BMC Pregnancy and Childbirth

RESEARCH

Open Access

Development and validation of a machine learning model for predicting intrapartum fever using pre-labor analgesia clinical indicators: a multicenter retrospective study

Bo Liu[1†], Liang Ling[2†], Fei Jia[1†], Dayuan Wei[2], Huiru Li[1], Yuanling Li[2], Hongquan Xiao[2], Mengqiao Wang[3], Chunping Li[4], Gang Zhang[2] and Jian Zhang[2*]

Abstract

Background Labor anesthesia is commonly used for pain relief during labor, but it can increase the risk of intrapartum fever. Currently, there are no reliable tools to predict which parturients might develop fever before labor anesthesia. The prediction model we developed aims to predict the incidence of intrapartum fever before labor analgesia.

Methods This study retrospectively analyzed the clinical data of parturients who underwent labor analgesia at Chengdu Jinjiang District Maternal & Child Health Hospital and Sichuan Jinxin Xinan Women's & Children's Hospital from January 2021 to June 2023. After the data were processed, the parturients were randomly divided into training and validation cohorts at an 8:2 ratio. The least absolute shrinkage and selection operator method was used for feature selection. Six machine learning models were developed and subjected to comprehensive analysis to assess and validate their predictive capabilities, ultimately selecting the best-performing model.

Results The study included a total of 5,052 parturients, with 418 (8.27%) parturients experiencing intrapartum fever. The predictive factors were primiparity, estimated neonatal weight, degree of uterine dilatation, presence of anemia, number of vaginal examinations, and height. The multilayer perceptron model emerged as the best-performing predictive model, achieving an area under the curve of 0.707, a sensitivity of 0.753, and a specificity of 0.584.

Conclusions The multilayer perceptron model, utilizing readily available pre-labor analgesia variables, demonstrates potential for predicting intrapartum fever. In comparison to existing tools, this model may enable earlier identification of high-risk parturients, supporting timely interventions and potentially enhancing maternal and neonatal health outcomes.

Keywords Intrapartum fever, Labor analgesia, Predictive models, Machine learning

†Bo Liu, Liang Ling, and Fei Jia contributed equally to this work.

*Correspondence:
Jian Zhang
anesethesiology@zju.edu.cn
Full list of author information is available at the end of the article

Introduction

Combined spinal–epidural anesthesia is recognized as the most effective and safe method of labor analgesia [1–4]. The popularity of combined spinal–epidural anesthesia is accompanied by a significant increase in the prevalence of intrapartum fever, which is reported to be approximately 6.4% to 33.2% in the literature [5]. Intrapartum fever increases the incidence of cesarean section, antibiotic treatment, low Apgar scores, neonatal intensive care unit (NICU) admission and neonatal encephalopathy [6–10]. At present, the mechanisms of combined spinal–epidural anesthesia-related maternal fever is unclear, and effective measures to treat intrapartum fever are not available. Early identification of high-risk parturients and early intervention could reduce the occurrence of intrapartum fever and adverse effects on parturients and newborns [11–13]. Therefore, identifying high-risk parturients as early as possible is clinically important.

Statistical models have been used to predict whether parturients will develop a fever during labor [14–17]. Jiang et al. sought to build a predictive model for intrapartum fever [14]. However, the included subjects were parturients who did not receive labor analgesia, and those who underwent cesarean section were excluded. Of note, Jiang et al. identified some predictors of the occurrence of maternal fever both before and after labor analgesia. Sun et al. used the pulse perfusion index to predict intrapartum fever [15], but the clinical indicators measured before the start of labor analgesia were not available. Some studies have attempted to use the white blood cell (WBC) count on admission to predict intrapartum fever, but the area under the curve (AUC) was 0.55 [16]. These limitations highlight a significant gap in the field: there is currently no effective predictive model based solely on pre-labor analgesia indicators. Most existing models rely heavily on intrapartum data, which restricts their application for early interventions.

Our study aims to develop and validate a machine learning model to predict the risk of intrapartum fever based on pre-labor analgesia clinical indicators. Compared to traditional methods, this model offers three key advantages: Firstly, it enables early identification of high-risk parturients, allowing for personalized interventions; secondly, it uses pre-labor data to facilitate timely decision-making and prevent complications; thirdly, it focuses on easily accessible variables, improving practicality and adaptability across diverse medical settings. By leveraging advanced machine learning techniques, the model captures complex relationships and significantly enhances prediction accuracy, supporting better maternal and neonatal health outcomes.

Methods

Study design

This study was a retrospective cohort design, analyzing the clinical data of parturients who received labor analgesia at Chengdu Jinjiang District Women & Children Health Hospital and Sichuan Jinxin Xinan Women & Children's Hospital from January 2021 to June 2023.

Participants

A total of 4038 parturients who received combined spinal–epidural labor anesthesia at Jinjiang District Women & Children Health Hospital from January 2021 to March 2022 and 1014 parturients who received combined spinal–epidural labor anesthesia at Jinxin Xinan Women & Children's Hospital from January 2023 to June 2023 were included. Participants were selected from the electronic medical records system based on the following exclusion criteria: a prenatal diagnosis of infectious disease, long-term steroid use or a history of prenatal nonsteroidal antipyretic analgesic use, a basal body temperature greater than 37.2 °C and incomplete clinical data.

Variables and measurements

The following variables were collected as potential predictors of intrapartum fever: age (years), height (cm), weight (kg), body mass index (BMI, kg/m^2), body surface area (m^2), gestational week (weeks), primiparity, gestational diabetes mellitus (GDM), pregnancy-induced hypertension, anemia, hepatitis B, hypothyroidism, maximum body temperature during labor (°C), premature rupture of membranes (PROM), and pre-labor analgesia data such as the use of oxytocin and MgSO4, the number of vaginal examinations, and the degree of cervical dilatation before analgesia (cm). Laboratory test results included WBC count ($\times 10^9$/L), neutrophil (NEUT) count ($\times 10^9$/L), NEUT (%), lymphocyte (LYM) count ($\times 10^9$/L), LYM (%), NLR (Lymphocyte Ratio), and PLR (Platelet–Lymphocyte Ratio). Estimated fetal weight before delivery (g) was also recorded.

Outcomes

The primary outcome of this study is the AUC of the model, which is used to evaluate the overall predictive performance of the model at different thresholds, reflecting the model's ability to predict fever during labor.

In order to comprehensively evaluate the performance of the model, we also focus on the following secondary outcome indicators: sensitivity, positive predictive value (PPV), negative predictive value (NPV), and F1 score. By analyzing these indicators, we can gain a deeper understanding of the overall performance of the model and provide more comprehensive support for clinical decision-making.

