



OPEN ACCESS

EDITED BY

Karel Allegaert,
KU Leuven, Belgium

REVIEWED BY

Fatemeh Darsareh,
Hormozgan University of Medical Sciences, Iran
Sulaiman Salim Said Al Mashrafi,
RMIT University, Australia

*CORRESPONDENCE

Haixia Zhang,
✉ zhx_510@hotmail.com
Hongliang Mei,
✉ liulidemao@126.com

RECEIVED 19 April 2025

ACCEPTED 19 May 2025

PUBLISHED 11 June 2025

CITATION

Guo X, Zhang H and Mei H (2025) Machine learning algorithms to predict epidural-related maternal fever: a retrospective study. *Front. Pharmacol.* 16:1614770. doi: 10.3389/fphar.2025.1614770

COPYRIGHT

© 2025 Guo, Zhang and Mei. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Machine learning algorithms to predict epidural-related maternal fever: a retrospective study

Xiaohui Guo^{1,2,3}, Haixia Zhang^{1,3,4*} and Hongliang Mei^{1,3,4*}

¹Department of Pharmacy, Nanjing Drum Tower Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, Jiangsu, China, ²Department of Pharmacy, Hainan Women and Children's Medical Center, Haikou, Hainan, China, ³China Hospital Reform and Development Research Institute of Nanjing University, Nanjing Drum Tower Hospital, Nanjing, China, ⁴Nanjing Medical Center for Clinical Pharmacy, Nanjing, Jiangsu, China

Introduction: The epidural-related maternal fever (ERMF) induced by patient-controlled epidural analgesia (PCEA) remains unpredictable. Our objective is to develop ERMF prediction models using real-world data, aiming to identify pertinent contributing factors and support obstetricians in making personalized clinical decisions.

Methods: Women who used patient-controlled epidural analgesia between October 2021 and March 2023 at a tertiary hospital in Jiangsu Province were retrospectively documented. The primary outcome was the occurrence of maternal fever associated with epidural use. We developed six machine learning (ML) models and assessed the area under curve (AUC) for characteristics of subjects' performance, calibration curves, and decision curve analyses.

Results: A total of 1,492 women were enrolled, with 24.3% experiencing ERMF (362 cases). The AUC ratios between the logistic regression (LR) model and the stochastic gradient descent (SGD) models showed statistical significance ($p < 0.05$), while the differences between the other models were not statistically significant. In comparison to the SVM model, the LR model exhibited better calibration (Brier score: 0.193; calibration slope: 0.715; calibration intercept: 0.062). Consequently, the LR model was selected as the prediction model. Furthermore, the LR-based nomogram identified eight significant predictors of ERMF, including neutrophil percentage, first stage of labor, amniotic fluid contamination during membrane rupture, artificial rupture of membranes, chorioamnionitis, post-analgesic antimicrobials, pre-analgesic oxytocin, post-analgesic oxytocin, and dinoprostone suppositories.

Conclusion: Optimally applying logistic regression models can enable rapid and straightforward identification of ERMF risk and the implementation of rational therapeutic measures, in contrast to machine learning models.

KEYWORDS

epidural-related maternal fever, machine learning, predictive model, Nomograms, risk assessment

1 Introduction

Epidural analgesia (EA) has gained widespread acceptance, chosen by a significant proportion of women in labor, ranging from 10% to 83% over the past 2 decades (Seijmonsbergen-Schermers et al., 2020). Patient-controlled epidural analgesia (PCEA) is commonly used for managing labor pain. However, recent years have raised increasing concerns about the safety of PCEA, with particular attention to epidural-related maternal fever (ERMF) (Sultan and Segal, 2020). ERMF, marked by a temperature of $\geq 38^{\circ}\text{C}$, occurs in approximately 11%–33% of women who use PCEA (Sultan et al., 2016; Del Arroyo et al., 2019). This condition can disrupt obstetric management during labor, potentially necessitating additional antibiotics or a cesarean section. Moreover, it may lead to adverse neonatal outcomes, including neonatal infections, brain damage, and long-term learning disabilities (Impey et al., 2008; Ren et al., 2021). Therefore, early prediction of ERMF is of utmost importance.

