ID, Dhillon-Smith RK, et al. Miscarriage matters: the epidemiological, physical, psychological, and economic costs of early pregnancy loss. *Lancet.* 2021;397(10285):1658–1667. doi:10.1016/S0140-6736(21)00682-6
4. Lei D, Zhang XY, Zheng PS. Recurrent pregnancy loss: fewer chromosomal abnormalities in products of conception? A meta-analysis. *J Assist Reprod Genet.* 2022;39(3):559–572. doi:10.1007/s10815-022-02414-2
5. Turocy JM, Rackow BW. Uterine factor in recurrent pregnancy loss. *Semin Perinatol.* 2019;43(2):74–79. doi:10.1053/j.semperi.2018.12.003

6. Amrane S, McConnell R. Endocrine causes of recurrent pregnancy loss. *Semin Perinatol.* 2019;43(2):80–83. doi:10.1053/j.semperi.2018.12.004
7. Giouleka S, Tsakiridis I, Arsenaki E, et al. Investigation and management of recurrent pregnancy loss: a comprehensive review of guidelines. *Obstetrical Gynecol Surv.* 2023;78(5):287–301. doi:10.1097/OGX.0000000000001133
8. Carp H. Immunotherapy for recurrent pregnancy loss. *Best Pract Res Clin Obstet Gynaecol.* 2019;60:77–86. doi:10.1016/j.bpobgyn.2019.07.005
9. Alecsandru D, Klimczak AM, Garcia Velasco JA, Pirtea P, Franasiak JM. Immunologic causes and thrombophilia in recurrent pregnancy loss. *Fertil Sterility.* 2021;115(3):561–566. doi:10.1016/j.fertnstert.2021.01.017
10. Cai R, Yang Q, Liao Y, Qin L, Han J, Gao R. Immune treatment strategies in unexplained recurrent pregnancy loss. *Am J Reprod Immunol.* 2025;93(2):e70060. doi:10.1111/aji.70060
11. Sheng YR, Hu WT, Shen HH, et al. An imbalance of the IL-33/ST2-AXL-efferocytosis axis induces pregnancy loss through metabolic reprogramming of decidual macrophages. *Cell Mol Life Sci.* 2022;79(3):173. doi:10.1007/s00018-022-04197-2
12. Salem L, Eltaieb E, Abdelmaksoud MF. Association of interleukin-33 with recurrent pregnancy loss in Egyptian women. *Eur Cytokine Netw.* 2022;33(2):23–42. doi:10.1684/een.2022.0478
13. Valero-Pacheco N, Tang EK, Massri N, et al. Maternal IL-33 critically regulates tissue remodeling and type 2 immune responses in the uterus during early pregnancy in mice. *Proc Natl Acad Sci USA.* 2022;119(35):e2123267119. doi:10.1073/pnas.2123267119
14. Levinson T, Wasserman A. C-Reactive Protein Velocity (CRPv) as a new biomarker for the early detection of acute infection/inflammation. *Int J Mol Sci.* 2022;23(15):8100. doi:10.3390/ijms23158100
15. Zhou HH, Tang YL, Xu TH, Cheng B. C-reactive protein: structure, function, regulation, and role in clinical diseases. *Front Immunol.* 2024;15:1425168. doi:10.3389/fimmu.2024.1425168
16. Weghofer A, Barad DH, Darmon SK, Kushnir VA, Albertini DF, Gleicher N. Euploid miscarriage is associated with elevated serum C-reactive protein levels in infertile women: a pilot study. *Arch Gynecol Obstetrics.* 2020;301(3):831–836. doi:10.1007/s00404-020-05461-1
17. Lotfy AM, Taha WS, Abdelmoaty MA. Evaluation of serum level of C-reactive protein (CRP) and its correlation with fetal ultrasound parameters in the prediction of threatened miscarriage in the first trimester. *Qatar Med J.* 2024;2024(1):9. doi:10.5339/qmj.2024.9
18. Braun AS, Vomstein K, Reiser E, et al. NK and T cell subtypes in the endometrium of patients with recurrent pregnancy loss and recurrent implantation failure: implications for pregnancy success. *J Clin Med.* 2023;12(17):5585. doi:10.3390/jcm12175585
19. Wang F, Jia W, Fan M, et al. Single-cell immune landscape of human recurrent miscarriage. *Genomics Proteomics Bioinformatics.* 2021;19(2):208–222. doi:10.1016/j.gpb.2020.11.002
20. Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. *New Engl J Med.* 2023;388(13):1201–1208. doi:10.1056/NEJMra2302038
