Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

An interpretable machine learning model for predicting adverse perinatal outcomes in pregnant women with preterm premature rupture of membranes

Hui Chen

q183187527340163.com

Xiaolan Clinical Institute of Shantou University Medical College

Qiu Mo

Xiaolan Clinical Institute of Shantou University Medical College

Qiu-Xia Feng

Xiaolan Clinical Institute of Shantou University Medical College

Jun-Ping Zhang

Linfen centralLinfen Central Hospital

Qing-Lan Lin

Xiaolan Clinical Institute of Shantou University Medical College

Yan-Ping Yu

Xiaolan Clinical Institute of Shantou University Medical College

Research Article

Keywords: preterm premature rupture of membranes, adverse perinatal outcomes, machine learning, interval from PPROM to delivery, SHapley Additive exPlanation

Posted Date: September 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10834878/v1

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

Additional Declarations: No competing interests reported.

An interpretable machine learning model for predicting adverse perinatal outcomes in pregnant women with preterm premature rupture of membranes

Hui Chen[1]*, Qiu Mo[1], Qiu-Xia Feng[1], Jun-Ping Zhang[2], Qing-Lan Lin [1], Yan-Ping Yu[1]*

[1] Department of Obstetrics and Gynecology, Xiaolan Clinical Institute of Shantou University Medical College, Zhongshan, People's Republic of China

2 Department of Obstetrics and Gynecology, Linfen Central Hospital, Linfen, People's Republic of China

Hui Chen and Qiu Mo contributed equally to this work.

* Correspondence:

Hui Chen

Department of Obstetrics and Gynecology, Xiaolan Clinical Institute of Shantou University Medical College, 65# Jucheng Avenue Middle, Xiaolan, Zhongshan, Guangdong, 528415, People's Republic of China, Tel. +86-076088662120. Email: q18318752734@163.com;

Yan-Ping Yu

Department of Obstetrics and Gynecology, Xiaolan People's Hospital of Zhongshan, 65# Jucheng Avenue Middle, Xiaolan, Zhongshan, Guangdong, 528415, People's Republic of China, Tel. +86-076088662120. Email: YYP631024@163.com

Abstract

Background: To establish and validate a machine learning (ML) model for predicting adverse perinatal outcomes (APOs) in pregnant women with preterm premature rupture of membranes (PPROM).

Methods: From March 2021 to December 2024, a total of 528 pregnant women with PPROM were retrospectively recruited from two centres. The primary endpoints were APOs, including acute respiratory distress syndrome, sepsis, grade III–IV intraventricular haemorrhage, necrotising enterocolitis and neonatal mortality. Five ML methods were used to construct models to predict the APOs. Model performance was evaluated using the area under the curve (AUC). Restricted cubic spline (RCS) analyses were performed to evaluate the association between the interval from PPROM to delivery and the APOs. SHapley Additive exPlanations (SHAP) analysis was performed to interpret the best-performing model.

Results: Among the five ML models, the multilayer perceptron (MLP) model exhibited robust performance for predicting survival in both the test cohort (AUC = 0.802) and the validation cohort (AUC = 0.847). SHAP analysis indicated that the five most important variables are gestational age at PPROM, Apgar score at 5 min, PT, interval from PPROM to delivery and Apgar score at 1 min. In the subgroup analysis of gestational age at PPROM at 34–36^{+6} weeks, the RCS revealed that a nonlinear relationship was found between the interval from PPROM to delivery and the APOs, with a threshold of approximately 26 hours. Below the threshold value, the effect size was 1.07 (95% CI, 1.01–1.14; P = 0.034).

Conclusions: We developed and validated an explainable ML model for predicting APOs among pregnant women with PPROM. For pregnant women who experience PPROM at 34–36^{+6} weeks of gestation, delivery should be pursued within the first 26 hours.

Keywords: preterm premature rupture of membranes; adverse perinatal outcomes; machine learning; interval from PPROM to delivery; SHapley Additive exPlanation

Introduction

Preterm premature rupture of membranes (PPROM) accounts for 30%–40% of all premature deliveries[1-3], leading to complications in 25%–30% of newborns[4], and severely affecting foetal and maternal outcomes[5-8]. PPROM is the leading cause of neonatal mortality and is associated with many adverse perinatal outcomes (APOs), such as acute respiratory distress syndrome, sepsis, bronchopulmonary dysplasia, neurodevelopmental impairment, intraventricular haemorrhage and necrotising enterocolitis[9-12]. Therefore, it is urgent to strengthen the understanding of the risk factors and mechanisms for APOs in pregnant women with PPROM and identifying effective prediction methods to develop personalised diagnosis and treatment plans.

Many studies have predicted the risk of APOs among pregnant women with PPROM. Maternal or foetal biomarkers have good predictive value for neonatal sepsis and necrotising enterocolitis[13-15]. The serum level of syndecan-1 in pregnant women has potential predictive value for APOs[16]. Kim et al. reported that semiquantitative expression of metalloproteinase-8 can predict adverse pregnancy outcomes in patients with premature rupture of membranes[17]. Although these studies predict postpartum complications among pregnant women with PPROM from different perspectives, the predictive performance is not particularly satisfactory, and comprehensive predictive models are lacking.

Machine learning (ML) algorithms have been widely applied in the medical field, especially in the diagnosis and prognosis of diseases[18-21]. Many robust ML prediction models have shown good predictive performance for adverse obstetric events[22, 23] and complications [24]. Traditional prediction models use linear interaction relationships between variables to predict diseases; however, many of the relationships between variables are nonlinear. This is precisely the advantage of ML methods as alternatives to traditional linear models, gradually overcoming the limitations of linear models[25, 26]. However, ML models for predicting postpartum APOs among pregnant women with PPROM are lacking.

The aim of this study was to establish and validate an interpretable ML model to accurately predict the risk of APOs among pregnant women with PPROM. We used the SHapley Additive exPlanation (SHAP) method to clarify feature importance and explain the model's prediction results.

Materials and Methods

Study population

We retrospectively collected clinical data from pregnant women with PPROM and their newborns admitted to Zhongshan Fifth People's Hospital and Linfen Central Hospital between March 2021 and December 2024. The inclusion criteria were women with premature rupture of membranes and delivery between 24^{+0} and 36^{+6} weeks of gestation from March 2021 to December 2024. The exclusion criteria were as follows: 1) pregnant women with twin or multiple births; 2) intrauterine foetal demise; and 3) loss of clinical data for women or neonates. In all, 528 patients who met the inclusion criteria were recruited from Xiaolan Clinical Institute of Shantou University Medical College (Center 1, n=375) and Linfen Central Hospital (Center 2, n=153), Detailed information on the enrolment procedure is displayed in Fig. 1.

