International Journal of General Medicine

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Machine Learning-Driven Intrapartum Fever Prediction: A Comprehensive Large Retrospective Cohort Study Integrating Inflammatory and Obstetric Markers

Hong Jiang, Na Li

Department of Obstetrics, Maternal and Child Health Hospital of Hubei Province, Wuhan, Hubei, 430070, People's Republic of China

Correspondence: Na Li, Department of Obstetrics, Maternal and Child Health Hospital of Hubei Province, No. 745 Wuluo Road, Hongshan District, Wuhan, Hubei, 430070, People's Republic of China, Email 13296609466@163.com

Objective: To develop and validate a machine learning-driven predictive model for intrapartum fever in parturients receiving neuraxial labor analgesia, integrating comprehensive clinical and hematological markers.

Methods: Among 15,760 parturients (2022–2024), 11,032 (70%) were allocated to the training cohort (834 [7.6%] febrile cases) and 4728 (30%) to the testing cohort (364 [7.7%] febrile cases). A three-stage variable screening process was applied, including Pearson correlation analysis (|r| > 0.15), LASSO regression with 10-fold cross-validation, and SHAP value analysis (top 75% importance). Seven machine learning algorithms, namely Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), eXtreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), Multilayer Perceptron (MLP), and Elastic Net (ENET), were evaluated via accuracy, ROC AUC, and cost-benefit analysis.

Results: Key predictors were neutrophil-lymphocyte ratio (NLR, SHAP=0.27), white blood cell count (WBC, SHAP=0.22), and primiparity (SHAP=0.18), with fever cases showing elevated NLR (7.71 vs 4.60, P<0.001) and vaginal exams (3.24 vs 2.29, P<0.001). Notably, the Random Forest (RF) model achieved a high test AUC of 0.94 but a reduced specificity of 0.57, which may increase false-positive risks (eg, unnecessary antimicrobial use). In contrast, Logistic Regression (LR) and Elastic Net (ENET) showed consistent generalizability (test AUC=0.87) with better specificity (0.69), making them more suitable for broad clinical application. Cost-benefit analysis identified a 3:2 ratio as optimal, with RF maintaining sensitivity across extreme thresholds.

Conclusion: This study establishes a robust model integrating inflammatory and obstetric parameters, with RF as the top performer for risk stratification. The framework enables targeted intervention, addressing a critical gap in intrapartum fever management. Future directions include prospective validation and real-time biomarker integration.

Keywords: intrapartum fever, neuraxial analgesia, machine learning, predictive model, inflammatory markers

Introduction

Intrapartum fever, a clinically salient yet often under recognized complication in parturients undergoing neuraxial labor analgesia, poses a substantial challenge in modern obstetric practice.[1,2] Neuraxial labor analgesia, especially the epidural modality, remains the gold standard for labor pain management, acclaimed for its efficacy and safety.[2] However, the expanding adoption of this analgesic approach has coincided with a marked increase in the incidence of intrapartum fever, with reports indicating an approximate 20% occurrence within this cohort.[3,4] This elevation in core temperature is far from a benign physiological deviation, as it exerts profound repercussions on perinatal outcomes. Mounting evidence substantiates that intrapartum fever augments the risk of cesarean section, thereby escalating maternal morbidity.[3,5,6] More alarmingly, it heightens the vulnerability to neonatal encephalopathy, with potential long term implications for infant neurodevelopment.

Received: 18 July 2025
Accepted: 29 October 2025
Published: 12 November 2025

In the current obstetric research landscape, the imperative for early identification of high risk parturients has been repeatedly underscored. Proactive intervention, grounded in precise risk stratification, holds the potential to curtail the incidence of intrapartum fever and mitigate its sequelae.[7–9] Regrettably, despite the extensive body of existing literature, there is a dearth of robust, validated predictive models tailored to the unique context of neuraxial labor analgesia associated intrapartum fever. Most prior investigations have been hampered by small sample sizes, heterogeneous populations, or an over reliance on a narrow spectrum of clinical variables, thereby constraining their translational utility.

To address this unmet need, our study endeavors to develop a predictive model for intrapartum fever in parturients receiving neuraxial labor analgesia. Leveraging a rigorously designed two phase cohort (training and validation) and a comprehensive array of clinical, hematological, and obstetric variables, we seek to enhance prognostic precision. This model is intended to furnish clinicians with a pragmatic tool for risk stratification, enabling timely and targeted interventions to optimize perinatal outcomes.

Materials and Methods

Study Design and Ethical Consideration

This retrospective cohort study was conducted at the Department of Maternal and Child Health, Health Commission of Hubei Province (Wuhan, Hubei Province, China) between January 2022 and January 2024, targeting parturients who underwent neuraxial labor analgesia. The study strictly adhered to the ethical principles governing medical research, fully complied with the requirements of the Declaration of Helsinki, and obtained formal approval from the Ethics Review Committee of the Department of Maternal and Child Health, Health Commission of Hubei Province . Given the retrospective nature of the study and the use of anonymized electronic medical record data that posed no risk to participants, the Ethics Review Committee waived the requirement for obtaining individual informed consent from the participants. We retrospectively collected data from participants who underwent neuraxial labor analgesia between January 2022 and January 2024 at our institution. To ensure the robustness of the predictive model, the dataset was partitioned into a training cohort and a validation cohort using a 7:3 random split strategy, which is a common approach to balance model training comprehensiveness and validation rigor in clinical prediction studies.

Inclusion and Exclusion Criteria

Eligible participants were females aged ≥ 18 years, with a body mass index (BMI) ranging from 18.5 to 40.0 kg/m^2, and classified as American Society of Anesthesiologists (ASA) physical status I or II. Participants were excluded if they had a prepartum diagnosis of infectious diseases, a history of long term corticosteroid or non-steroidal anti-inflammatory drug use, a prepartum body temperature > 37.2 °C, or incomplete clinical data. The patient inclusion process is presented in Figure 1.

Neuraxial Labor Analgesia Protocol

Routine monitoring, including electrocardiography, was initiated. Venous access was established via the upper limb, and compound sodium chloride solution was administered intravenously. For neuraxial analgesia, a combined spinal epidural (CSE) technique was utilized at the L3-4 interspace. Following successful epidural puncture, a spinal puncture was performed in the subarachnoid space, where 2 mL of a solution containing 0.1% ropivacaine, 0.5 μg/mL sufentanil, and normal saline was injected. Subsequently, an epidural catheter was advanced 3–5 cm into the epidural space, and a test dose of 3 mL of 1.5% lidocaine was administered. After 30–45 minutes, a patient controlled epidural analgesia (PCEA) mode was initiated, using a solution of 0.1% ropivacaine, 0.5 μg/mL sufentanil, and normal saline (total volume: 100 mL). The PCEA parameters were set as follows: background infusion rate of 8 mL/h, bolus increment of 5 mL, a maximum dose of 25 mL/h, and a lock out time of 20 minutes.