Nevertheless, the etiology of ERMF remains elusive, and previous studies have identified predictors of ERMF in patients receiving PCEA. Women who developed intrapartum fever had a significantly longer first stage of labor and a higher incidence of cesarean section, assisted vaginal delivery, intrapartum hemorrhage, and amniotic fluid clouding (Impey et al., 2008). Risk factors for epidural fever include prolonged labor, a low parity, ≥ 7 cervical examinations, higher EA dosages, extended EA duration, and total labor duration (Wang et al., 2023). Women with placental neutrophilic inflammation displayed significantly higher fever rates (Sharma et al., 2014). However, all predictors in these studies depended on traditional statistical methods, overlooking possible nonlinear relationships or multicollinearity among variables. Hence, it is vital to underscore that, even when accounting for all predictors, the method lacks a firmly established accuracy in perfectly predicting the onset of ERMF.

In the era of continually accumulating patient clinical data, contemporary healthcare decision-making demands precise, innovative, and prediction-driven decision support. Machine learning (ML) algorithms, as a statistical methodology, offer advantages over traditional statistical modeling methods due to their capacity for generalization and differentiation. Machine learning models excel at extracting implicit information from extensive datasets, potentially addressing challenges associated with using traditional logistic regression models to resolve controversies related to influential factors (Goecks et al., 2020).

Artificial intelligence (AI) is widely employed in healthcare research, and real-world studies involving machine learning algorithms have shown promising performance. AI has found application in areas like prognostic assessment for disease treatment and histologic imaging for disease diagnosis (Roozbeh et al., 2025; Banaei et al., 2025; Darsareh et al., 2023). However, models using ML algorithms to predict the influencing factors of ERMFs have not been developed yet. Consequently, this study aims to construct a predictive model for ERMF based on real-world data with the goal of identifying relevant influencing factors and assisting obstetricians in making personalized clinical decisions.

2 Methods

Study Design: The study was approved by the Medical Ethics Committee of the Affiliated Drum Tower Hospital of Nanjing University Medical School (approval number 2022–106-04). Written informed consent was obtained from patients or their representatives for participation in this study.

Patients: We conducted a retrospective analysis of 1,492 women who received PCEA between October 2021 and March 2023 at a tertiary care hospital in Jiangsu Province. The inclusion criteria were as follows (Seijmonsbergen-Schermers et al., 2020): Women aged over 18 and up to 50 years (Sultan and Segal, 2020); Singleton pregnancies with vaginal delivery; and (Sultan et al., 2016) Meeting the clinical criteria for administering analgesia compatible with PCEA (Laboring women requesting analgesia without the following exclusion criteria may receive PCEA: Patient refusal, Mental illness, Severe psychological distress, Spinal trauma or deformity, Puncture site infection or sepsis, Hemodynamic instability, Coagulation disorders.). The exclusion criteria were as follows (Seijmonsbergen-Schermers et al., 2020): Body temperature $\geq 37.5^{\circ}\text{C}$ before PCEA (Sultan and Segal, 2020); PECA after the cervix has opened to 10 cm (Sultan et al., 2016); More than two features were missing in a single sample (Del Arroyo et al., 2019); More than 20% missing values for a Individual features.

Data collection and definitions: The primary outcome of this study was defined as the occurrence of fever ($\geq 38^{\circ}\text{C}$) in perinatal women following PCEA (3). Fever data were retrieved from nursing records, considering the highest temperature reading in cases of multiple recordings. Patient demographic information, clinical data, and laboratory test results were obtained from the electronic healthcare information system. All data underwent manual processing and verification, with missing values imputed using the K-nearest neighbor imputation (KNNI) method, which calculates the average of the K nearest eigenvalues to fill in missing data points based on eigenvalues in proximity.

Before applying machine learning algorithms, features were normalized through mean-centering and scaling to unit variance. Variable selection followed a rigorous methodology, involving a systematic review of relevant studies and expert consultations to ensure comprehensive coverage. The study involved 45 variables, categorized as five clinical variables (age, BMI, parturition, gestational week, and cervical dilatation at the onset of analgesia), 11 disease variables, 10 surgical variables, two biochemical criteria, and 27 medication-related variables. Relevant demographic information encompassed age, height, weight, BMI, gestational week, primiparity, transplacental delivery, fetal weight, and comorbid conditions such as hypertension, diabetes mellitus, Hypothyroidism, chorioamnionitis, Group B *Streptococcus* colonization, acute chorioamnionitis, and hepatitis B in pregnancy. Records of medication administered both before and after PCEA initiation included oxytocin, antibiotics (cephalosporin antibiotics), pethidine, diazepam, nifedipine, dinoprostone suppositories, misoprostol tablets, magnesium sulfate, and others. Laboratory findings obtained immediately before PCEA initiation included white blood cell count and neutrophil percentage. Additional clinical data encompassed factors such as cervical dilatation, number of vaginal examinations, artificial rupture of membranes,

balloons for uterine cervical ripening, first stage of labor, second stage of labor, third stage of labor, total duration of labor, amniotic fluid contamination, Time between rupture of fetal membranes and labor onset, Intrapartum hemorrhage, etc.