Statistical analysis

Statistical analyses were performed via SPSS software version 25.0 and R4.3.1. The Kolmogorov–Smirnov test was used to evaluate the normal distribution of continuous variables. For normally distributed continuous variables, data are presented as the mean ± standard deviation (SD) and were analyzed via the independent-sample t test. Nonnormally distributed continuous variables are expressed as the median (M) and interquartile range (IQR). Categorical variables are described as percentages (%), with group comparisons conducted using the chi-square test. A P-value of less than 0.05 was considered statistically significant.

To ensure dataset integrity, cases with missing variables were excluded based on pre-defined criteria. This approach was chosen due to the minimal proportion of missing data, avoiding potential bias from imputation and ensuring reliable analysis. The study cohort was randomly divided into a training set ($n=4041$) and a test set ($n=1011$) at an 8:2 ratio. In machine learning, it is common to use 80% of the data for training and 20% for validation to ensure the model learns effectively while allowing unbiased performance evaluation [18]. Within the training set, the least absolute shrinkage and selection operator (LASSO) algorithm was employed to identify candidate predictors. Six machine learning models were subsequently developed: support vector machine (SVM), logistic regression (LR), multilayer perceptron (MLP), random forest (RF), extreme gradient boosting (XGB), and decision tree (DT). Optimal hyperparameters for each model were determined using tenfold cross-validation. Model performance was independently validated via the test set. The model with the highest predictive accuracy was selected on the basis of the area under the receiver operating characteristic (ROC) curve (AUC). In this study, we utilized ROC curve analysis in R4.3.1 to determine the optimal cutoff point for each model. By assessing the sensitivity and specificity at various thresholds, we calculated the Youden index for each threshold. The threshold with the highest Youden index was selected as the optimal cutoff point, allowing us to achieve the best balance between sensitivity and specificity. This approach ensures maximal identification of truly high-risk individuals in our classification task while minimizing false positives to the greatest extent possible. The final model's predictive performance and clinical utility were then assessed, and variable importance was quantified. A detailed outline of the analytical process is illustrated in Fig. 1.

We assume that the incidence of intrapartum fever is 20%. This prediction model includes 25 predictors, each requiring at least 10 participants [19]. Therefore, we need at least $(25\times10\ /\ 0.2)=1250$ participants. To ensure sufficient events for the separate validation cohort, the planned sample size was increased to a minimum of 5000 participants.

The assumed 20% incidence of intrapartum fever in our sample size calculation was based on the wide range of prevalence reported in the literature, which varies from 6.4% to 33.2%. To provide a balanced and conservative

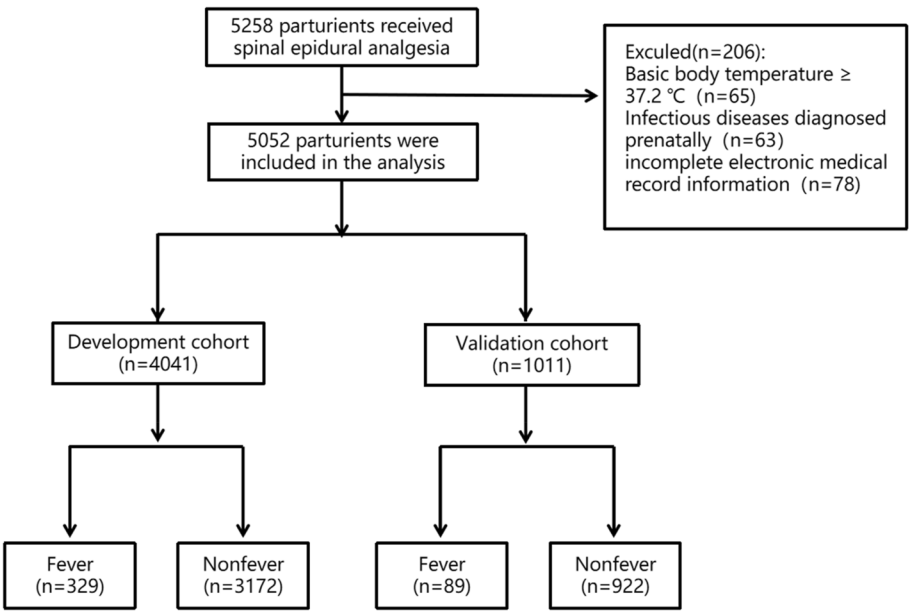

Fig. 1 Flowchart of the machine learning model: This figure illustrates the selection process of study participants

estimate, we selected an approximate average value for the calculation.While the observed incidence of 8.27% in our study was lower than the assumed 20%, this discrepancy may be attributed to the specific characteristics of our study population.

Ethical considerations

This study was approved by the Ethics Committee of Chengdu Jinjiang District Women & Children Health Hospital (approval number: 2024 trial (20)) and Sichuan Jinxin Xinan Women & Children's Hospital (approval number: 2024 trial (2)). Given the retrospective nature of the study, which involved the analysis of anonymized data without direct patient intervention, the ethics committees waived the requirement for informed consent. The justification for this waiver is based on the minimal risk posed to participants, as the data collected were de-identified and stored securely to protect patient confidentiality. There were no additional risks identified related to data usage, as the study strictly adhered to data protection regulations and ethical guidelines.All procedures were conducted following relevant ethical guidelines and regulations, in line with the principles outlined in the World Medical Association's Declaration of Helsinki [20].

Results

A total of 5258 parturients were initially enrolled in this study. 65 parturients were excluded due to a prenatal body temperature ≥ 37.2 °C. 63 parturients were excluded because of a prenatal diagnosis of infectious disease, 78 parturients were excluded because of incomplete electronic medical records, and 5052 parturients were eligible for the study. A total of 5052 participants were randomly divided into a development cohort ($n=4041$) and a validation cohort ($n=1011$) at an 8:2 ratio. A total of 4041 parturients were included in the development cohort (Fig. 1). The overall incidence rate of intrapartum fever among the participants was 8.27% (418/5052).

Baseline data analysis

Univariate analysis revealed that several factors were significantly associated with intrapartum fever. Parturients with intrapartum fever had higher BMI ($P=0.030$), estimated neonatal weight ($P<0.001$), and the number of vaginal examinations ($P<0.001$). The proportions of primiparity ($P<0.001$) and anemia ($P<0.001$) were also significantly higher in the fever group compared to the no-fever group. In contrast, parturients with intrapartum fever were significantly shorter in height ($P=0.021$) and had lower degrees of cervical dilatation before labor analgesia ($P<0.001$). These findings highlight potential clinical risk factors that may contribute to the development of intrapartum fever (Table 1).