21. MacEachern SJ, Forkert ND. Machine learning for precision medicine. *Genome.* 2021;64(4):416–425. doi:10.1139/gen-2020-0131
22. Practice Committee of the American Society for Reproductive Medicine. Definitions of infertility and recurrent pregnancy loss: a committee opinion. *Fertil Sterility.* 2020;113(3):533–535. doi:10.1016/j.fertnstert.2019.11.025
23. Austin PC, White IR, Lee DS, van Buuren S. Missing data in clinical research: a tutorial on multiple imputation. *Can J Cardiol.* 2021;37(9):1322–1331. doi:10.1016/j.cjca.2020.11.010
24. Schober P, Vetter TR. Logistic regression in medical research. *Anesthesia Analg.* 2021;132(2):365–366. doi:10.1213/ANE.0000000000005247
25. Hu J, Szymczak S. A review on longitudinal data analysis with random forest. *Briefings Bioinf.* 2023;24(2). doi:10.1093/bib/bbad002
26. Koh J. Gradient boosting with extreme-value theory for wildfire prediction. *Extremes.* 2023;26(2):273–299. doi:10.1007/s10687-022-00454-6
27. Tetko IV, van Deursen R, Godin G. Be aware of overfitting by hyperparameter optimization! *J Cheminf.* 2024;16(1):139. doi:10.1186/s13321-024-00934-w
28. Ozler S, Oztas E, Guler BG, Caglar AT. Increased levels of serum IL-33 is associated with adverse maternal outcomes in placenta previa accreta. *J Matern Fetal Neonatal Med.* 2021;34(19):3192–3199.
29. Chen H, Zhou X, Han TL, Baker PN, Qi H, Zhang H. Decreased IL-33 production contributes to trophoblast cell dysfunction in pregnancies with preeclampsia. *Mediators Inflamm.* 2018;2018:9787239. doi:10.1155/2018/9787239
30. Pieczyńska J, Placzowska S, Pawlik-Sobecka L, Kokot I, Sozański R, Grajeta H. Association of dietary inflammatory index with serum IL-6, IL-10, and CRP concentration during pregnancy. *Nutrients.* 2020;12(9):2789. doi:10.3390/nu12092789
31. Cornish EF, McDonnell T, Williams DJ. Chronic inflammatory placental disorders associated with recurrent adverse pregnancy outcome. *Front Immunol.* 2022;13:825075. doi:10.3389/fimmu.2022.825075
32. Cheng H, Zhu Z, Chi P, et al. Association of systemic chronic inflammation during pregnancy in different periods and its trajectories with preterm birth. *Am J Reprod Immunol.* 2024;91(5):e13848. doi:10.1111/aji.13848
33. Chen J, Khalil RA. Matrix metalloproteinases in normal pregnancy and preeclampsia. *Prog Mol Biol Transl Sci.* 2017;148:87–165. doi:10.1016/bs.pmbts.2017.04.001
34. McKelvey KJ, Ariyakumar G, McCracken SA. Inflammatory and immune system markers. *Methods Mol Biol.* 2018;1710:85–101.
35. Ticconi C, Inversetti A, Marraffa S, et al. Chronic endometritis and recurrent reproductive failure: a systematic review and meta-analysis. *Front Immunol.* 2024;15:1427454. doi:10.3389/fimmu.2024.1427454
36. Jiang S, He F, Gao R, et al. Neutrophil and neutrophil-to-lymphocyte ratio as clinically predictive risk markers for recurrent pregnancy loss. *Reprod Sci.* 2021;28(4):1101–1111. doi:10.1007/s43032-020-00388-z
37. Gurbanova A, Feyaerts D, Benner M, et al. Profiling NK cell receptor repertoire of women with idiopathic recurrent pregnancy loss. *J Reprod Immunol.* 2025;173:104803. doi:10.1016/j.jri.2025.104803
38. Mishra S, Ashish A, Rai S, et al. The impact of inflammatory cytokines on recurrent pregnancy loss: a preliminary investigation. *Reprod Sci.* 2025;32(3):804–814. doi:10.1007/s43032-025-01786-x
39. Li M, Zhou R, Yu D, Chen D, Zhao A. A nomogram and risk stratification to predict subsequent pregnancy loss in patients with recurrent pregnancy loss. *Hum Reprod.* 2024;39(10):2221–2232. doi:10.1093/humrep/deae181

International Journal of General Medicine

Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>

Dovepress

Taylor & Francis Group