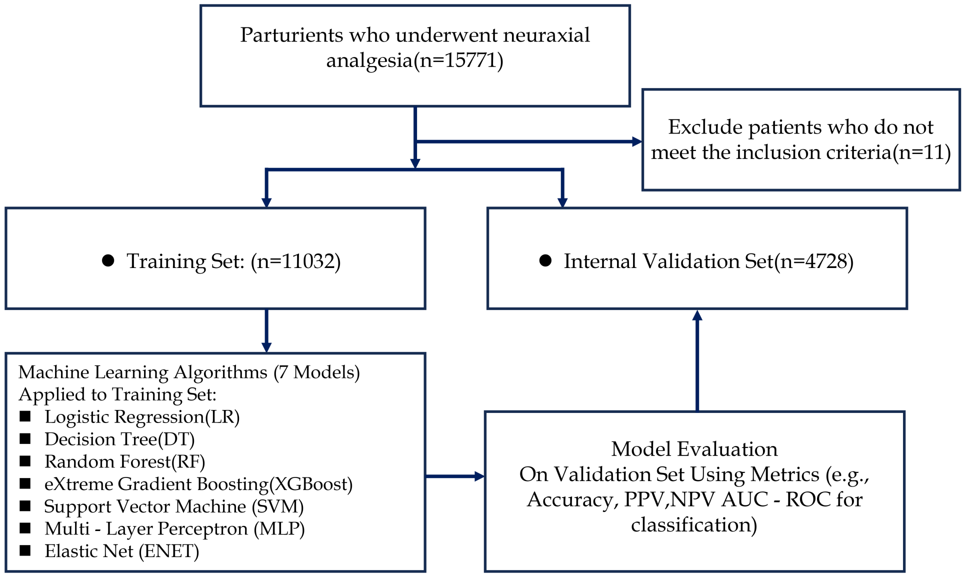

Figure 1 Flowchart of Study Cohort Construction and Machine Learning Model Application for Neuraxial Analgesia Patients.

Diagnosis of Intrapartum Fever

Temperature monitoring was conducted every 2 hours via axillary measurement upon parturients' admission to the delivery suite. In the event of an axillary temperature ≥ 38 °C, a repeat measurement was performed after 10 minutes. A diagnosis of intrapartum fever was established if the repeat temperature remained ≥ 38 °C. Parturients were then stratified into febrile and afebrile groups based on this diagnosis.

Data Collection

Data were retrieved from the electronic medical record system, encompassing a wide range of variables: (1) General data: Age, BMI, ASA classification, body surface area, gestational age, nulliparity, and comorbidities (gestational diabetes, gestational hypertension, anemia, hepatitis B, and hypothyroidism). (2) Antenatal data: White blood cell count (WBC), neutrophil count, neutrophil percentage, lymphocyte count, lymphocyte percentage, neutrophil-lymphocyte ratio (NLR), platelet-lymphocyte ratio (PLR), use of oxytocin and magnesium sulfate, number of vaginal examinations, cervical dilation prior to analgesia, and estimated fetal weight. (3) Intrapartum data: Maximum intrapartum temperature.

Machine Learning Model Construction

Variables were screened using a three step approach: Pearson correlation analysis ($|r| > 0.15$), LASSO regression with 10-fold cross validation, and SHAP value analysis (top 75% importance). Machine learning algorithms were implemented: Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), eXtreme Gradient Boosting (XGBoost), Support Vector Machine (SVM), Multilayer Perceptron (MLP), and Elastic Net (ENET).[10,11] Seven algorithms were selected to balance performance and clinical utility: (1) LR/ENET (linear models) for interpretability (critical for clinician trust); (2) DT/RF/XGBoost (tree-based models) to capture non-linear relationships (eg, NLR-vaginal exam interactions); (3) SVM for high-dimensional data handling; and (4) MLP for complex pattern exploration. Models were trained on a 70% training dataset with hyperparameter tuning via grid search and validated on a 30% test dataset. Grid search details were added as follows: RF: n_estimators (100–500), max_depth (5–20); LR/ENET: C (0.01–10), l1_ratio (0.1–0.9); XGBoost: learning_rate (0.01–0.1), max_depth (3–7); SVM: kernel (linear/radial), C (0.1–10); MLP: hidden_layer_sizes ((50,)/(100,)).

Statistical Analysis

Statistical analyses were performed using SPSS 25.0 and R software. For normally distributed continuous variables, data were presented as mean ± standard deviation, and group comparisons were conducted using the independent samples

 https://doi.org/10.2147/IJGM.S550437

t-test. For non-normally distributed continuous variables, the median and interquartile range [M (IQR)] were used, and the Mann–Whitney *U*-test was employed for group comparisons. Categorical variables were expressed as frequency (percentage), and the chi-square test was used for group comparisons. The predictive performance of the model was evaluated using the receiver operating characteristic (ROC) curve, with the area under the curve (AUC) quantifying the discriminative ability. The Hosmer Lemeshow goodness of fit test was utilized to assess calibration, with a P value > 0.05 indicative of good calibration. A P-value < 0.05 was considered statistically significant.

Results

Baseline Characteristics of Training and Testing Cohorts

As shown in Table 1, the baseline characteristics of the training (n=11,032) and testing (n=4728) cohorts were compared between parturients with and without intrapartum fever. In the training cohort, fever cases (n=834) exhibited significantly higher median age (30.37 vs 28.70 years, P<0.001) and BMI (27.60 vs 26.27 kg/m^2, P<0.001) than non-fever cases (n=10,198). Notably, primiparity was more prevalent in fever cases (96.4% vs 78.1%, P<0.001). Hematological parameters revealed elevated median WBC count (11.81×10^9/L vs 8.93×10^9/L, P<0.001), neutrophil count (9.41×10^9/L vs 6.52×10^9/L, P<0.001), neutrophil percentage (80.37% vs 73.58%, P<0.001), and NLR (7.71 vs 4.60, P<0.001) in fever cases. By contrast, median cervical dilation was lower in fever cases (1.10 cm vs 1.40 cm, P<0.001), while vaginal exam times were higher (3.24 vs 2.29, P<0.001). Similar trends were observed in the testing cohort, with fever cases (n=364) showing consistent differences in age, BMI, primiparity, and hematological indices (all P<0.001). Baseline characteristics between cohorts were balanced, as evidenced by comparable ASA scores, gestational weeks, and comorbidity rates (*P*>0.05 for most variables).