Data collection

We retrospectively collected data from 528 pregnant women and their foetuses included in this study, including age, gravidity, parity, prepregnancy body mass index (prepregnancy BMI), HbA1c level, birth weight, Apgar score at 1 min, Apgar score at 5 min, interval from PPROM to delivery (h), D-dimer level, prothrombin time (PT), PCT, hypertensive pregnancy disorder, primigravida status, neonatal sex, caesarean delivery status and gestational age at PPROM. Additionally, inflammatory markers such as neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR),

systemic immune-inflammation index (SII; where SII = platelet × neutrophil/lymphocyte), prognostic nutritional index (PNI; where PNI = lymphocyte ×0.05+ albumin×10), systemic inflammatory response index (SIRI; where SIRI = monocyte × neutrophil/lymphocyte) and lymphocyte-to- monocyte ratio (LMR) were used.

Outcomes

APO was defined as any of the following: acute respiratory distress syndrome, sepsis[10], grade III–IV intraventricular haemorrhage, necrotising enterocolitis[27] and neonatal mortality.

Model Development and Evaluation

Five ML models, namely, logistic regression (LR), extreme gradient boosting (XGB), light gradient boosting machine (LightGBM), adaptive boosting (AdaBoost) and multilayer perceptron (MLP), were used to predict the risk of APOs in PPROM. Patients were randomly split into a training cohort (70%) and a test cohort (30%). Tenfold cross-validation was used to improve the models. To evaluate the predictive performance of each model, we performed a quantitative analysis using the area under the receiver operating characteristic curve (AUC). Concurrently, metrics, including accuracy, sensitivity, specificity, and F1 score, were calculated. Decision curve analysis (DCA) was utilised to assess the clinical utility of each model by quantifying the net benefit at different threshold probabilities. SHapley Additive exPlanations (SHAP) analysis was used to interpret model predictions and address the “black-box” nature of ML models. SHAP values quantify the contribution of each feature to the predicted outcome by reflecting the magnitude and direction of changes in the model output, thereby enabling the ranking of feature importance and improving model interpretability.

Statistical analysis

Continuous variables with a normal distribution are expressed as medians with

interquartile ranges (IQRs) and were compared using t tests. Categorical variables are presented as percentages and were compared by the chi-square test or Fisher's exact test. The data were divided into four quantiles on the basis of order of the interval from PPROM to delivery (h) values from lowest (Q1) to highest (Q4), thereby converting them into a categorical variable. A multivariate model employing restricted cubic splines (RCS) was used to analyse the association between interval from PPROM to delivery (h) and APOs. A two-tailed P value < 0.05 was considered to indicate statistical significance. SPSS (version 26.0, IBM), Python (version 3.8; http://www.python.org) and R 4.1.2 statistical packages were used to analyse the data.

Results

Baseline characteristics of the study population

The baseline clinical characteristics of 528 pregnant women with PPROM and their newborns are outlined in Table 1. In the training cohort, test cohort, and validation cohort, 32 (12.2%), 13 (11.5%), and 11 (7.2%) neonates, respectively, experienced APOs. Across all three cohorts, no significant differences were observed between the APO and non-APO groups in terms of age, gravidity, parity, HbA1c, D-dimer, PCT, NLR, PLR, SII, PNI, SIRI, LMR, hypertensive pregnancy disorder, primigravida or neonatal sex (P > 0.05). However, significant differences were noted in gestational age at PPROM, Apgar score at 5 min and interval from PPROM to delivery (p < 0.05). The remaining variables were significantly different in one or two of the three cohorts (p < 0.05).

Model performance and interpretability

We carried out ten rounds of 10-fold internal cross-validation to construct five ML models (Supplementary Fig. 1). As illustrated in Table 2 and Fig. 2A–B, among all five ML models, the MLP model achieved the highest AUC (0.802), accuracy (0.779), specificity (0.790), and F1 score (0.419) in the test cohort. In the validation cohort, the MLP model outperformed the other four ML models in terms of the AUC (0.847), accuracy (0.758), sensitivity (0.909), specificity (0.746), and F1 score (0.351).

The DCA curves for the test and validation sets are shown in Fig. 2C and 2D, respectively. These curves indicate that the MLP model has a significant net gain advantage over the other models across a broad range of threshold probabilities.

The SHAP plot displays the five strongest predictors for APOs, ranked in order as gestational age at PPROM, Apgar score at 5 min, PT, interval from PPROM to delivery, and Apgar score at 1 min (Fig. 3A). The SHAP plot illustrates the contribution of each feature to the prediction of APOs in the MLP model for two pregnant women with PPROM. The specific numerical values of each feature and their corresponding SHAP values visually illustrate the positive and negative influences of these features on the prediction outcomes (Fig. 3B-C).

Subgroup analysis

We further analysed the subgroup with gestational age at PPROM between 34 and 36^{+6} weeks, comprising a total of 403 individuals across the three cohorts. We investigated the association between the interval from PPROM to delivery and APOs, further subdivision of the interval from PPROM to delivery into four quartiles. Notably, compared to patients in the first quartile, patients in the fourth quartile of Model 3 (OR, 4.532; 95% CI, 1.309-15.696, P=0.017) had significantly increased APO risk (Table 3). The results indicate that the interval from PPROM to delivery has a nonlinear relationship with APOs (p-nonlinear = 0.029) (Fig. 4). By employing a two-stage linear regression model, the threshold for the interval from PPROM to delivery was determined to be 26 hours. Above this threshold, a stable association existed between gestational age at PPROM and APO incidence (OR, 1.00; 95% CI, 1.00–1.01; P = 0.729). Conversely, below this threshold, each unit increase in the interval from PPROM to delivery was associated with a 7% increase in the risk of developing APOs (OR, 1.07; 95% CI, 1.01–1.14; P= 0.034) (Supplementary Table S1).

Discussion

Early prediction of postpartum APOs among pregnant women with PPROM is crucial for optimising treatment decisions. In this multicentre study, we developed an ML model based on five risk factors and validated its predictive performance in an external validation cohort. These findings indicate that our model can be used as a personalised tool to predict the occurrence of APOs among pregnant women with PPROM after delivery, providing guidance for clinical doctors' treatment decisions.