LASSO regression for selecting risk factors

LASSO regression applies a penalty function to shrink variable coefficients, setting some to zero for variable selection. This penalty is controlled by the hyperparameter λ. Using tenfold cross-validation, we identified the λ value that minimizes the mean squared error (MSE) (left dashed line in Fig. 2A) and the λ value one standard error away from the minimum MSE (right dashed line in Fig. 2A). We selected the latter as the optimal λ value to achieve a more parsimonious predictive model. Figure 2B shows the LASSO coefficient paths for 25 features associated with intrapartum fever across different log λ values, with each curve representing one feature. As the penalty increases, model compression intensifies, enhancing the selection of important variables. On the basis of the optimal λ determined in Fig. 2A, six predictors were ultimately identified: height, primiparity, anemia, number of vaginal examinations, size of uterine dilatation, and estimated neonatal weight. These factors align with the high-risk factors identified in the baseline data analysis.

Building machine learning models

To optimize the hyperparameters of the six machine learning models, we employed a tenfold cross-validation approach. In this process, each model underwent systematic validation and evaluation. Specifically, the dataset was divided into ten mutually exclusive subsets, with nine subsets used for training and the remaining subset used for testing in each iteration. By calculating the AUC values from the test results of all ten subsets, we obtained the performance of the models under different hyperparameter configurations. Ultimately, the optimal hyperparameter configuration for each model was determined on the basis of the average of these AUC values, allowing us to identify the configuration that exhibited the best overall predictive performance (Fig. 2C). This approach ensures the stability and generalizability of the models.

Model validation

The trained models were applied to the validation cohort. The AUC values of the ROC curves for the six models, ranked from highest to lowest, were as follows: MLP (0.707 [95% CI, 0.658–0.7516), LR (0.695 [95% CI, 0.642–0.747]), XGB (0.662 [95% CI, 0.609–0.714]), DT (0.629 [95% CI, 0.580–0.679]), RF (0.592 [95% CI, 0.528–0.656]), and SVM (0.580 [95% CI, 0.525–0.634]) (Fig. 3). The DeLong test was used to compare the AUC values of the ROC curves for the MLP model and the other five models. The specific comparison results are indicated in the lower right corner of Fig. 3. According to the results, there was no statistically significant difference in the AUC values among the MLP, LR and XGB models ($P>0.05$). However, the MLP model demonstrated

Table 1 Baseline Data Analysis of the Study Subjects

Variable	Fever ($n=329$)	No fever ($n=3712$)	*P* value
Age (years)	28.90 ± 4.04	28.99 ± 3.84	0.710
Height (cm)	158.93 ± 5.04	159.60 ± 4.88	**0.021**
Weight (kg)	66.79 ± 8.68	66.36 ± 8.29	0.390
BMI (kg/m^2)	26.43 ± 3.13	26.04 ± 2.94	**0.030**
Body surface area (m^2)	1.80 ± 0.13	1.80 ± 0.13	0.925
Gestational age (w)	39.44 ± 1.35	39.32 ± 1.27	0.151
WBC count (10^9/L)	9.30 ± 2.78	9.07 ± 2.33	0.137
LYM count (10^9/L)	1.51 ± 0.44	1.54 ± 0.44	0.327
LYM (%)	17.07 ± 5.36	17.61 ± 5.49	0.085
NEUT count (10^9/L)	7.08 ± 2.54	6.83 ± 2.14	0.083
NEUT (%)	74.79 ± 6.93	74.53 ± 6.36	0.505
PLR (%)	126.72 ± 51.94	121.39 ± 52.32	0.076
NLR (%)	5.03 ± 2.30	4.83 ± 2.57	0.136
Estimated neonatal weight (g)	3281.31 ± 388.48	3208.59 ± 374.65	**< 0.001**
Primiparity (%)	309 (93.92)	2940 ((79.20)	**< 0.001**
GDM (%)	66 (20.02)	836 (22.52)	0.338
Hypertension during pregnancy (%)	16 (4.86)	138 (3.72)	0.374
Anemia (%)	115 (34.95)	858 (23.11)	**< 0.001**
Hepatitis B (%)	18 (5.47)	168 4.53)	0.518
Hypothyroidism (%)	31 (9.42)	294 (7.92)	0.393
PROM (%)	133 (40.43)	1352 (36.42%)	0.166
Size of uterine dilatation[a]	1.22 ± 0.65	1.46 ± 0.94	**< 0.001**
Number of vaginal examinations[a]	2.98 ± 1.69	2.52 ± 1.28	**< 0.001**
Oxytocin use[a]	140 (42.55)	1403 (37.80)	0.100
$MgSO_4$ use[a]	18 (5.47)	225 (6.06)	0.756

[*] These variables were obtained before labor analgesia

Abbreviations: BMI body mass index, *WBC* white blood cell, *LYM* lymphocyte, *NEUT* neutrophil, *PLR* platelet–lymphocyte ratio, *NLP* lymphocyte ratio, *GDM* gestational diabetes, *PROM* premature rupture of membranes

greater sensitivity, negative predictive values (NPV), positive predictive values (PPV), and F1 scores than the LR and XGB models did on the test set (Table 2). In summary, the predictive model established using the MLP algorithm exhibited the best clinical performance.

The MLP model demonstrated the best performance in this study. This type of artificial neural network excels at capturing complex nonlinear relationships within data, making it particularly suitable for complex medical datasets. In combination with multiple layers of neurons, the MLP processes input data through weighted transformations, enabling it to learn and identify patterns more effectively than linear models do. Unlike simpler models, the MLP can concurrently learn from the subtle interactions among a variety of predictive variables, aiding in the identification of risk factors that may lead to intrapartum fever. The backpropagation algorithm used in the MLP iteratively optimizes weights, enhancing the model's

(See figure on next page.)