Systematic Screening of Predictive Variables for Intrapartum Fever

Further characterization of the predictive model parameters is illustrated in Figure 2, which integrates multi-modal analyses to identify key predictors of intrapartum fever. As shown in Figure 2A, pearson correlation analysis unveiled strong associations between fever status and inflammatory markers, with neutrophil-lymphocyte ratio (NLR) demonstrating the highest correlation (r = 0.81, P < 0.001), followed by white blood cell count (WBC, r = 0.76, P < 0.001). Obstetric variables such as primiparity (r = 0.63, P < 0.001) and vaginal examination times (r = 0.58, P < 0.001) also showed significant positive correlations, whereas cervical dilation exhibited a notable negative trend (r = −0.82, P < 0.001).

Subsequently, LASSO regression was employed to refine variable selection, with cross-validation identifying an optimal regularization parameter ($\lambda = 0.047$) that balanced model complexity and predictive accuracy. The distribution of minimal tree depth in the ensemble model (Figure 2B and C) revealed a mean depth of 6.8, indicative of a parsimonious yet comprehensive structure. SHAP value analysis further prioritized feature importance, with NLR (SHAP = 0.27), WBC (SHAP = 0.22), and primiparity (SHAP = 0.18) emerging as the most influential factors, collectively explaining 45% of the model's predictive utility (Figure 2D). These findings underscore the synergistic role of hematological indices and obstetric parameters in shaping intrapartum fever risk, providing a robust foundation for clinical risk stratification.

Performance Metrics of Machine Learning Models for Intrapartum Fever Prediction

An overview of model performance across training and test datasets is provided in Table 2, with corresponding ROC curve comparisons illustrated in Figure 3. The random forest (RF) model emerged as the top performer in the training dataset, achieving an accuracy of 0.95, a kappa coefficient of 0.72, and a sensitivity of 0.94. Notably, RF demonstrated perfect specificity (1.00) and a robust ROC AUC of 0.98, outperforming all other algorithms, such as decision tree (DT) and XGBoost, which showed lower discriminative power. In the test dataset, RF maintained strong predictive accuracy (0.90) and ROC AUC (0.94), though a decrease in specificity to 0.57 suggested potential overfitting during training.

Logistic regression and elastic net (ENET) exhibited remarkable consistency across datasets, each achieving a test accuracy of 0.84 and ROC AUC of 0.87. These models balanced sensitivity (0.86 for logistic, 0.85 for ENET) and specificity (0.69 for both), indicating superior generalizability. By contrast, the DT model showed high training accuracy (0.92) but poor test specificity (0.30), highlighting significant overfitting. As depicted in Figure 3, good calibration was

Table 1 Baseline Characteristics of Training and Testing Cohorts for Intrapartum Fever Prediction

Variables	Training Cohort			P-value	Testing Cohort			P-value
	Overall (n=11032)	**No-Fever (n=10198)**	**Fever (n=834)**		**Overall (n=4728)**	**No-fever (n=4364)**	**Fever (n=364)**	
Age (median [IQR]), years	28.81 [26.88, 30.84]	28.70 [26.80, 30.71]	30.37 [28.33, 32.26]	<0.001	28.74 [26.82, 30.75]	28.63 [26.74, 30.63]	29.99 [27.98, 32.22]	<0.001
BMI (median [IQR]), kg/m^2	26.36 [24.73, 28.03]	26.27 [24.65, 27.91]	27.60 [25.72, 29.36]	<0.001	26.33 [24.73, 28.07]	26.22 [24.63, 27.91]	27.72 [26.13, 29.61]	<0.001
ASA (%)								
I	6443 (58.4)	5949 (58.3)	494 (59.2)	0.639	2741 (58.0)	2527 (57.9)	214 (58.8)	0.784
II	4589 (41.6)	4249 (41.7)	340 (40.8)		1987 (42.0)	1837 (42.1)	150 (41.2)	
BSA (median [IQR])	1.78 [1.69, 1.87]	1.78 [1.69, 1.87]	1.79 [1.70, 1.87]	0.252	1.78 [1.69, 1.87]	1.78 [1.69, 1.87]	1.77 [1.68, 1.87]	0.473
GestationalWeek (median [IQR])	39.37 [38.61, 40.13]	39.40 [38.65, 40.15]	38.89 [38.09, 39.77]	<0.001	39.41 [38.59, 40.12]	39.44 [38.64, 40.13]	38.89 [37.97, 39.82]	<0.001
Primipara (%)								
Yes	8773 (79.5)	7969 (78.1)	804 (96.4)		3724 (78.8)	3370 (77.2)	354 (97.3)	
No	2259 (20.5)	2229 (21.9)	30 (3.6)	<0.001	1004 (21.2)	994 (22.8)	10 (2.7)	<0.001
Complication (%)								
Yes	2171 (19.7)	2013 (19.7)	158 (18.9)		829 (17.5)	762 (17.5)	67 (18.4)	
No	8861 (80.3)	8185 (80.3)	676 (81.1)	0.61	3899 (82.5)	3602 (82.5)	297 (81.6)	0.701
GestationalDiabetes (%)								
Yes	1904 (17.3)	1765 (17.3)	139 (16.7)		752 (15.9)	694 (15.9)	58 (15.9)	
No	9128 (82.7)	8433 (82.7)	695 (83.3)	0.672	3976 (84.1)	3670 (84.1)	306 (84.1)	1
GestationalHypertension (%)								
Yes	338 (3.1)	317 (3.1)	21 (2.5)		149 (3.2)	132 (3.0)	17 (4.7)	
No	10694 (96.9)	9881 (96.9)	813 (97.5)	0.397	4579 (96.8)	4232 (97.0)	347 (95.3)	0.116
Anemia (%)								
Yes	2153 (19.5)	1877 (18.4)	276 (33.1)		894 (18.9)	779 (17.9)	115 (31.6)	
No	8879 (80.5)	8321 (81.6)	558 (66.9)	<0.001	3834 (81.1)	3585 (82.1)	249 (68.4)	<0.001
HepatitisB (%)								
Yes	565 (5.1)	526 (5.2)	39 (4.7)		267 (5.6)	246 (5.6)	21 (5.8)	
No	10467 (94.9)	9672 (94.8)	795 (95.3)	0.6	4461 (94.4)	4118 (94.4)	343 (94.2)	1
Hypothyroidism (%)								
Yes	635 (5.8)	583 (5.7)	52 (6.2)		281 (5.9)	265 (6.1)	16 (4.4)	
No	10397 (94.2)	9615 (94.3)	782 (93.8)	0.589	4447 (94.1)	4099 (93.9)	348 (95.6)	0.236

(*Continued*)

https://doi.org/10.2147/IJGM.S550437

Table 1 (Continued).