To date, studies concerning the prediction of postpartum complications in women with PPROM have been reported[9, 28-31]. However, the majority of these studies have relied on model construction using a single method (traditional LR), which may prove inadequate for handling the complex relationships between variables and disease, thereby compromising predictive performance. Among the five ML models examined in this study, the MLP model distinguished itself with the highest area under the curve (AUC), exhibiting both excellent accuracy and net benefit. This model also performed most impressively on the independent external validation set. Through its multilayer architecture and nonlinear activation functions (such as sigmoid and ReLU), the MLP can approximate arbitrary nonlinear functions, effectively handling complex regression problems and classification tasks. The MLP ML method has been shown to hold significant value in the medical field[32, 33].

In this study, the MLP algorithm was employed to construct a final model incorporating five features. These characteristics are readily obtained from the hospital records of patients with PPROM, indicating that the model is an effective tool for predicting the risk of APOs in women with PPROM. In our model, the predictors of APOs were Apgar score at 1 min, 2 Apgar score at 5 min, interval from PPROM to delivery, PT and gestational age at PPROM. Tariku et al.’s study included pregnant women who experienced PPROM before 37 weeks of gestation. Analysis revealed that gestational age at PPROM is an independent risk factor for adverse postnatal outcomes[34]. Our ML model reached similar conclusions. Recep et al. analysed the neonatal mortality rate among 183 pregnant women with PPROM following delivery

and reported that it was closely associated with low gestational age, which aligns with our conclusions[35]. ML techniques are often termed “black boxes”, as the process from prediction to outcome is virtually impossible to explain[36]. This opacity may cause clinicians to hesitate to implement medical conclusions derived from ML. However, a significant advantage of this study lies in our use of the SHAP method to rationally reveal the “black box” operational mechanisms of the ML model. The SHAP method provides both global and local explanations, clearly elucidating the model's functionality and revealing the importance ranking of five clinical predictive factors within the model.

In this study, the interval from PPROM to delivery was identified as a predictor of APOs. To our knowledge, whether women with PPROM at 34 to 36^{+6} weeks should receive expectant management remains a matter of ongoing debate[37]. The American College of Obstetricians and Gynaecologists guidelines recommend that if PPROM occurs after 34 weeks of pregnancy, the induction of labour should be considered[38]. The guidelines of the Royal College of Obstetricians and Gynaecologists in the UK clearly state that childbirth should be considered at 34 weeks of pregnancy and recommend that PPROM patients receiving expectant treatment at 34 weeks of pregnancy be informed of postpartum neonatal respiratory problems, neonatal intensive care unit (NICU) occupancy rates, and a reduced risk of caesarean section [9]. However, one randomized controlled trial (PPROMT trial) suggested that expectant management should be adopted for pregnant women with PPROM occurring between 34 and 36^{+6} weeks, while closely monitoring both maternal and foetal health[10]. The aforementioned study merely conducted a conclusive analysis and did not specify the precise timeframe after PPROM, beyond which antenatal management becomes unsuitable. Our study involved a grouped analysis of specific time intervals from PPROM to delivery. RCS analysis confirmed a significant nonlinear correlation between the interval from PPROM to delivery and APO risk, and when the interval from PPROM to delivery was less than 26 hours, APO risk increased significantly. Our research recommends that pregnant women with PPROM at 34 to 36^{+6} weeks should

be delivered immediately, particularly within the first 26 hours.

Additionally, this study has several limitations. First, as a retrospective multicentre study, potential selection bias may inevitably exist. Second, the independent dataset used for external validation was relatively small, and with only one external validation cohort available, adequately assessing the extrapolation capabilities of the modelled approach is challenging. Consequently, future prospective multicentre studies on a greater scale are needed to validate our ML model. Finally, this study employed a single-modality approach utilising solely clinical variables. Future research may incorporate multimodal data, including deep learning techniques applied to imaging data, such as prenatal colour Doppler ultrasound.

Conclusions

In summary, we developed and validated a prediction model for APOs in patients with PPROM at diagnosis. The combined use of the five predictors showed good performance. This study provides valuable insights for pregnant women with PPROM, especially those with a gestational age of 34–36^{+6} weeks. Nevertheless, future prospective studies are needed to validate the prediction model.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

None.

Funding

This research received no external funding.

Author Contributors

HC and QM were responsible for study design and drafting the manuscript. QXF, JPZ, and QLL were responsible for data collection. HC were responsible for statistics analysis. HC and QM conducted the data analysis and developed and validated the ML models. YPY critically reviewed the manuscript. HC and QM accessed and verified the underlying data. All authors read and approved the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare that they have no competing interests.

Ethics approval and consent to participate

Ethics approval was obtained from the institutional review boards of the Linfen Central Hospital and Xiaolan Clinical Institute of Shantou University Medical College (Approval No. ZSXL-LL2025-075). This study protocol complied with the guidelines of the Declaration of Helsinki and was approved by the institutional review board of all participating centres. The requirement for informed consent was waived owing to the retrospective study design.

Reference

1. Cox SM, Williams ML, Leveno KJ. The natural history of preterm ruptured membranes: what to expect of expectant management. Obstet Gynecol. 1988; 71(4):558-562.
2. Gibbs RS, Blanco JD. Premature rupture of the membranes. Obstet Gynecol. 1982; 60(6):671-679.
3. Mercer BM, Goldenberg RL, Meis PJ, Moawad AH, Shellhaas C, Das A, Menard MK, Caritis SN, Thurnau GR, Dombrowski MP, et al. The Preterm Prediction Study: prediction of preterm premature rupture of membranes through clinical findings and ancillary testing. The National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. Am J Obstet Gynecol. 2000; 183(3):738-745.
4. Mercer BM, Miodovnik M, Thurnau GR, Goldenberg RL, Das AF, Ramsey RD, Rabello YA, Meis PJ, Moawad AH, Iams JD, et al. Antibiotic therapy for reduction of infant morbidity after preterm premature rupture of the membranes. A randomized controlled trial. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. Jama. 1997; 278(12):989-995.
5. Furman B, Shoham-Vardi I, Bashiri A, Erez O, Mazor M. Clinical significance and outcome of preterm prelabor rupture of membranes: population-based study. Eur J Obstet Gynecol Reprod Biol. 2000; 92(2):209-216.
6. Goldenberg RL, Nelson KG, Davis RO, Koski J. Delay in delivery: influence of gestational age and the duration of delay on perinatal outcome. Obstet Gynecol. 1984; 64(4):480-484.
7. Mercer BM. Preterm premature rupture of the membranes. Obstet Gynecol. 2003; 101(1):178-193.
8. Bond DM, Middleton P, Levett KM, van der Ham DP, Crowther CA, Buchanan SL, Morris J. Planned early birth versus expectant management for women with

preterm prelabour rupture of membranes prior to 37 weeks' gestation for improving pregnancy outcome. Cochrane Database Syst Rev. 2017; 3(3):Cd004735.