Fig. 2 **A** LASSO regression cross-validation plot: LASSO regression cross-validation plot showing the binomial deviance across different log(λ) values. Red points represent the mean deviance, and the error bars indicate standard error (SE). Vertical dashed lines denote the selected optimal λ values. **B** LASSO regression variable shrinkage plot: LASSO regression coefficient shrinkage plot illustrating the trend of coefficients of various features as λ increases. Each curve represents a feature, demonstrating how the coefficients shrink towards zero with increasing penalty. **C** AUC values of the 6 models in tenfold cross-validation:Comparison of AUC values of six machine learning models in ten-fold cross-validation, illustrating the predictive performance of each model. Color coding is used to differentiate models, with higher AUC values indicating better predictive capability.Abbreviations: LASSO Least Absolute Shrinkage and Selection Operator, LR Logistic Regression, SVM Support Vector Machine, MLP Multilayer Perceptron, RF Random Forest XGB Extreme Gradient Boosting, DT Decision Tree, AUC Area Under the Receiver Operating Characteristic Curve

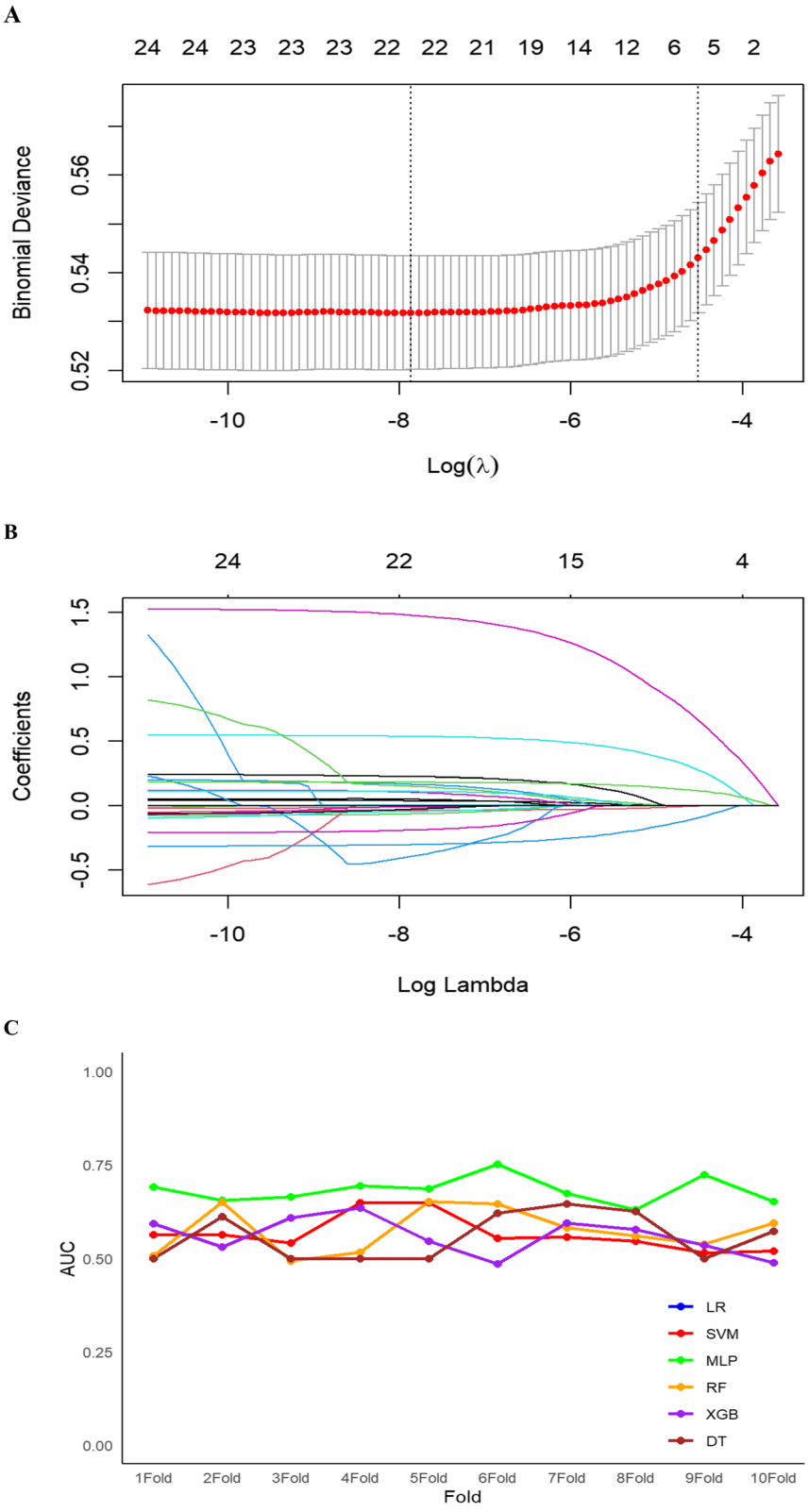

Fig. 2 (See legend on previous page.)

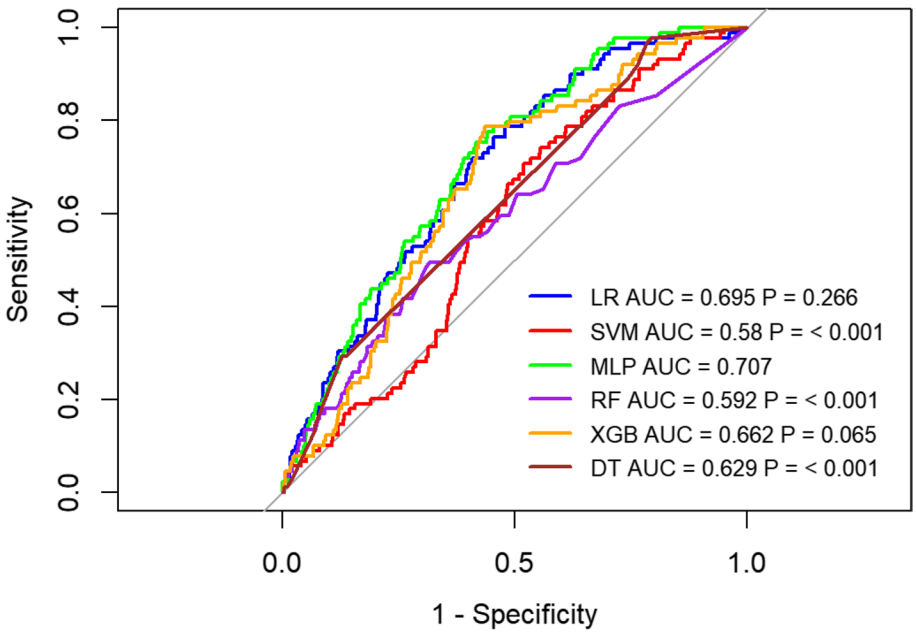

Fig. 3 ROC curves of the six models obtained using the validation cohort: The figure displays the ROC curves of six different machine learning models evaluated on the validation cohort to predict the occurrence of intrapartum fever, along with the AUC values for each model. Abbreviations: LR Logistic Regression, SVM Support Vector Machine, MLP Multilayer Perceptron, RF Random Forest, XGB Extreme Gradient Boosting, DT Decision Tree, AUC Area Under the Receiver Operating Characteristic Curve