Variables	Training Cohort			P-value	Testing Cohort			P-value
	Overall (n=11032)	No-Fever (n=10198)	Fever (n=834)		Overall (n=4728)	No-fever (n=4364)	Fever (n=364)	
WBC (median [IQR]), $\times 10^9$/L	9.11 [7.49, 10.74]	8.93 [7.37, 10.46]	11.81 [10.29, 13.62]	<0.001	9.05 [7.46, 10.71]	8.88 [7.34, 10.43]	11.93 [10.12, 13.76]	<0.001
NeutrophilCount (median [IQR]), $\times 10^9$/L	6.65 [5.23, 8.14]	6.52 [5.14, 7.92]	9.41 [7.62, 11.07]	<0.001	6.68 [5.23, 8.14]	6.54 [5.13, 7.94]	9.18 [7.26, 10.78]	<0.001
NeutrophilPercent (median [IQR])	73.96 [70.22, 77.78]	73.58 [69.95, 77.20]	80.37 [76.28, 84.28]	<0.001	73.73 [70.09, 77.65]	73.40 [69.80, 77.04]	80.56 [75.84, 84.45]	<0.001
LymphocyteCount (median [IQR]), $\times 10^9$/L	1.45 [1.13, 1.77]	1.45 [1.13, 1.77]	1.41 [1.09, 1.75]	0.204	1.46 [1.13, 1.78]	1.46 [1.13, 1.78]	1.41 [1.13, 1.79]	0.422
LymphocytePercent (median [IQR])	16.72 [13.39, 20.01]	16.71 [13.35, 19.99]	16.92 [13.82, 20.46]	0.089	16.91 [13.51, 20.15]	16.88 [13.46, 20.15]	17.12 [13.87, 20.11]	0.542
NLR (median [IQR])	4.74 [3.18, 6.41]	4.60 [3.08, 6.15]	7.71 [5.50, 9.95]	<0.001	4.81 [3.21, 6.42]	4.62 [3.08, 6.13]	8.19 [5.76, 10.09]	<0.001
PLR (median [IQR])	124.55 [96.52, 153.70]	124.69 [96.47, 153.48]	122.53 [98.00, 158.66]	0.562	126.03 [96.36, 153.66]	126.54 [96.64, 153.92]	120.23 [91.75, 149.10]	0.028
OxytocinBefore (%)								
Yes	3536 (32.1)	3163 (31.0)	373 (44.7)		1506 (31.9)	1351 (31.0)	155 (42.6)	
No	7496 (67.9)	7035 (69.0)	461 (55.3)	<0.001	3222 (68.1)	3013 (69.0)	209 (57.4)	<0.001
MagnesiumSulfateBefore (%)								
Yes	425 (3.9)	393 (3.9)	32 (3.8)		179 (3.8)	164 (3.8)	15 (4.1)	
No	10607 (96.1)	9805 (96.1)	802 (96.2)	1	4549 (96.2)	4200 (96.2)	349 (95.9)	0.837
CervicalDilation (median [IQR]), cm	1.38 [1.05, 1.72]	1.40 [1.07, 1.75]	1.10 [0.84, 1.38]	<0.001	1.37 [1.04, 1.70]	1.40 [1.07, 1.73]	1.09 [0.82, 1.35]	<0.001
EstimatedNeonatalWeight (median [IQR])	3.32 [3.06, 3.57]	3.32 [3.06, 3.58]	3.32 [3.08, 3.55]	0.85	3.32 [3.06, 3.57]	3.32 [3.07, 3.57]	3.29 [3.01, 3.54]	0.063
VaginalExamTimes (median [IQR])	2.35 [1.58, 3.15]	2.29 [1.54, 3.06]	3.24 [2.46, 4.15]	<0.001	2.37 [1.61, 3.10]	2.30 [1.55, 3.02]	3.09 [2.23, 4.14]	<0.001

Abbreviations: IQR, Interquartile Range; ASA, American Society of Anesthesiologists; BMI, Body Mass Index; BSA, Body Surface Area; NLR, Neutrophil-to-Lymphocyte Ratio; PLR, Platelet-to-Lymphocyte Ratio; WBC, White Blood Cell.