9. Birgisdottir BT, Andersson T, Varli IH, Saltvedt S, Abtahi F, Åden U, Holzmann M. Risk factors and risk-indicating model for early-onset neonatal sepsis after preterm prelabor rupture of membranes: A historical cohort study. Acta Obstet Gynecol Scand. 2025; 104(8):1475-1486.
10. Morris JM, Roberts CL, Bowen JR, Patterson JA, Bond DM, Algert CS, Thornton JG, Crowther CA. Immediate delivery compared with expectant management after preterm pre-labour rupture of the membranes close to term (PPROMT trial): a randomised controlled trial. Lancet. 2016; 387(10017):444-452.
11. Younge NE, Saha S, Brumbaugh JE, Klein JM, Bell EF, Colaizy TT, Hughes BL, Malcolm WF, Goldberg RN, Wyckoff MH, et al. Outcomes of extremely preterm infants exposed to prolonged prelabor rupture of membranes before 24 weeks of gestation. Am J Obstet Gynecol. 2025; 233(2):131.e131-131.e114.
12. Pharande P, Mohamed AL, Bajuk B, Lui K, Bolisetty S. Preterm infant outcomes in relation to the gestational age of onset and duration of prelabour rupture of membranes: a retrospective cohort study. BMJ Paediatr Open. 2017; 1(1):e000216.
13. Matyas M, Ilyés T, Valeanu M, Crăciun AM, Hăşmăşanu M, Grosu N, Zaharie G. The predictive value of maternal and neonatal inflammatory biomarkers for necrotizing enterocolitis. Eur J Pediatr. 2025; 184(5):316.
14. Musilova I, Leskova J, Kacerovsky M, Ledvinova J, Parmar A, Hornychova H, Stichhauer R, Maly J, Musilova I. Role of Intrauterine Infectious and Inflammatory Complications in the Development of Necrotizing Enterocolitis in a Newborn from Preterm Prelabor Rupture of Membranes' Gestation: A Retrospective Cohort Study. Fetal Diagn Ther. 2025; 52(5):468-482.
15. Grill A, Goeral K, Leitich H, Farr A, Berger A, Rittenschober-Boehm J. Maternal biomarkers in predicting neonatal sepsis after preterm premature

rupture of membranes in preterm infants. Acta Paediatr. 2024; 113(5):962-972.

16. Ağaoğlu RT, Çirkin Tekeş G, Oğuz Y, Akın F, Bayrak A, Filiz AA, Akbulut Ö V, Pay RE, Yakut Yücel K. Maternal serum syndecan-1 levels in preterm prelabor rupture of membranes and potential predictor of adverse perinatal outcomes: A case-control study. Int J Gynaecol Obstet. 2025; 171(1):326-336.
17. Ji Kim H, Choi J, Ji Oh E, Lee KN, Yoon Park J, Joon Oh K. Semi-quantitative metalloproteinase-8 rapid test for the prediction of adverse pregnancy outcomes in patients with preterm premature rupture of membranes. Eur J Obstet Gynecol Reprod Biol. 2024; 297:65-71.
18. Sun Y, Zhang L, Dong D, Li X, Wang J, Yin C, Poon LC, Tian J, Wu Q. Application of an individualized nomogram in first-trimester screening for trisomy 21. Ultrasound Obstet Gynecol. 2021; 58(1):56-66.
19. Namburete AI, Stebbing RV, Kemp B, Yaqub M, Papageorghiou AT, Alison Noble J. Learning-based prediction of gestational age from ultrasound images of the fetal brain. Med Image Anal. 2015; 21(1):72-86.
20. Lipschuetz M, Guedalia J, Rottenstreich A, Novoselsky Persky M, Cohen SM, Kabiri D, Levin G, Yagel S, Unger R, Sompolinsky Y. Prediction of vaginal birth after cesarean deliveries using machine learning. Am J Obstet Gynecol. 2020; 222(6):613.e611-613.e612.
21. Hou F, Zhu Y, Zhao H, Cai H, Wang Y, Peng X, Lu L, He R, Hou Y, Li Z, et al. Development and validation of an interpretable machine learning model for predicting the risk of distant metastasis in papillary thyroid cancer: a multicenter study. EClinicalMedicine. 2024; 77:102913.
22. Guedalia J, Sompolinsky Y, Novoselsky Persky M, Cohen SM, Kabiri D, Yagel S, Unger R, Lipschuetz M. Prediction of severe adverse neonatal outcomes at the second stage of labour using machine learning: a retrospective cohort study. Bjog. 2021; 128(11):1824-1832.
23. Meng X, Wang L, Wu M, Zhang N, Li X, Wu Q. Construction of a Prediction Model for Adverse Perinatal Outcomes in Foetal Growth Restriction Based on a Machine Learning Algorithm: A Retrospective Study. Bjog. 2025;

132(10):1514-1523.

24. Mustafa HJ, Kalafat E, Prasad S, Heydari MH, Nunge RN, Khalil A. Prediction of hypertension and diabetes in twin pregnancy using machine learning model based on characteristics at first prenatal visit: national registry study. Ultrasound Obstet Gynecol. 2025; 65(5):613-623.

25. Punyapet P, Suwanrath C, Chainarong N, Sawaddisan R, Vichitkunakorn P. Predictors of adverse perinatal outcomes in fetal growth restriction using a combination of maternal clinical factors and simple ultrasound parameters. Int J Gynaecol Obstet. 2023; 162(2):669-675.

26. Powel JE, Zantow EW, Bialko MF, Farley LG, Lawlor ML, Mullan SJ, Vricella LK, Tomlinson TM. Predictive index for adverse perinatal outcome in pregnancies complicated by fetal growth restriction. Ultrasound Obstet Gynecol. 2023; 61(3):367-376.

27. Wang Y, Song J, Sun H, Xu F, Li K, Nie C, Zhang X, Peng X, Xia L, Shen Z, et al. Erythropoietin prevents necrotizing enterocolitis in very preterm infants: a randomized controlled trial. J Transl Med. 2020; 18(1):308.