Model development: The dataset was divided into training and test sets using an 8:2 ratio for model development. We applied the least absolute shrinkage and selection operator (LASSO) algorithm to reduce data dimensionality. Furthermore, all features selected by LASSO were incorporated into six models to evaluate ERMF. Model construction for machine learning algorithms was carried out using Python's scikit-learn library, an open-source Python package specifically designed for predictive model building in machine learning. The scikit-learn package facilitates the construction of supervised machine learning models, encompassing techniques such as logistic regression (LR), support vector machines (SVM), stochastic gradient descent (SGD) (Nguyen et al., 2020), random forest classifiers (RFC) (Lai et al., 2023), random forest (RF) (Rigatti, 2017), extreme gradient boosting (XGboost) (Nwanosike et al., 2022) and multi-layer perceptron (MLP). The data from the split training set was utilized to build the machine learning algorithm model with the application of 5-fold cross-validation. Hyperparameters employed in model construction were established through grid search with 5-fold cross-validation. The algorithms were implemented in Python version 3.7, making use of the scikit-learn library version 0.24.1 and the XGBoost library version 1.2.1.

Model Evaluation and Feature Importance: Following model acquisition, we assessed predictive performance using the area under curve (AUC) score derived from the receiver operating characteristic curve (ROC) (Mandrekar, 2010). This score depends on sensitivity and specificity measurements at different thresholds, typically ranging from 0.5 to 1. A value approaching one indicates superior model performance. The AUCs and their corresponding 95% confidence intervals (CIs) were computed for each of the six models. The ROC-AUCs were statistically compared for significance using the DeLong test in R software, with a p-value below 0.05 indicating a significant difference in model performance. Additional evaluation metrics for model performance included accuracy, precision, and recall.

The calibration of the ML model on the test set was assessed by calculating the Brier score, calibration slope, and calibration intercept. The Brier score quantified the difference between the estimated and observed risk of ERMF, and models with a calibration slope of one and calibration intercept of 0 indicated perfect calibration (Oosterhoff et al., 2022). Decision curve analysis was employed to evaluate clinical utility, which considers a weighted average of true positives and false positives, by calculating net gains across a range of threshold probabilities.

Statistical analyses: The nomogram was constructed using a fitted LR model, employing the rms package. All analyses were performed using SPSS version 26.0. Descriptive statistics were applied to all variables. Subsequently, univariate analyses were carried out using the t-test for continuous variables with a normal distribution and Mann-Whitney U-tests for those with a non-normal distribution. Categorical variables were evaluated using

either Fisher's exact test or the χ^2 test. All tests were two-sided, and p-values below 0.05 were deemed statistically significant.

3 Results

3.1 Study population

Figure 1 illustrates that 1,492 patients met the inclusion criteria and were enrolled in this study. Patients with ERMF accounted for 24.3% of the sample (362 patients); a comparable proportion of ERMF patients was maintained between the training set and the test set (24.6% vs 22.7%, $p > 0.05$). Baseline statistics for both the ERMF and non-ERMF groups are presented in Table 1. Patient characteristics were balanced between the training set ($n = 1,193$, 80%) and the test set ($n = 299$, 20%) (Supplementary Table S1).

3.2 Feature selection

Table 1 demonstrates that 21 features in the univariate analyses exhibited significant differences between patients with and without ERMF ($p < 0.05$). Six features with non-zero coefficients were subsequently removed from the LASSO regression. The final fifteen variables incorporated into the ML model were: neutrophil percentage, first stage of labor, amniotic fluid contamination during membrane rupture, artificial rupture of membranes, chorioamnionitis, post analgesic antimicrobials, pre analgesic oxytocin, post analgesic oxytocin, and dinoprostone suppositories (Table 2).