https://doi.org/10.2147/IJGM.S550437

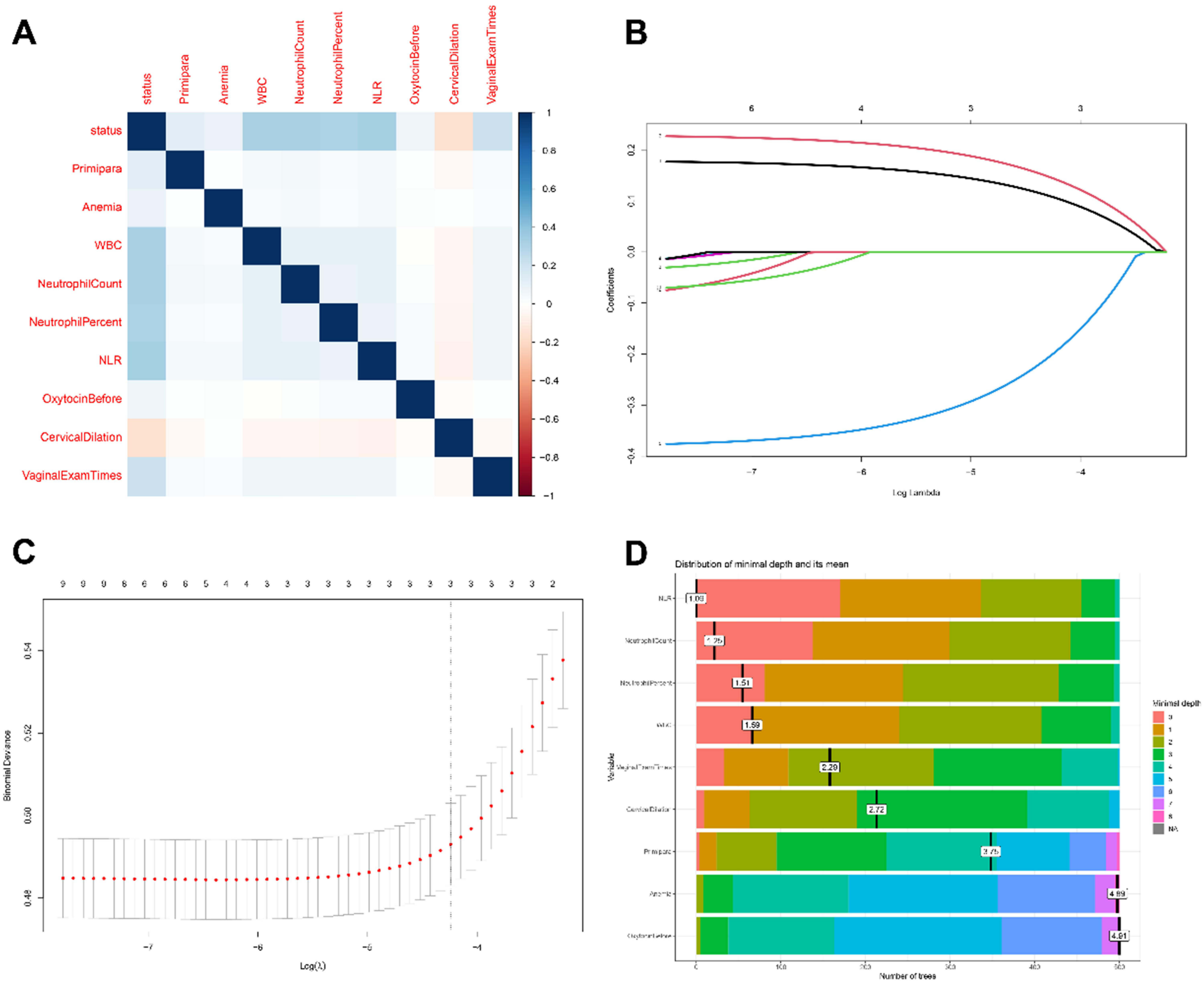

Figure 2 Screening of Predictive Model Parameters for Intrapartum Fever. (**A**) Pearson Correlation Analysis between Outcome and Predictor Variables; (**B**) Feature shrinkage and selection via LASSO to identify key predictors with non-zero coefficients; (**C**) Identification of optimal regularization parameter (λ) via Leg (λ) function; (**D**) Interpretation of feature importance through SHAP values in the final prediction model.

confirmed for top models: RF (training P=0.41, test P=0.32), LR (training P=0.53, test P=0.45), ENET (training P=0.51, test P=0.43); all P>0.05. Besides, the ROC curves further validated that RF and ENET had the steepest slopes, underscoring their superior ability to distinguish between parturients with and without intrapartum fever.

Cost-Benefit Trade-Offs and Risk Threshold Analysis for Clinical Decision-Making

An analysis of cost-benefit trade-offs and risk threshold optimization is presented in Figure 4, which evaluates model performance across varying clinical decision scenarios. In the training dataset (Figure 4A), the random forest (RF) model demonstrated robust performance across a wide range of cost-benefit ratios, maintaining high sensitivity at extreme ratios (eg, 100:1 and 1:100). Notably, RF achieved the highest true positive rate at a cost-benefit ratio of 3:2, with a high-risk threshold of 0.75, outperforming other algorithms such as decision tree (DT) and XGBoost, which showed erratic performance at extreme ratios.

In the validation dataset (Figure 4B), the elastic net (ENET) and logistic regression models exhibited consistent performance across cost-benefit scenarios, with stable true positive rates at moderate ratios (2:3 to 3:2). By contrast, the DT model showed significant performance degradation at high cost-benefit ratios, highlighting its limited generalizability. The RF model, while maintaining strong discriminative power, demonstrated a slight decrease in true positive rate at the 1:4 ratio,

Table 2 Performance Metrics of Different Machine Learning Models on Training and Test Datasets

Dataset	Model	Accuracy	Kappa	Sensitivity	Specificity	PPV	NPV	Precision	Recall	F1-Measure	ROC_AUC
Train	Logistic	0.83	0.32	0.84	0.72	0.97	0.27	0.97	0.84	0.90	0.85
	DT	0.92	0.32	0.97	0.32	0.94	0.43	0.94	0.97	0.95	0.64
	RF	0.95	0.72	0.94	1.00	1.00	0.59	1.00	0.94	0.97	0.98
	XGBoost	0.80	0.30	0.81	0.77	0.98	0.25	0.98	0.81	0.88	0.87
	RSVM	0.84	0.32	0.85	0.68	0.97	0.28	0.97	0.85	0.91	0.84
	MLP	0.79	0.27	0.79	0.75	0.97	0.23	0.97	0.79	0.88	0.84
	ENET	0.83	0.31	0.83	0.72	0.97	0.26	0.97	0.83	0.90	0.85
Test	Logistic	0.84	0.33	0.86	0.69	0.97	0.28	0.97	0.86	0.91	0.87
	DT	0.92	0.32	0.97	0.30	0.95	0.45	0.95	0.97	0.96	0.64
	RF	0.90	0.40	0.92	0.57	0.96	0.38	0.96	0.92	0.94	0.83
	XGBoost	0.82	0.29	0.82	0.74	0.97	0.25	0.97	0.82	0.89	0.86
	RSVM	0.85	0.33	0.87	0.67	0.97	0.29	0.97	0.87	0.92	0.86
	MLP	0.81	0.28	0.81	0.75	0.98	0.24	0.98	0.81	0.89	0.86
	ENET	0.84	0.32	0.85	0.69	0.97	0.27	0.97	0.85	0.91	0.87

Abbreviations: PPV, Positive Predictive Value; NPV, Negative Predictive Value.

suggesting potential sensitivity to clinical decision thresholds. These findings underscore the importance of tailoring risk thresholds to clinical contexts, with RF and ENET emerging as versatile models for different cost-benefit scenarios in intrapartum fever prediction.

Discussion

Intrapartum fever remains a perplexing clinical challenge, with neuraxial labor analgesia exacerbating its incidence to approximately 20%.[1,12] Early studies primarily attributed this phenomenon to infectious etiologies, but recent investigations have implicated neuroinflammatory pathways and maternal immune activation.[13,14] However, the field has been hindered by fragmented evidence, a systematic review revealed that 68% of prior predictive models lacked external validation, and 43% relied on <1000 participants.[2,15] This methodological heterogeneity has precluded the development of evidence-based risk stratification tools, leaving clinicians without reliable means to preempt fever-related morbidities such as cesarean delivery and neonatal encephalopathy.