Table 2 Comparison of the Overall Predictive Performance Across Six Models

	LR	SVM	MLP	RF	XGB	DT
Sensitivity (95% CI)	0.629 (0.521–0.727)	0.045 (0.010–0.094)	0.753 (0.660–0.835)	0.326 (0.237–0.419)	0.494 (0.390–0.600)	0.292 (0.195–0.391)
Specificity (95% CI)	0.640 (0.607–0.670)	0.983 (0.974–0.990)	0.584 (0.553–0.614)	0.807 (0.780–0.832)	0.707 (0.677–0.736)	0.864 (0.842–0.885)
NPV (95% CI)	0.947 (0.929–0.963)	0.914 (0.896–0.931)	0.961 (0.944–0.975)	0.925 (0.906–0.942)	0.935 (0.917–0.954)	0.927 (0.910–0.944)
PPV (95% CI)	0.144 (0.110–0.183)	0.200 (0.045–0.400)	0.149 (0.116–0.183)	0.140 (0.097–0.187)	0.140 (0.104–0.177)	0.172 (0.113–0.234)
Accuracy (95% CI)	0.639 (0.609–0.667)	0.900 (0.882–0.919)	0.599 (0.569–0.629)	0.764 (0.737–0.790)	0.688 (0.658–0.717)	0.814 (0.789–0.837)
F1 Score (95%CI)	0.235 (0.183–0.288)	0.073 (0.019–0.147)	0.249 (0.200–0.296)	0.196 (0.140–0.256)	0.218 (0.165–0.270)	0.217 (0.145–0.285)

Abbreviations: NPV Negative Predictive Value, *PPV* Positive Predictive Value

accuracy and robustness. Consequently, the MLP exhibits its superior predictive ability for identifying pregnant women at high risk for intrapartum fever.

Evaluation of the model

In the calibration curve of the validation cohort, the MLP model is close to the diagonal line (Fig. 4A), indicating that the predicted probabilities are relatively consistent with the actual observed probabilities. With respect to the curve shape, the curve of the MLP model features relatively favorable extensibility and ductility. This implies that the model can maintain a relatively stable performance in the context of diverse prediction scenarios. In the decision curve analysis (DCA) (Fig. 4B), the MLP model has a certain net benefit within the entire threshold range, indicating that this model can provide some value in prediction. Compared with other models, the MLP model shows advantages in terms of certain thresholds and cost – benefit ratios. Overall, the MLP model shows good performance in DCA, highlighting

its potential application in clinical decision-making. The SHapley Additive exPlanations (SHAP) variable importance plot (Fig. 4C) sequentially displays the ranking of risk factors: primiparity > estimated neonatal weight > size of uterine dilatation > anemia > number of vaginal examinations > height. In this plot, a deeper purple color indicates higher variable values, whereas a deeper yellow color indicates lower variable values. All six risk factors were positively correlated with the outcome of intrapartum fever.

Discussion

In this retrospective study, we developed and externally validated an MLP model to predict the risk of intrapartum fever prior to labor analgesia. By utilizing six easily obtainable variables identified through LASSO regression—primiparity, estimated neonatal weight, degree of uterine dilatation, presence of anemia, number of vaginal examinations, and height—our model exhibited strong potential in predicting the occurrence of intrapartum fever before labor analgesia.Overall, our findings highlight the potential utility of the MLP model in identifying patients at greater risk of intrapartum fever, thereby enabling timely interventions and improved patient care.

Identifying high-risk cases of intrapartum fever before labor analgesia is crucial. Early detection allows healthcare teams to intervene promptly, improving maternal and neonatal care. Since intrapartum fever can lead to negative outcomes such as cesarean delivery, low Apgar scores, and NICU admissions, it is crucial to identify these high-risk individuals in a timely manner [6]. By implementing personalized monitoring and early interventions, providers can closely observe high-risk parturients and take appropriate actions when necessary to reduce potential complications.

Traditional models often identify fever risk after labor analgesia because they rely on real-time data available during labor, which limits the timeliness of risk identification and may miss the optimal intervention window [14, 21]. the MLP model developed in this study leverages six easily obtainable pre-labor variables to predict risks before labor analgesia. This forward-looking capability enables clinicians to develop personalized management plans early in labor, enhancing both the timeliness and effectiveness of interventions. Unlike traditional models, which assume linear relationships between variables, the MLP model captures complex nonlinear interactions using multiple layers of neurons, significantly improving predictive accuracy. Additionally, the model's simplicity in requiring only six variables makes it practical for resource-limited settings.

The selection of predictive indicators in this study was based on both previous literature and clinical experience [14, 22]. Indicators such as primiparity, estimated fetal weight, anemia, and the number of vaginal examinations were consistent with those reported in earlier studies [14, 23]. Additionally, this study identified new predictive factors, including the degree of cervical dilation and height. This may be related to differences in sample characteristics, variable selection, and sample size.

Multiple studies have confirmed that primiparity is a risk factor for intrapartum fever [14, 24, 25]. Combined spinal–epidural anesthesia-related intrapartum fever is associated with noninfectious inflammation [26, 27]. Both labor and childbirth are accompanied by increased inflammation levels [28]. Primiparas are more sensitive to proinflammatory cytokines and may have greater levels of inflammation during labor than multiparas due to a longer duration of labor [29, 30]. The duration of labor has been proven to be an independent risk factor [31, 32]. The greater the degree of cervical dilatation before the start of labor analgesia is, the shorter the duration of labor analgesia that parturients may experience. Research has demonstrated that the duration of labor analgesia is closely related to the occurrence of fever [33]. However, previous studies have rarely included cervical dilatation as a specific indicator when exploring factors influencing labor analgesia and fever occurrence. This study highlights the importance of cervical dilatation as a potential predictive factor, filling a gap in prior research and offering new insights into the relationship between labor analgesia and fever.

Anemia significantly affects the immune system functionality by diminishing the body's oxygen transport

(See figure on next page.)