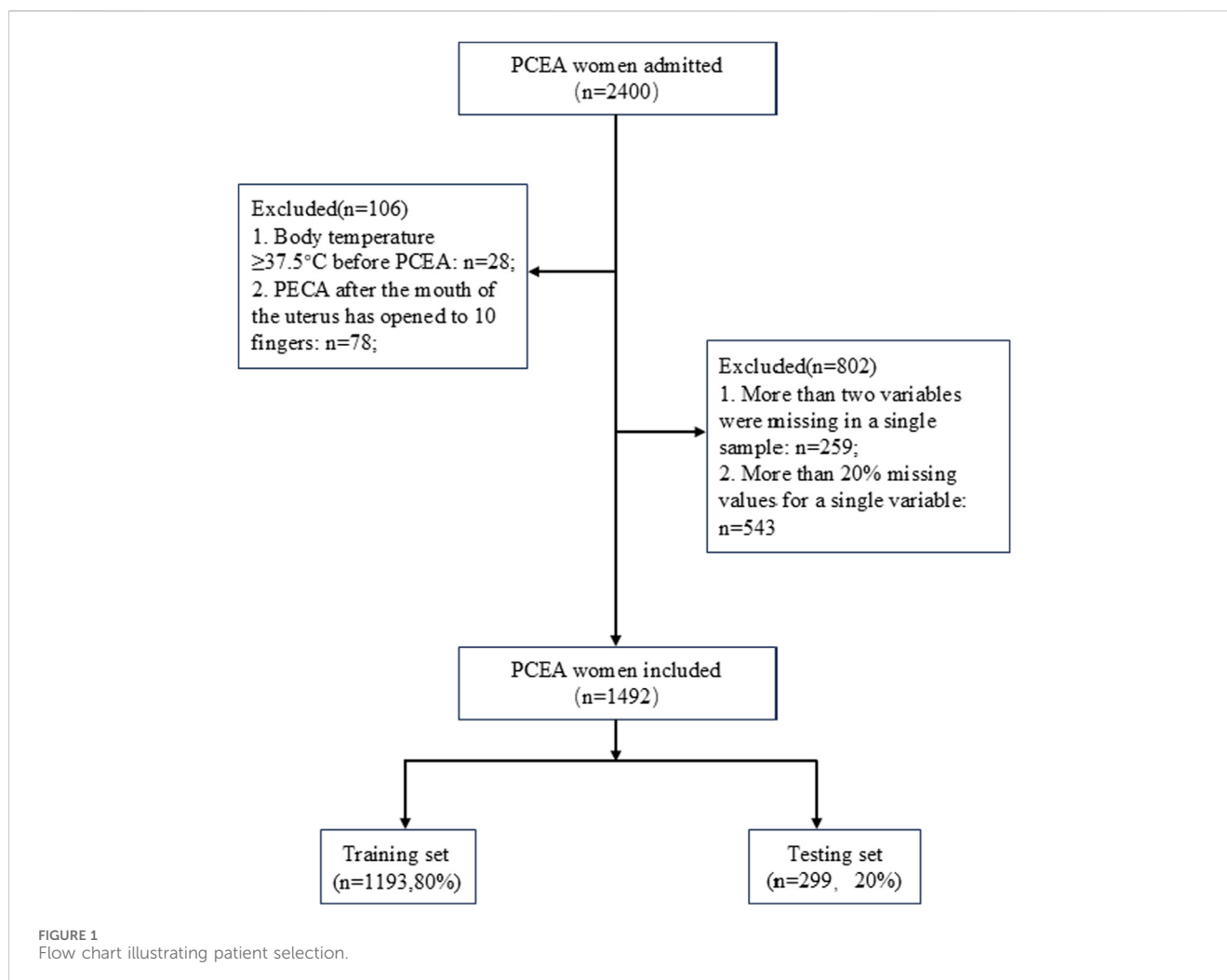
3.3 Model performance

The model hyperparameters are presented in Supplementary Table S2. The ROC of each model on the training set is depicted in Figure 2. Table 3 displays the performance metrics for the test set, encompassing AUC, sensitivity, Brier score, calibrated slope, and calibrated intercept.