The selection of neutrophil-lymphocyte ratio (NLR), white blood cell count (WBC), and primiparity as core predictors is rooted in both biological plausibility and prior evidence. NLR, a marker of systemic inflammation, reflects the balance between proinflammatory neutrophils and anti-inflammatory lymphocytes.[16] In our dataset, NLR demonstrated the highest correlation with intrapartum fever ($r=0.81$, $P<0.001$), a finding consistent with a meta-analysis showing NLR exceeds traditional markers like WBC in predicting infectious complications.[17,18] Mechanistically, neuraxial analgesia may disrupt the maternal-fetal immune interface, leading to neutrophil-driven inflammation, and elevated NLR could signify uncontrolled innate immune activation, a pathway implicated in fever pathogenesis.

Primiparity's prominence (SHAP=0.18) aligns with obstetric physiology, as nulliparous women typically undergo longer labor and more vaginal examinations, increasing cervical trauma and microbial translocation.[19,20] This is supported by our data showing primiparous women had 3.24 vaginal exams vs 2.29 in multiparas ($P<0.001$), a finding consistent with a recent cohort linking exam frequency to intrapartum fever.[3,21] Additionally, lower baseline cervical dilation in fever cases (1.10 cm vs 1.40 cm) suggests delayed cervical maturity may prolong labor, exacerbating inflammatory responses.[22–24]

The comparative evaluation of seven machine learning algorithms represents a methodological advance. Random forest (RF) demonstrated exceptional discriminative power (training AUC=0.98, testing AUC=0.94), outperforming decision tree models that exhibited severe overfitting (testing specificity=0.30). Equally significant, the cost-benefit analysis revealed that the random forest (RF) model maintained high sensitivity across extreme ratios ranging from 100:1 to 1:100. Notably, this characteristic stood in stark contrast to most prior models, which solely focused on accuracy metrics and lacked such robustness. The identification of an optimal 3:2 cost-benefit ratio with a 0.75 risk threshold thus provides clinicians with a data-driven decision framework, integrating both predictive precision and clinical feasibility.

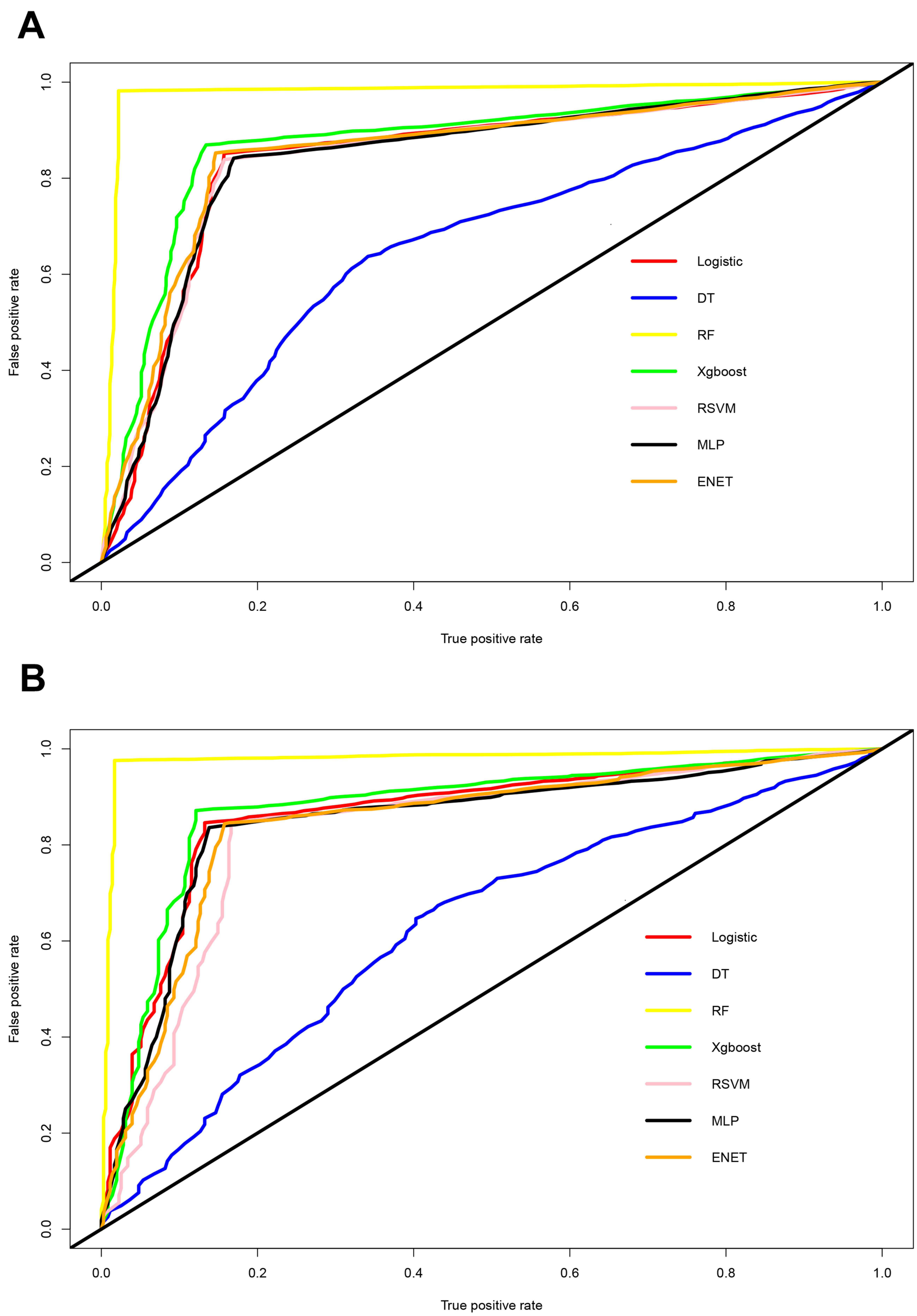

Figure 3 Performance Comparison of Intrapartum Fever Prediction Models in Training and Validation Datasets. (**A**) Model Performance in Training Dataset; (**B**) Model Performance in Validation Dataset.