Fig. 4 **A** Calibration curve: The calibration curve shows the relationship between the predicted probabilities of intrapartum fever and the actual observed outcomes for different models. The diagonal line represents the ideal calibration scenario, indicating that the predicted probabilities closely align with the observed probabilities. **B** DCA Curve: DCA illustrating the standardized net benefit across varying high-risk thresholds for different models. The MLP model shows a consistent net benefit across the range of thresholds compared to the other models. **C** SHAP Plot: SHAP variable importance plot ranking the features based on their impact on the model's output. Features are ranked as follows: Primiparity > Estimated Neonatal Weight > Size of Uterine Dilatation > Anemia > Number of Vaginal Examinations > Height. The color gradient indicates the feature value, with deeper purple representing higher values and deeper yellow indicating lower values. Abbreviations: LR Logistic Regression, SVM Support Vector Machine, MLP Multilayer Perceptron, RF Random Forest, XGB Extreme Gradient Boosting, DT Decision Tree, AUC Area Under the Receiver Operating Characteristic Curve, SHAP SHapley Additive exPlanations

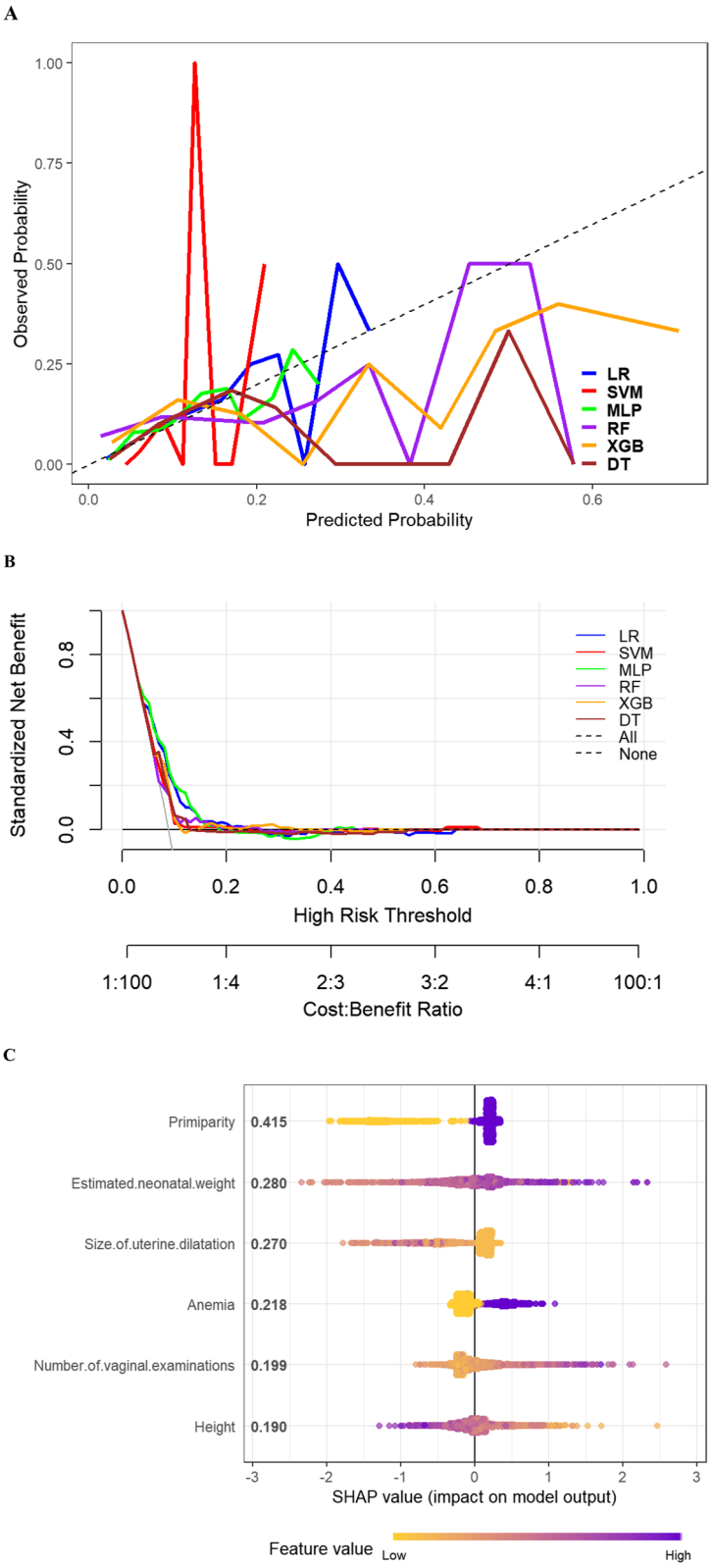

Fig. 4 (See legend on previous page.)

capacity. Insufficient oxygen delivery impairs the activity and efficacy of immune cells, resulting in a weakened immune response that increases the susceptibility of individuals with anemia to infections [34]. Additionally, anemia is often closely linked to iron deficiency, and iron is essential for optimal immune system operation. Iron deficiency further compromises immune defenses, reducing the body's ability to combat pathogens effectively [35].

Frequent vaginal examinations prior to labor analgesia may increase the risk of intrapartum fever. Each examination can potentially disrupt the protective barriers of the vagina and cervix, thereby increasing the likelihood of microorganisms entering the uterine cavity and leading to local infections and systemic inflammatory responses [36]. Furthermore, despite the use of sterile techniques, repeated examinations may still facilitate the spread of microorganisms.

Given that the hourly infusion rate of local anesthetic drugs remains constant and taller parturients have relatively longer spinal canals, the number of spinal segments blocked by the anesthetic is reduced. This decrease in blocked segments leads to a weakened blockade of the sympathetic nerves, thereby preserving the parturients' autonomous body temperature regulation and reducing the likelihood of noninfectious fever. Interestingly, previous studies have not extensively discussed height as a significant predictor of intrapartum fever, which may be related to our larger sample size and the fact that many earlier studies did not include height after accounting for BMI. Our findings highlight a potential physiological mechanism linking height to reduced fever risk, offering a novel perspective on this variable's role in clinical outcomes [13]. The greater the weight of the fetus is, the more likely dystocia due to cephalic presentation is to occur, which ultimately leads to prolonged labor [37, 38]. The duration of labor has been confirmed as an independent risk factor for fever by several studies [31, 32].

SHAP analysis clearly illustrates the impact of key features on prediction outcomes, enhancing the transparency of the model. For instance, it highlights the importance of factors such as primiparity and neonatal birth weight, enabling healthcare teams to more accurately identify high-risk parturients and implement personalized interventions. Additionally, the visualizations provided by SHAP facilitate effective communication between clinicians and patients, helping parturients better understand their individual risks and actively participate in their care. Therefore, SHAP analysis significantly improves the clinical applicability of the model and contributes to optimizing risk management.

The MLP model shows moderate sensitivity in identifying high-risk individuals but relatively low specificity and PPV, leading to potential false positives by misclassifying low-risk parturients as high-risk. However, its high NPV ensures reliable identification of low-risk cases, which is crucial for resource allocation and minimizing unnecessary interventions. In clinical practice, overidentification (false positives) is often more acceptable than underidentification (false negatives), especially in conditions like intrapartum fever that may cause serious outcomes for mothers and newborns. Early risk detection enables timely monitoring and intervention, helping prevent complications. While false positives may lead to additional noninvasive measures like temperature monitoring or increased clinical evaluations, these are manageable. The model's insights support clinicians in communicating with patients about the need for monitoring and encourage active participation in care, improving overall risk management and decision-making.