As depicted in Table 3 and Figure 2, the discriminative performance was observed in LR (AUC, 0.791), SVM (AUC, 0.786), SGD (AUC, 0.766), RFC (AUC, 0.752), XGB (AUC, 0.749), and MLP (AUC, 0.749) on the testing set. DeLong's test revealed a statistically significant difference in AUC between the LR model and the SGD model, ($P < 0.05$). No statistically significant differences were observed among the remaining models (Supplementary Table S3).

In the test set, Brier scores ranged from 0.155 to 0.193, calibrated slopes ranged from 0.306 to 0.867, and calibrated intercepts ranged from -0.062 to 0.109 (Figure 3A; Table 3). Decision curve analyses demonstrated that the SVM and RFC models exhibited a higher net benefit and a default strategy of treating all patients or no patients compared to other ML models (Figure 3B).

Considering all factors, the LR model's AUC was statistically different from that of the RFC, and the LR also demonstrated superior calibration compared to the SVM (Brier score, 0.193; calibration slope, 0.715; calibration intercept, -0.062). Consequently, the LR model was selected as the predictive model.



3.4 Development a nomogram for the prediction of ERMF with the LR model

To predict ERMF, we constructed an LR-based nomogram. The nomogram identifies eight key predictors for ERMF prediction: neutrophil percentage, first stage of labor, amniotic fluid contamination during membrane rupture, artificial rupture of membranes, chorioamnionitis, post-analgesic antimicrobials, pre-analgesic oxytocin, post-analgesic oxytocin, and dinoprostone suppositories (refer to Figure 4).

4 Discussion

The etiology of maternal fever induced by epidural labor analgesia remains uncertain, and the application of specific preventive or therapeutic measures for epidural-associated maternal fever presents challenges. Several studies have investigated factors influencing the incidence of ERMF to varying degrees, with the aim of identifying effective preventive strategies (Ren et al., 2021; Wang et al., 2023; Yin and Hu, 2022). Machine learning has seen extensive application in obstetrics and gynecology

in recent years. However, its utilization in the context of epidural labor analgesia and its association with ERMF remains unexplored.

In this retrospective study, we compared the performance of traditional binomial logistic regression with machine learning algorithmic prediction models for forecasting ERMF in pregnant women who underwent PCEA, utilizing real-world research data. The results indicated that the LR model achieved the highest AUC among all tested models. Additionally, the application of a nomogram to depict pertinent factors influencing ERMF, including neutrophil percentage, first stage of labor, Amniotic fluid contamination during membranes rupture, artificial rupture of membranes, chorioamnionitis, post-analgesic antimicrobials, pre analgesic oxytocin, post analgesic oxytocin, and dinoprostone suppositories, underscores the importance of further analyzing these results. Furthermore, the AUC across all examined models ranged from 0.749 to 0.791, indicating that several models in this study exhibit strong generalization capabilities within this cohort. In this study, we found no evidence to support that machine learning-based ERMF prediction models have better performance than logistic regression-based models.

The LR model exhibited superior sensitivity but lower accuracy compared to the machine learning model, achieving 44% precision.

TABLE 1 Demographic and clinical data of the patients.