28. Ohuma EO, Jabin N, Young MF, Epie T, Martorell R, Peña-Rosas JP, Garcia-Casal MN, Papageorghiou AT, Kennedy SH, Villar J. Association between maternal haemoglobin concentrations and maternal and neonatal outcomes: the prospective, observational, multinational, INTERBIO-21st fetal study. Lancet Haematol. 2023; 10(9):e756-e766.

29. Zhu H, Wang Y, Yin H, Liu F, Ma Y, Li X. Risk factors associated with respiratory distress syndrome in late preterm infants. Pak J Med Sci. 2024; 40(9):1947-1952.

30. Chen H, Tan Q, Lai S, Mai H, Wang D. Association between glycated hemoglobin and the risk of neonatal respiratory distress syndrome in preterm premature rupture of membranes pregnancies. Sci Rep. 2024; 14(1):27122.

31. Faucett AM, Metz TD, DeWitt PE, Gibbs RS. Effect of obesity on neonatal outcomes in pregnancies with preterm premature rupture of membranes. Am J Obstet Gynecol. 2016; 214(2):287.e281-287.e285.

32. Liang L, Mai S, Sethi G, Luo Y, Ma Z, Wu L, Lu D, Han J, Cai R, Zhou Y, et al. Fungal microbiota signatures anticipate neoadjuvant immunochemotherapy outcomes in esophageal cancer. J Immunother Cancer. 2025; 13(10).
33. Ren D, Wang Y, Chen L, He J, Shen T. Machine learning-based radiomics approach assessing preoperative non-contrast CT for microsatellite instability prediction in colon cancer. Front Physiol. 2025; 16:1672636.
34. Abebe TA, Nima DD, Mariye YF, Leminie AA. Determinants for perinatal adverse outcomes among pregnant women with preterm premature rupture of membrane: A prospective cohort study. Front Reprod Health. 2022; 4:1052827.
35. Ağaoğlu RT, Öztürk Ö, Ulusoy CO, Günday F, Sarikaya Kurt D, Aksu M, Hizli B, Yakut Yücel K. Perinatal outcomes and predictors of neonatal mortality in preterm premature rupture of membranes: a tertiary center experience. BMC Pregnancy Childbirth. 2025; 25(1):585.
36. The Lancet Respiratory M. Opening the black box of machine learning. Lancet Respir Med. 2018; 6(11):801.
37. van der Ham DP, Vijgen SM, Nijhuis JG, van Beek JJ, Opmeer BC, Mulder AL, Moonen R, Groenewout M, van Pampus MG, Mantel GD, et al. Induction of labor versus expectant management in women with preterm prelabor rupture of membranes between 34 and 37 weeks: a randomized controlled trial. PLoS Med. 2012; 9(4):e1001208.
38. ACOG Practice Bulletin No. 80: premature rupture of membranes. Clinical management guidelines for obstetrician-gynecologists. Obstet Gynecol. 2007; 109(4):1007-1019.

Table 1 Clinical characteristics of pregnant women with PPROM in different cohorts.

Characteristic	Training cohort			Test cohort			Validation cohort		
	APO (-) (N = 230)	APO (+) (N = 32)	p value	APO (-) （N = 100）	APO (+) （N = 13）	p value	APO (-) （N = 142）	APO (+) （N = 11）	p value
Age, years	30.0 (27.0, 34.0)	31.5 (28.0, 34.0)	0.462	30.0 (28.0, 33.0)	32.0 (30.0, 35.0)	0.285	31.0 (27.0, 34.0)	32.0 (28.0, 36.0)	0.513
Gravidity	2.00 (1.00, 4.00)	2.00 (1.50, 4.50)	0.463	2.00 (1.00, 4.00)	2.00 (1.00, 2.00)	0.070	2.00 (1.00, 4.00)	3.00 (2.00, 3.00)	0.736
Parity	2.00 (1.00, 2.00)	2.00 (1.00, 2.00)	0.686	2.00 (1.00, 2.00)	2.00 (1.00, 2.00)	0.615	2.00 (1.00, 2.00)	2.00 (1.00, 3.00)	0.523
Prepregnancy BMI (kg/m^2)	21.0 (19.3, 23.2)	21.1 (18.3, 24.0)	0.777	20.0 (18.0, 23.4)	23.5 (21.3, 24.1)	0.012	20.4 (18.9, 23.5)	24.2 (23.3, 25.6)	0.006
HbA1c (%)	5.00 (4.80, 5.20)	5.05 (4.90, 5.20)	0.307	5.00 (4.80, 5.20)	5.10 (4.90, 5.20)	0.545	5.00 (4.70, 5.20)	5.00 (4.80, 5.00)	0.898
Birth weight (g)	2,400 (2,080, 2,700)	2,020 (1,560, 2,445)	<0.001	2,345 (2,060, 2,655)	2,100 (1,920, 2,190)	0.014	2,500 (2,180, 2,730)	2,270 (1,390, 2,730)	0.222
Apgar score at 1 min	10.00 (10.00, 10.00)	9.00 (9.00, 10.00)	<0.001	10.00 (10.00, 10.00)	10.00 (9.00, 10.00)	0.163	10.00 (10.00, 10.00)	9.00 (6.00, 10.00)	<0.001
Apgar score at 5 min	10.00 (10.00, 10.00)	10.00 (9.00, 10.00)	<0.001	10.00 (10.00, 10.00)	10.00 (9.00, 10.00)	<0.001	10.0000 (10.0000, 10.0000)	10.0000 (9.0000, 10.0000)	<0.001
Interval from PPROM to delivery (h)	15 (8, 32)	31 (11, 115)	0.015	14 (8, 34)	61 (19, 160)	0.012	14 (7, 43)	42 (22, 74)	0.015