To integrate the MLP model into clinical workflows, it can be embedded into the hospital's electronic medical record system. Clinical staff can input patient data during admission, and the system will immediately generate fever risk assessment results to present to doctors. This assessment mechanism not only enhances decision-making efficiency but also assists healthcare providers in formulating personalized management plans, optimizing the care process for parturients. It is recommended to start with a pilot implementation in the obstetrics department, combined with training and feedback mechanisms to optimize the model's performance. While this model can help improve maternal and infant health, attention should also be paid to issues such as data integration, privacy protection, and staff training.

This study has several limitations. Firstly, the data were collected from a specific region, which may limit the applicability of the results to other populations or healthcare environments, affecting the model's performance across different demographic characteristics and healthcare systems. To enhance the generalizability of the research, future studies should consider adopting a multi-center design to collect broader regional data. Secondly, as a retrospective study, we face the risks of selection bias and information bias. Although we have made efforts to collect relevant variables, the indicators used may not capture all potential influencing factors, thus limiting the accuracy of the model. Therefore, conducting prospective validation studies will help further evaluate the model's effectiveness.Thirdly, while we considered multiple clinical indicators in model development, we did not include real-time data and other relevant indicators, such as cytokines. These biological markers may be crucial for identifying high-risk patients, and future research should explore the integration of these indicators to enhance the model's

predictive capability and clinical applicability. Finally, although this study provides a new perspective for the early identification of labor-related fever risk, it remains essential to assess the model's applicability and effectiveness in actual clinical settings. Enhancing patient education and involvement in decision-making, as well as providing personalized care based on risk assessment results, will contribute to improving maternal and neonatal health outcomes.

Conclusions

This study developed and validated a MLP model based on six pre-labor analgesia clinical indicators to predict intrapartum fever. The model demonstrates potential for early identification of high-risk parturients, allowing timely interventions to improve maternal and neonatal health outcomes. While further validation and optimization are needed, the model provides a foundation for integrating machine learning into obstetric care to enhance risk prediction and clinical decision-making.

Abbreviations

AUC	Area under the receiver operating characteristic curve
BMI	Body mass index
CI	Confidence interval
DCA	Decision curve analysis
DT	Decision Tree
GDM	Gestational diabetes mellitus
LASSO	Least Absolute Shrinkage and Selection Operator
LYM	Lymphocyte
MLP	Multilayer Perceptron
MSE	Mean squared error
NEUT	Neutrophil
NICU	Neonatal intensive care unit
NLR	Neutrophil–Lymphocyte Ratio
NPV	Negative Predictive Value
PLR	Platelet–Lymphocyte Ratio
PPV	Positive Predictive Value
PROM	Premature rupture of membranes
RF	Random Forest
ROC	Receiver Operating Characteristic
SVM	Support Vector Machine
SHAP	SHapley Additive exPlanations
XGB	Extreme Gradient Boosting
WBC	White blood cell

Acknowledgements

Not applicable.

Authors' contributions

BL and LL conceived the study design; BL, LL, and FJ drafted the manuscript; DYW, JZ, and CPL were responsible for the methodology; BL, LL, and CPL conducted the validation; LL, DYW, and GZ were involved in the formal analysis; YLL was responsible for the investigation; HRL, HQX, and CPL were involved in data curation. The manuscript was reviewed and approved by BL, and supervision was provided by JZ. All the authors have read and approved the final manuscript.

Funding

This study was supported by the Sichuan Science and Technology Support Plan (No. 2023YFQ0005) and the Chengdu Medical Research Project (No. 2023012).

Data availability

The datasets utilized and analyzed in this study can be obtained from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

This study received approval from the Ethics Committee of Chengdu Jinjiang District Women & Children Health Hospital (approval number: 2024 trial (20)) and Sichuan Jinxin Xinan Women & Children's Hospital (approval number: 2024 trial (2)). Owing to the retrospective design of the study, the requirement for written informed consent was waived by both ethics committees. We obtained administrative permission to access and utilize medical records from the respective hospitals, with authorization granted by the data protection officers. All procedures were conducted following relevant ethical guidelines and regulations, in line with the principles outlined in the World Medical Association's Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Anesthesiology, Chengdu Jinjiang District Women and Children Health Hospital, No. 3 Sanguantang Street, Jinjiang District, Chengdu 610011, Sichuan, China. [2]Department of Anesthesiology, Sichuan Women's and Children's Hospital/Women's and Children's Hospital,Chengdu Medical College, No. 290 West Second Street, Shayan Road, Wuhou District, Chengdu 610031, Sichuan, China. [3]Department of Epidemiology and Biostatistics, Chengdu Medical College, Jinniu District, Chengdu 610500, Sichuan, China. [4]Department of Anesthesiology, Sichuan Jinxin Xinan Women & Children's Hospital, Jinjiang District, Chengdu 610011, Sichuan, China.

Received: 21 November 2024 Accepted: 20 January 2025
Published online: 06 March 2025

References

1. Halliday Lucy, Nelson Scott M, Kearns Rachel J. Epidural analgesia in labor: A narrative review. Int J Gynaecol Obstet. 2022;159(2):356–364.
2. Sultan Pervez, Patel Selina D, Jadin Sandra, Carvalho Brendan, Halpern Stephen H. Transversus abdominis plane block compared with wound infiltration for postoperative analgesia following Cesarean delivery: a systematic review and network meta-analysis. Can J Anaesth. 2020;67(12):1710–1727.
3. Ashagrie Henos Enyew, Fentie Demeke Yilkal, Kassahun Habtamu Getinet. A review article on epidural analgesia for labor pain management: a systematic review. Int J Surg Open. 2020;24:100–04.
4. Gaetano R, Antonio S, Stefano C, et al. Transversus abdominis plane block versus wound infiltration for post-cesarean section analgesia: A systematic review and meta-analysis of randomized controlled trials. Int J Gynaecol Obstet. 2021;153(3):383–92.
5. Sultan Pervez, David Anna L, Fernando Roshan, Ackland Gareth L. Inflammation and Epidural-Related Maternal Fever: Proposed Mechanisms. Anesth Analg. 2016. 122(5):1546–1553.
6. Hensel Drew, Zhang Fan, Carter Ebony B, et al. Severity of intrapartum fever and neonatal outcomes. Am J Obstet Gynecol. 2022. 227(3):513.e1–513.e8.
7. Sophie J, Enrico L, Christiana N, et al. Epidural-Related Fever and Maternal and Neonatal Morbidity: A Systematic Review and Meta-Analysis. Neonatology. 2020;117(3):259–70.
8. Eran A, Lina S, Yossi T, et al. Intrapartum fever and the risk for perinatal complications - the effect of fever duration and positive cultures. J Matern Fetal Neonatal Med. 2018;31(11):1418–25.
9. Törnell S, Ekéus C, Hultin M, Håkansson S, Thunberg J, Högberg U. Low Apgar score, neonatal encephalopathy and epidural analgesia during