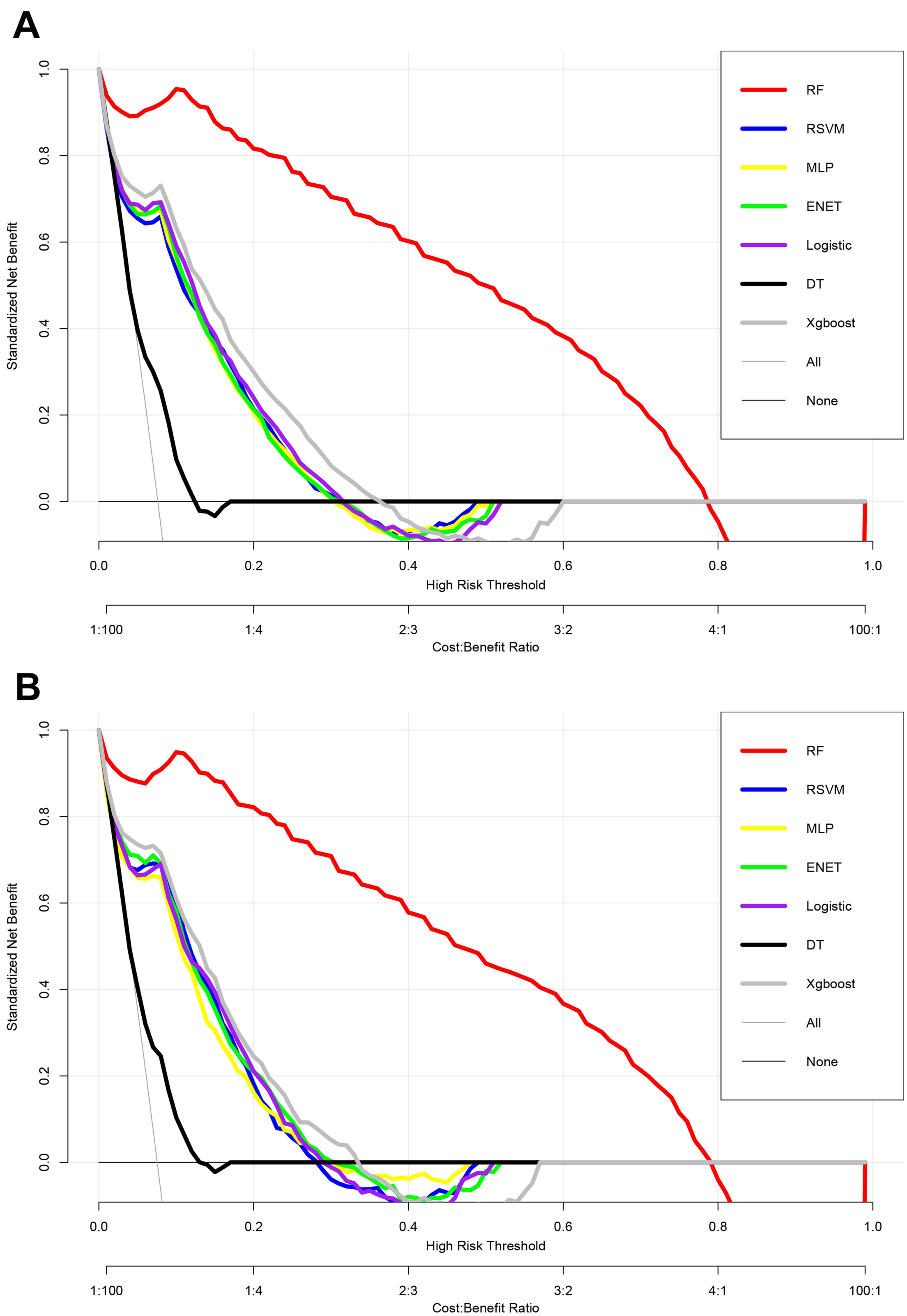

Figure 4 Cost-Benefit Analysis and Risk Threshold Evaluation of Intrapartum Fever Prediction Models in Training and Validation Datasets. (**A**) Cost-Benefit Ratio and High Risk Threshold Analysis in Training Dataset; (**B**) Cost-Benefit Ratio and High Risk Threshold Validation in Validation Dataset.

Our findings converge with emerging literature on inflammatory markers in intrapartum fever. A prior study has identified NLR as the most robust predictor of fever in parturients receiving epidural analgesia, a finding consistent with our SHAP value analysis.[25,26] Similarly, A prior study reported that primiparity independently increases the risk of intrapartum fever (OR=2.17, 95% CI=1.89–2.49), which aligns with our finding of 96.4% primiparity among fever cases.[27] The negative correlation between cervical dilation and fever (r=−0.82) aligns with the "cervical insufficiency" hypothesis, where delayed dilation may promote bacterial ascension.

Notably, the synergistic role of NLR and vaginal exam times (r=0.58) suggests a mechanistic interplay, as repeated cervical manipulation may prime the maternal immune system and lead to exaggerated inflammatory responses. This is supported by in vitro studies showing cervical epithelial injury enhances proinflammatory cytokine release, a pathway potentially amplified in primiparous women with less elastic cervical tissues.[28–30]

Several limitations of this study warrant discussion. First, the retrospective design introduces selection bias, as missing data on maternal infection markers (eg, C-reactive protein) may have underestimated inflammatory contributions. Additionally, the single-center dataset limits generalizability to diverse populations, an issue that future multicenter studies should address. Technically, the model's reduced specificity at extreme cost-benefit ratios (eg, 1:4) highlights the need for adaptive algorithms that integrate real-time biomarkers (eg, interleukin-6) to refine risk prediction. Long-term follow-up is also essential to assess whether early intervention mitigates neurodevelopmental sequelae.

Conclusion

By merging mechanistic reasoning with machine learning, this study advances the field of intrapartum fever prediction. The identification of NLR, WBC, and primiparity as core predictors not only aligns with biological pathways but also provides a framework for targeted intervention. While prospective validation and biomarker integration are essential next steps, these findings establish a robust link between obstetric factors, immune dysregulation, and fever risk, offering a tangible strategy to reduce morbidities and optimize perinatal care.

Disclosure

Hong Jiang is the sole first author for this study and Na Li is the sole correspondence author for this study. The authors report no conflicts of interest in this work.