D-dimer (mg/L)	1,068 (716, 1,487)	910 (654, 1,632)	0.549	1,142 (774, 1,547)	908 (697, 1,093)	0.180	1,252 (776, 1,795)	648 (561, 1,827)	0.253
PT(s)	11.00 (10.60, 11.40)	11.40 (10.90, 11.75)	0.009	11.05 (10.60, 11.50)	11.10 (10.80, 11.40)	0.576	11.00 (10.70, 11.50)	11.60 (11.00, 11.70)	0.041
PCT (μg/L)	0.05 (0.04, 0.07)	0.06 (0.04, 0.07)	0.939	0.060 (0.040, 0.070)	0.050 (0.030, 0.071)	0.619	0.05 (0.04, 0.07)	0.06 (0.05, 0.07)	0.544
NLR	4.27 (3.36, 5.58)	4.57 (3.84, 6.32)	0.187	4.40 (3.46, 5.97)	3.99 (3.02, 5.62)	0.259	4.13 (3.30, 5.63)	4.19 (3.43, 6.15)	0.364
PLR	127 (103, 161)	140 (111, 171)	0.515	145 (109, 185)	105 (83, 149)	0.050	133 (101, 162)	143 (113, 176)	0.427
SII	938 (675, 1,251)	974 (709, 1,292)	0.799	1,028 (757, 1,290)	855 (572, 1,265)	0.223	870 (675, 1,236)	1,125 (854, 1,432)	0.101
PNI	341 (326, 358)	342 (317, 361)	0.568	341 (325, 364)	356 (323, 360)	0.637	343 (325, 355)	351 (338, 364)	0.245
SIRI	2.38 (1.62, 3.37)	2.54 (1.72, 3.48)	0.635	2.35 (1.72, 3.66)	2.53 (1.72, 4.36)	0.777	2.37 (1.69, 3.59)	2.39 (1.94, 3.80)	0.727
LMR	3.05 (2.31, 3.83)	2.90 (2.15, 3.45)	0.367	2.88 (2.30, 3.70)	3.25 (2.27, 4.15)	0.743	2.93 (2.30, 3.66)	2.91 (2.48, 3.17)	0.902
Hypertensive pregnancy disorder			0.999			0.528			0.999
No	215 (93.5%)	30 (93.8%)		95 (95.0%)	12 (92.3%)		132 (93.0%)	11 (100.0%)	
Yes	15 (6.5%)	2 (6.3%)		5 (5.0%)	1 (7.7%)		10 (7.0%)	0 (0.0%)	
Primigravida			0.691			0.829			0.755
No	128 (55.7%)	19 (59.4%)		57 (57.0%)	7 (53.8%)		77 (54.2%)	7 (63.6%)	
Yes	102 (44.3%)	13 (40.6%)		43 (43.0%)	6 (46.2%)		65 (45.8%)	4 (36.4%)	

Neonatal sex			0.582			0.547		0.334
Female	96 (41.7%)	15 (46.9%)		45 (45.0%)	7 (53.8%)		52 (36.6%)	6 (54.5%)	
Male	134 (58.3%)	17 (53.1%)		55 (55.0%)	6 (46.2%)		90 (63.4%)	5 (45.5%)	
Caesarean delivery			0.079			0.031		0.055
No	145 (63.0%)	15 (46.9%)		70 (70.0%)	5 (38.5%)		85 (59.9%)	3 (27.3%)	
Yes	85 (37.0%)	17 (53.1%)		30 (30.0%)	8 (61.5%)		57 (40.1%)	8 (72.7%)	
Gestational age at PPROM			<0.001			0.004		0.010
< 34(weeks)	46 (20.0%)	19 (59.4%)		21 (21.0%)	8 (61.5%)		25 (17.6%)	6 (54.5%)	
34–36^{+6}(weeks)	184 (80.0%)	13 (40.6%)		79 (79.0%)	5 (38.5%)		117 (82.4%)	5 (45.5%)	

Abbreviations: APO, adverse perinatal outcome; LMR, lymphocyte-to-monocyte ratio; NLR, neutrophil-to-lymphocyte ratio; PLR, platelet-to-lymphocyte ratio; PNI, prognostic nutritional index; SII, systemic immune-inflammation index; SIRI, systemic inflammation response index; PPROM, premature rupture of membranes; BMI, body mass index; PT, prothrombin time.

Table 2 Performance of the models in predicting adverse perinatal outcomes.

Model	AUC	Accuracy	Sensitivity	Specificity	F1-score
Test cohort					
LR	0.799	0.637	0.846	0.610	0.349
XGBoost	0.770	0.637	0.846	0.610	0.349
LightGBM	0.735	0.637	0.846	0.610	0.349

AdaBoost	0.683	0.522	0.846	0.480	0.289
MLP	0.802	0.779	0.692	0.790	0.419
Validation cohort					
LR	0.842	0.745	0.909	0.732	0.339
XGBoost	0.757	0.608	0.909	0.585	0.250
LightGBM	0.773	0.647	0.909	0.627	0.270
AdaBoost	0.788	0.725	0.818	0.718	0.300
MLP	0.847	0.758	0.909	0.746	0.351

Abbreviations: LR, logistic regression; XGBoost, extreme gradient boosting; LightGBM, light gradient boosting machine; AdaBoost, adaptive boosting; MLP, multilayer perceptron; AUC, area under the curve

Table 3 Association of Interval from PPROM to delivery and adverse perinatal outcomes.

Interval from PPROM to delivery (h)	**Model 1 OR (95%CI)**	**p value**	**Model 2 OR (95%CI)**	**p value**	**Model 3 OR (95%CI)**	**p value**
Quartile						
Q1	Ref		Ref		Ref	
Q2	0.625 (0.145-2.689)	0.528	0.667 (0.152-2.926)	0.591	1.134 (0.231-5.568)	0.877
Q3	0.800 (0.209-3.067)	0.745	0.874 (0.225-3.403)	0.846	1.524 (0.3.7-7.562)	0.606
Q4	2.503 (0.837-7.488)	0.101	3.095 (1.002-9.557)	0.049	4.532 (1.309-15.696)	0.017

Model 1: was not adjusted for any covariates.

Model 2: was adjusted for age, prepregnancy BMI and hypertensive pregnancy disorder.

Model 3: was adjusted for age, prepregnancy BMI and hypertensive pregnancy disorder, gestational age at PPROM, D-dimer, PT, gravidity, parity, caesarean delivery, Apgar score at 1 min, Apgar score at 5 min, birth weight, primigravida, Neonatal sex, NLR, PLR, SII, PNI, SIRI, LMR, HbA1c and primigravida.

Q1: ⩽7 h; Q2: > 7 h and ⩽13 h; Q3: > 13 and ⩽ 26 h; Q4: > 26 h

Abbreviations: OR, odds ratio; CI, confidence interval; PPROM, preterm premature rupture of membranes.