labour: a Swedish registry-based study. Acta Anaesthesiol Scand. 2015;59(4):486–95.
10. Dior Uri P, Kogan Liron, Eventov-Friedman Smadar, et al. Very High Intrapartum Fever in Term Pregnancies and Adverse Obstetric and Neonatal Outcomes. Neonatology. 2016;109(1):62–68.
11. Laura G. Maternal fever in labor: etiologies, consequences, and clinical management. Am J Obstet Gynecol. 2023;228(5S):S1274–82.
12. Christopher G, Roger N, Latha H, Julie R, Cynthia S, Laura G. Maternal Epidural Steroids to Prevent Neonatal Exposure to Hyperthermia and Inflammation. Am J Perinatol. 2019;36(8):828–34.
13. Yuru F, Wenwen H, Shi F, et al. Programmed intermittent epidural bolus decreases the incidence of intra-partum fever for labor analgesia in primiparous women: a randomized controlled study. Arch Gynecol Obstet. 2019;300(6):1551–7.
14. Jiang Zhenfei Hu, Xiaoyi ZH, et al. Nomogram for perinatal prediction of intrapartum fever: a retrospective case-control study. BMC Pregnancy Childbirth. 2021;21(1):445.
15. Caixia S, Shujun R, Chen C, et al. Pulse perfusion index for predicting intrapartum fever during epidural analgesia. J Clin Anesth. 2022;80:110852.
16. Zhao W, Wang LZ, Chang XY, Zhang YF, Xiao F, Xia F. Maternal serum C-reactive protein and white blood cell count at hospital admission as predictors of intrapartum maternal fever: a retrospective case-control study in women having epidural labor analgesia. Int J Obstet Anesth. 2022;50:103537.
17. 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.
18. Jin Y, Lan A, Dai Y, Jiang L, Liu S. Development and testing of a random forest-based machine learning model for predicting events among breast cancer patients with a poor response to neoadjuvant chemotherapy. Eur J Med Res. 2023. 28(1): 394.
19. Menzenbach Jan, Kirfel Andrea, Guttenthaler Vera, et al. PRe-Operative Prediction of postoperative DElirium by appropriate SCreening (PROPDESC) development and validation of a pragmatic POD risk screening score based on routine preoperative data. J Clin Anesth. 2022. 78: 110684.
20. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013. 310(20): 2191-2194.
21. Debnath S, Koppel R, Saadi N, et al. Prediction of intrapartum fever using continuously monitored vital signs and heart rate variability. Di git Health. 2023. 9: 20552076231187594.
22. Chai J, Jia L, Cao H, et al. Association of lymphocyte count and incidence of maternal fever in epidural analgesia-involved labor. Ann Transl Med. 2020. 8(23): 1584.
23. Hui W, Zaiping Y, Siyi W, 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.
24. Seijmonsbergen-Schermers Anna E, van den Akker Thomas, Rydahl Eva, et al. Variations in use of childbirth interventions in 13 high-income countries: A multinational cross-sectional study. PLoS Med. 2020;17(5):e1003103.
25. Bai-Song Z, Bing Li, Qing-Ning W, Jun-Xiang J, Xing-Rong S. Time- and dose-dependent correlations between patient-controlled epidural analgesia and intrapartum maternal fever. BMC Anesthesiol. 2021;21(1):31.
26. Sharma Shiv K, Rogers Beverly B, Alexander James M, McIntire Donald D, Leveno KJ. A randomized trial of the effects of antibiotic prophylaxis on epidural-related fever in labor. Anesth Analg. 2014;118(3):604–10.
27. Chan James K, Roth Johannes, Oppenheim Joost J, et al. Alarmins: awaiting a clinical response. J Clin Invest. 2012;122(8):2711–2719.
28. Xianjin Z, Jiang Li, Deng Shengqiong Xu, Zhendong LZ. Ropivacaine at different concentrations on intrapartum fever, IL-6 and TNF-α in parturient with epidural labor analgesia. Exp Ther Med. 2019;17(3):1631–6.
29. Kuhlman Kate R, Robles Theodore F, Haydon Marcie D, Dooley Larissa, Boyle Chloe C, Bower Julienne E. Early life stress sensitizes individuals to the psychological correlates of mild fluctuations in inflammation. Dev Psychobiol. 2020;62(3):400–408.
30. Qinghua P, Zeyong Y, Zhang Wangping Wu, Xiaomin. Comparison of median effective concentration of ropivacaine in multiparas or primiparas during epidural labor analgesia: STROBE compliant. Medicine (Baltimore). 2020;99(1):e18673.
31. Harper Lorie M, Caughey Aaron B, Roehl Kimberly A, Odibo Anthony O, Cahill AG. Defining an abnormal first stage of labor based on maternal and neonatal outcomes. Am J Obstet Gynecol. 2014;210(6):536.e1-7.
32. Jie R, Tao W, Bo Y, et al. Risk Factors and Safety Analyses for Intrapartum Fever in Pregnant Women Receiving Epidural Analgesia During Labor. Med Sci Monit. 2021;27:e929283.
33. Yin Huifen Hu, Rong. A cohort study of the impact of epidural analgesia on maternal and neonatal outcomes. J Obstet Gynaecol Res. 2019;45(8):1435–41.
34. Manfred N, Günter W. Iron in infection and immunity. Mol Aspects Med. 2020;75:100864.
35. Jingyao Hu, Xinyu Li, Liang Y, Hangyu Li. Hypoxia, a key factor in the immune microenvironment. Biomed Pharmacother. 2022;151:113068.
36. Gluck Ohad, Mizrachi Yossi, Ganer Herman Hadas, Bar Jacob, Kovo Michal, Weiner Eran. The correlation between the number of vaginal examinations during active labor and febrile morbidity, a retrospective cohort study. BMC Pregnancy Childbirth. 2020;20(1):246.
37. Finnegan Catherine L, Burke Naomi, Breathnach Fionnuala, et al. Defining the upper limit of the second stage of labor in nulliparous patients. Am J Obstet Gynecol MFM. 2019;1(3):100029.
38. Maria B, Lars T, Marie O. Prolonged passive second stage of labor in nulliparous women-Prevalence and risk factors: A historical cohort study. Acta Obstet Gynecol Scand. 2022;101(5):499–505.

Publisher's Note