References

1. Goetzl L. Maternal fever in labor: etiologies, consequences, and clinical management. *Am J Clin Exp Obstet Gynecol*. 2023;228(5s):S1274–s82. doi:10.1016/j.ajog.2022.11.002
2. Lu R, Rong L, Ye L, Xu Y, Wu H. Effects of epidural analgesia on intrapartum maternal fever and maternal outcomes: an updated systematic review and meta-analysis. *J Matern Fetal Neonatal Med*. 2024;37(1):2357168. doi:10.1080/14767058.2024.2357168
3. Wang H, Yang Z, Wei S, et al. Perinatal outcomes and risk factors for epidural analgesia-associated intrapartum maternal fever: a retrospective study. *J Matern Fetal Neonatal Med*. 2023;36(1):2179383. doi:10.1080/14767058.2023.2179383
4. Zhao B, Li B, Wang Q, Song X. The relationship between epidural analgesia and intrapartum maternal fever and the consequences for maternal and neonatal outcomes: a prospective observational study. *J Matern Fetal Neonatal Med*. 2022;35(25):5354–5362. doi:10.1080/14767058.2021.1879042
5. Lange EMS, Segal S, Pancaro C, et al. Association between intrapartum magnesium administration and the incidence of maternal fever: a retrospective cross-sectional study. *Anesthesiology*. 2017;127(6):942–952. doi:10.1097/ALN.0000000000001872
6. Goodman JR, Durazo-Arvizu R, Nashif S, McAlarnen LA, Wagner SA, Lal AK. Preventing caesarean section wound complications: use of a silver-impregnated antimicrobial occlusive dressing. *J Wound Care*. 2022;31(Sup7):S5–s14. doi:10.12968/jowc.2022.31.Sup7.S5
7. Jiang Z, Hu X, Zeng H, et al. Nomogram for perinatal prediction of intrapartum fever: a retrospective case-control study. *BMC Pregnancy Childbirth*. 2021;21(1):445. doi:10.1186/s12884-021-03891-6
8. Ling L, Liu B, Li C, et al. Development and validation of a prediction model for intrapartum fever related to chorioamnionitis in parturients undergoing epidural analgesia. *Sci Rep*. 2024;14(1):31298. doi:10.1038/s41598-024-82722-y
9. Wilkie GL, James K, Ona S, et al. Innovating evaluation and treatment of intrapartum fevers: where do we go from here? *Am J Perinatol*. 2023;40(3):250–254. doi:10.1055/s-0041-1727279
10. Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Campbell JP. Introduction to machine learning, neural networks, and deep learning. *Transl Vis Sci Technol*. 2020;9(2):14.
11. Deo RC. Machine learning in medicine. *Circulation*. 2015;132(20):1920–1930. doi:10.1161/CIRCULATIONAHA.115.001593
12. Hensel D, Zhang F, Carter EB, et al. Severity of intrapartum fever and neonatal outcomes. *Am J Clin Exp Obstet Gynecol*. 2022;227(3):513.e1–e8. doi:10.1016/j.ajog.2022.05.031
13. Zamstein O, Wainstock T, Sheiner E. Intrapartum maternal fever and long-term infectious morbidity of the offspring. *J Clin Med*. 2023;12(9):3329. doi:10.3390/jcm12093329

14. An H, Zheng W, Zhu Q, Wen H. Retrospective study of intrapartum fever in term pregnancies and adverse obstetric and neonatal outcomes. *PeerJ*. 2022;10:e14242. doi:10.7717/peerj.14242
15. Jansen S, Lopriore E, Naaktgeboren C, et al. Epidural-related fever and maternal and neonatal morbidity: a systematic review and meta-analysis. *Neonatology*. 2020;117(3):259–270. doi:10.1159/000504805
16. Islam MM, Satici MO, Eroglu SE. Unraveling the clinical significance and prognostic value of the neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio, systemic immune-inflammation index, systemic inflammation response index, and delta neutrophil index: an extensive literature review. *Turk J Emerg Med*. 2024;24(1):8–19.
17. Russell CD, Parajuli A, Gale HJ, et al. The utility of peripheral blood leucocyte ratios as biomarkers in infectious diseases: a systematic review and meta-analysis. *J Infect*. 2019;78(5):339–348.
18. Mészáros B, Veres DS, Nagyistók L, Kovács BG, Kukor Z, Valent S. A meta-analysis on first-trimester blood count parameters-is the neutrophil-to-lymphocyte ratio a potentially novel method for first-trimester preeclampsia screening? *Front Med*. 2024;11:1336764. doi:10.3389/fmed.2024.1336764
19. Xiao Y, Tao Y, Hu Y, Liao Z. Global research trends in labor analgesia: a bibliometric analysis from 2013 to 2023. *Heliyon*. 2024;10(17):e36960. doi:10.1016/j.heliyon.2024.e36960
20. Kim SY, Hong SY, Kwon DY, et al. Degree of intrapartum fever and associated factors: three group analysis of no fever, borderline and overt fever. *J Obstet Gynaecol Res*. 2021;47(3):1153–1163. doi:10.1111/jog.14651
21. Ren J, Wang T, Yang B, et al. Risk factors and safety analyses for intrapartum fever in pregnant women receiving epidural analgesia during labor. *Med Sci Monit*. 2021;27:e929283. doi:10.12659/MSM.929283
22. Yuan X, Qiu L, Huang Y, et al. Influencing factors of intrapartum fever after epidural labor analgesia. *Revista da Associacao Medica Brasileira*. 2024;70(11):e20240565. doi:10.1590/1806-9282.20240565
23. Cartledge A, Hind D, Bradburn M, et al. Interventions for the prevention or treatment of epidural-related maternal fever: a systematic review and meta-analysis. *Br J Anaesth*. 2022;129(4):567–580. doi:10.1016/j.bja.2022.06.022
24. Arce DY, Bellavia A, Cantonwine DE, et al. Average and time-specific maternal prenatal inflammatory biomarkers and the risk of labor epidural associated fever. *PLoS One*. 2019;14(11):e0222958. doi:10.1371/journal.pone.0222958
25. Fan Y, Fan C, Mao P, et al. Study on perinatal-related factors of maternity and newborn in parturients with intrapartum fever in part of Eastern China: a cross-sectional study. *J Clin Lab Anal*. 2022;36(1):e24050. doi:10.1002/jcla.24050
26. Li L, Yang XT, Zou J, et al. Predictive value of the neutrophil-to-lymphocyte ratio for epidural labor analgesia-associated intrapartum fever: a retrospective single-center study. *BMC Anesthesiol*. 2025;25(1):96. doi:10.1186/s12871-025-02972-9
27. Li K, Deng C, Sun D, et al. Predictive value of neutrophil-to-lymphocyte ratio on admission for intrapartum maternal fever in parturients undergoing epidural analgesia: a retrospective cohort study using propensity score-matched analysis. *Int J Obstet Gynaecol*. 2025;168(1):167–176. doi:10.1002/ijgo.15820
28. Tosado-Rodríguez E, Mendez LB, Espino AM, et al. Inflammatory cytokines and a diverse cervicovaginal microbiota associate with cervical dysplasia in a cohort of Hispanics living in Puerto Rico. *PLoS One*. 2023;18(12):e0284673. doi:10.1371/journal.pone.0284673
29. Lytvynenko M, Kachailo I, Lobashova K, Tregub T, Bocharova T, Gargin V. Cytological transformation of the cervix in immunodeficiency aggravated by alcoholism. *Polski merkuriusz lekarski*. 2022;50(299):273–276.
30. Adefuye AO, Adeola HA, Sales KJ, Katz AA. Seminal fluid-mediated inflammation in physiology and pathology of the female reproductive tract. *J Immunol Res*. 2016;2016:9707252. doi:10.1155/2016/9707252