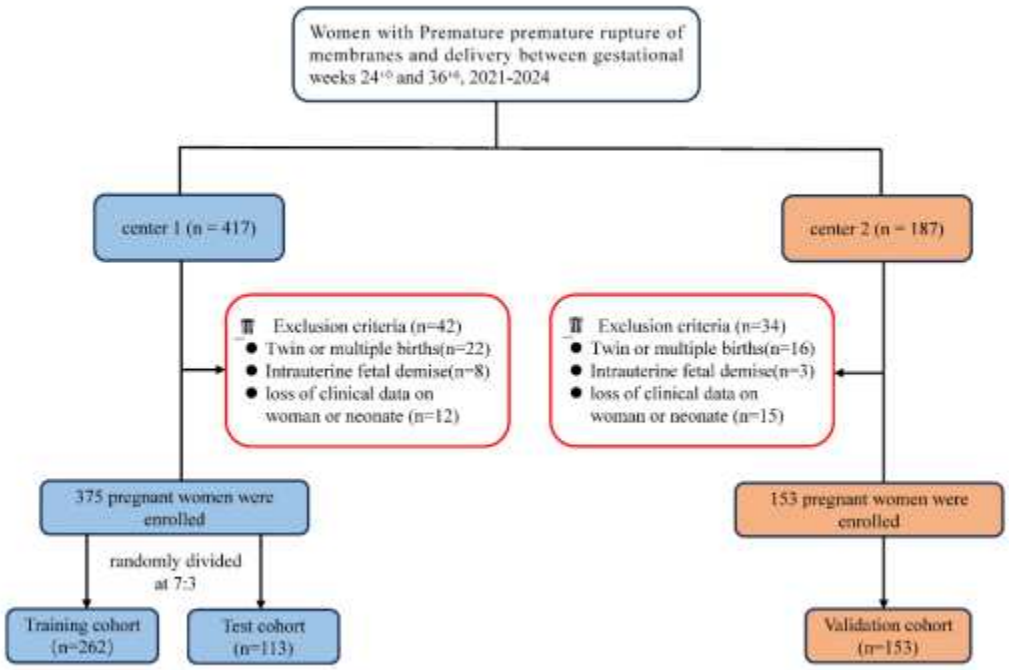


Fig. 1 Flow diagram of the study cohort selection process.

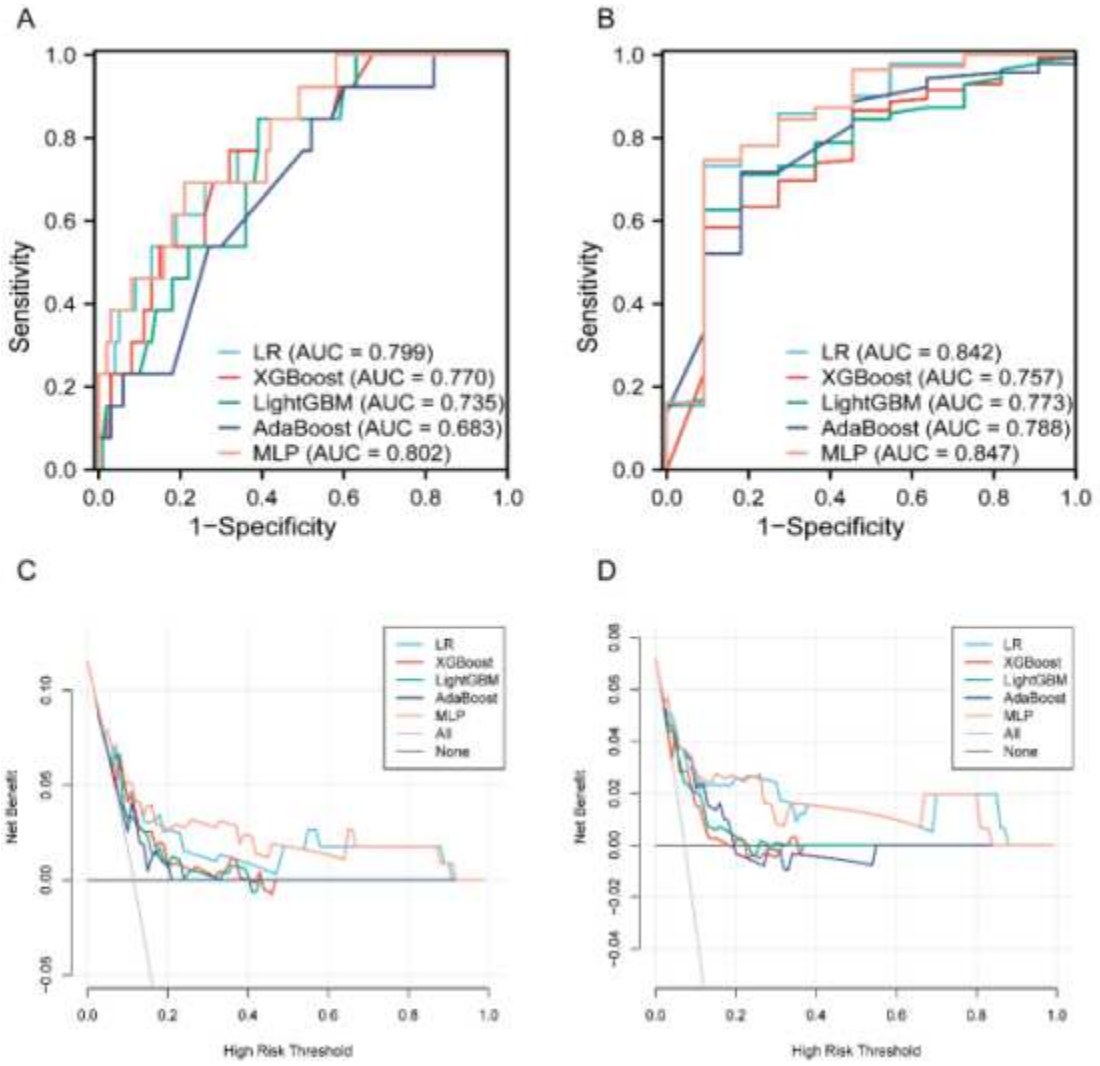


Fig. 2 Performance of the five machine learning models in predicting adverse perinatal outcomes among pregnant women with PPROM in the test cohort (A) and validation cohort (B). Decision curve analysis for the five machine learning models in the test cohort (C) and validation cohort (D).

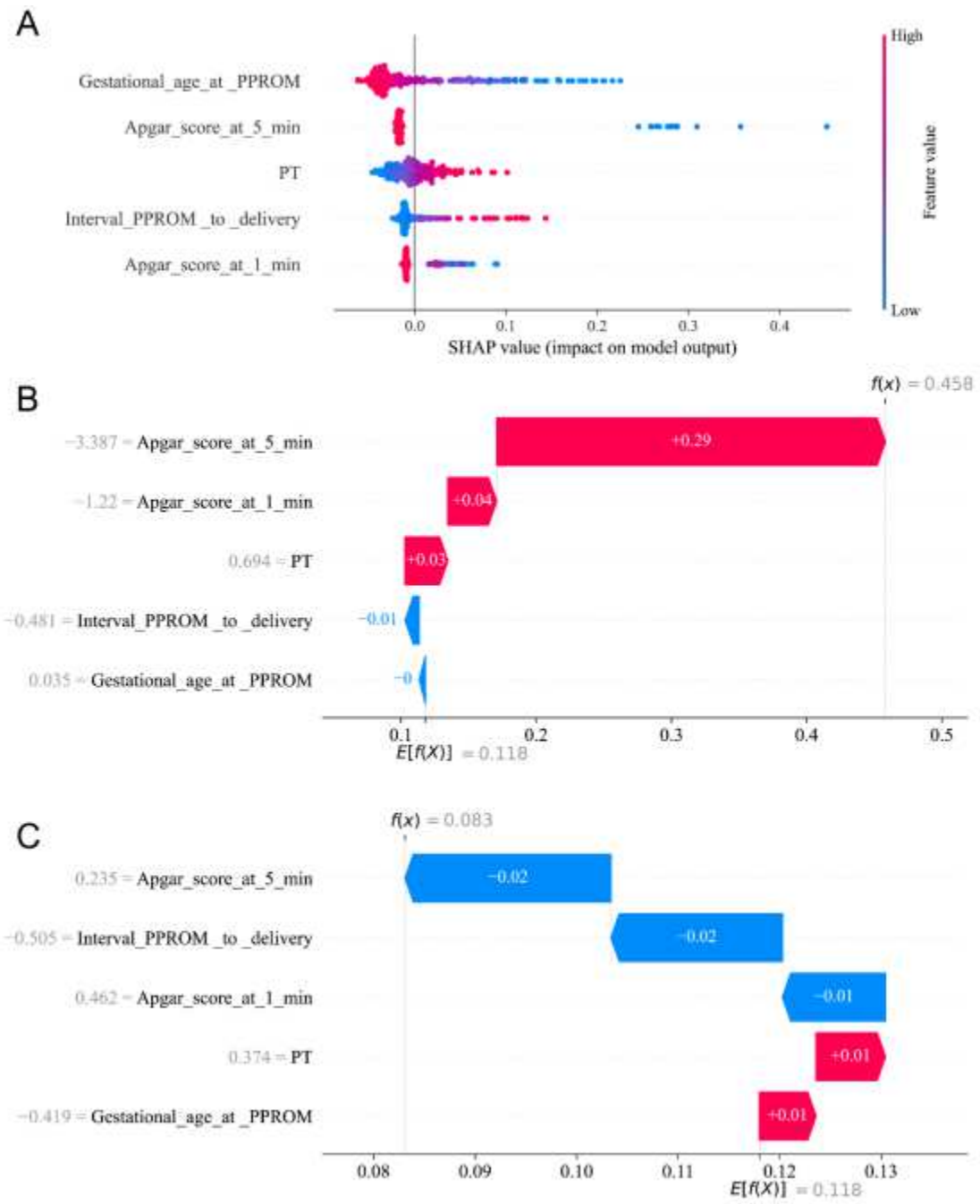


Fig. 3 SHAP diagram of the MLP model. (A) SHAP honeycomb diagram of the MLP model. (B) APO-positive and (C) APO-negative SHAP waterfall plots showing interpretable examples of single-sample feature prediction results.

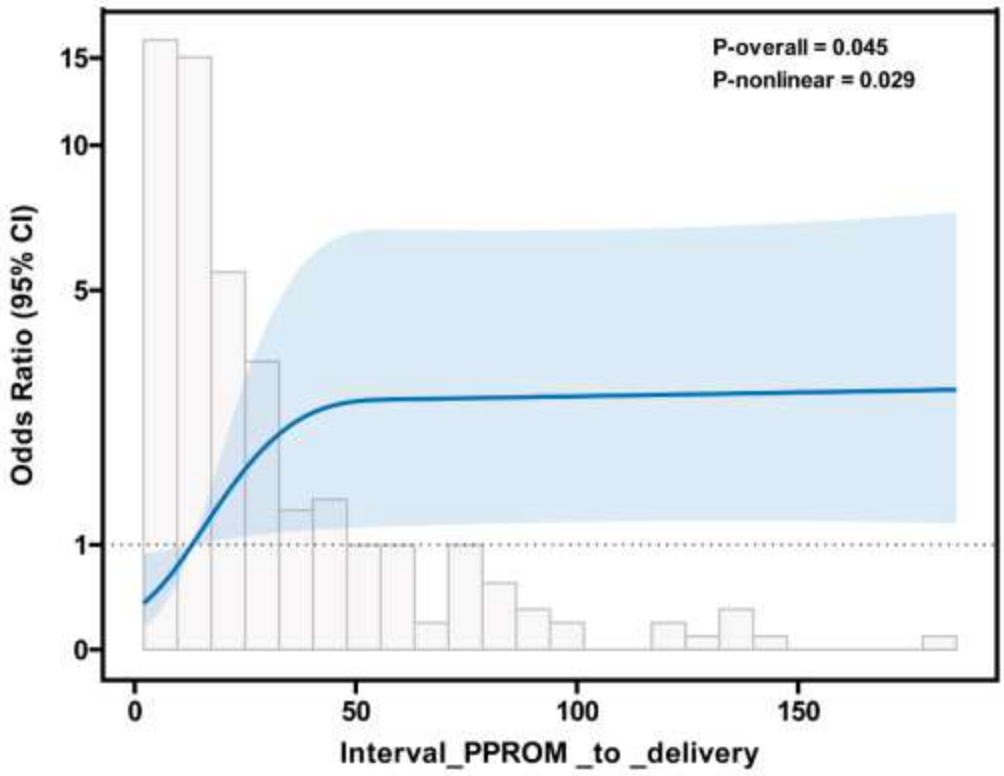


Fig. 4 Restricted cubic spline curves for the association between the interval from PPROM to delivery and adverse perinatal outcomes. The blue lines represent references for hazard ratios, and the blue areas represent 95% confidence intervals. The model was adjusted for age, prepregnancy BMI and hypertensive pregnancy disorder, gestational age at PPROM, D-dimer, PT, gravidity, parity, caesarean delivery, Apgar score at 1 min, Apgar score at 5 min, birth weight, primigravida, neonatal sex, NLR, PLR, SII, PNI, SIRI, LMR, HbA1c and primigravida.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementary.docx