Feature	ERMF		P
	Non-ERMF (n = 1,130)	ERMF (n = 362)	
Height (cm), Mean $\pm$ SD	162.44 $\pm$ 5.07	161.81 $\pm$ 4.97	0.037
Weight (kg), Mean $\pm$ SD	69.51 $\pm$ 10.34	69.33 $\pm$ 10.08	0.769
BMI (kg/m ²), Mean $\pm$ SD	26.32 $\pm$ 3.59	26.44 $\pm$ 3.30	0.588
Age (year), Mean $\pm$ SD	30.54 $\pm$ 3.34	30.63 $\pm$ 3.12	0.650
Gestational week (week), M (Q ₁ , Q ₃)	39.30 (38.60–40.10)	39.50 (39.10–40.20)	0.001
Intrapartum hemorrhage (mL), M (Q ₁ , Q ₃)	200.00 (150.00–400.00)	300.00 (150.00–400.00)	0.004
White blood cell count (10 ⁹ /L), M (Q ₁ , Q ₃)	9.30 (8.10–10.50)	9.41 (8.30–11.20)	0.029
Cervical dilatation (cm), M (Q ₁ , Q ₃)	2.00 (2.00–3.00)	2.00 (2.00–2.50)	0.775
Neutrophil percentage (%), M (Q ₁ , Q ₃)	76.48 (73.90–78.60)	76.62 (74.20–80.57)	0.034
Number of vaginal examinations (n), M (Q ₁ , Q ₃)	4.00 (3.00–5.00)	4.00 (3.00–6.00)	0.003
First stage of labor (hour), M (Q ₁ , Q ₃)	8.00 (5.87–10.50)	10.00 (8.00–12.96)	<0.001
Second stage of labor (hour), M (Q ₁ , Q ₃)	0.33 (0.02–0.75)	0.40 (0.02–0.73)	0.834
Third stage of labor (hour), M (Q ₁ , Q ₃)	0.08 (0.07–0.13)	0.08 (0.08–0.12)	0.237
Total duration of labor (hour), M (Q ₁ , Q ₃)	8.42 (6.24–11.08)	10.86 (8.45–13.46)	<0.001
Time between rupture of fetal membranes and labor onset (hour), M (Q ₁ , Q ₃)	9.11 (5.37–16.35)	10.79 (7.70–16.39)	<0.001
Primiparous woman, n (%)	1,050 (92.92)	352 (97.24)	0.003
Amniotic fluid contamination during membranes rupture, n (%)	64 (5.66)	32 (8.84)	0.032
Amniotic fluid contamination at delivery of the fetus, n (%)	180 (15.93)	72 (19.89)	0.080
Balloons for uterine cervical ripening, n (%)	35 (3.10)	15 (4.14)	0.336
Artificial rupture of membranes, n (%)	668 (59.12)	247 (68.23)	0.002
Gestational diabetes, n (%)	164 (14.51)	46 (12.71)	0.390
Gestational hypertension, n (%)	53 (4.69)	11 (3.04)	0.177
Hypothyroidism in pregnancy, n (%)	126 (11.15)	45 (12.43)	0.506
Giant fetus, n (%)	31 (2.74)	16 (4.42)	0.112
Chorioamnionitis, n (%)	5 (0.44)	8 (2.21)	0.005
Group B <i>Streptococcus</i> colonization, n (%)	22 (1.95)	7 (1.93)	0.987
Acute chorioamnionitis, n (%)	5 (0.44)	8 (2.21)	0.005
Hepatitis B in pregnancy, n (%)	7 (0.62)	4 (1.10)	0.557
Immune system diseases, n (%)	24 (2.12)	13 (3.59)	0.118
Preterm premature rupture of membranes, n (%)	285 (25.22)	67 (18.51)	0.009
Pre analgesic prophylactic antimicrobials, n (%)	107 (9.47)	51 (14.09)	0.013
Post analgesic prophylactic antimicrobials, n (%)	210 (18.58)	175 (48.34)	<0.001
Pre analgesic antimicrobials, n (%)	360 (31.86)	115 (31.77)	0.974
Post analgesic antimicrobials, n (%)	338 (29.91)	243 (67.13)	<0.001
Pre analgesic oxytocin, n (%)	864 (76.46)	253 (69.89)	0.012
Post analgesic oxytocin, n (%)	366 (32.39)	165 (45.58)	<0.001
Pre analgesic magnesium sulfate injection, n (%)	52 (4.60)	12 (3.31)	0.293

(Continued on following page)

TABLE 1 (Continued) Demographic and clinical data of the patients.

Feature	ERMF		P
	Non-ERMF (n = 1,130)	ERMF (n = 362)	
Post analgesic magnesium sulfate injection, n (%)	31 (2.74)	13 (3.59)	0.407
Pre analgesic diazepam, n (%)	94 (8.32)	33 (9.12)	0.636
Post analgesic diazepam, n (%)	7 (0.62)	2 (0.55)	1.000
Pre analgesic pethidine, n (%)	199 (17.61)	53 (14.64)	0.189
Post analgesic pethidine, n (%)	21 (1.86)	9 (2.49)	0.459
Misoprostol Tablets, n (%)	178 (15.75)	59 (16.30)	0.805
Dinoprostone Suppositories, n (%)	268 (23.72)	101 (27.90)	0.108
Pre analgesic nifedipine, n (%)	83 (7.35)	28 (7.73)	0.806
Post analgesic nifedipine, n (%)	45 (3.98)	12 (3.31)	0.564

(BMI, body mass index; ERMF, epidural-associated maternal fever; M, median; Q1, the first quartile; Q3, the third quartile.).

TABLE 2 LASSO estimates of selected variables.

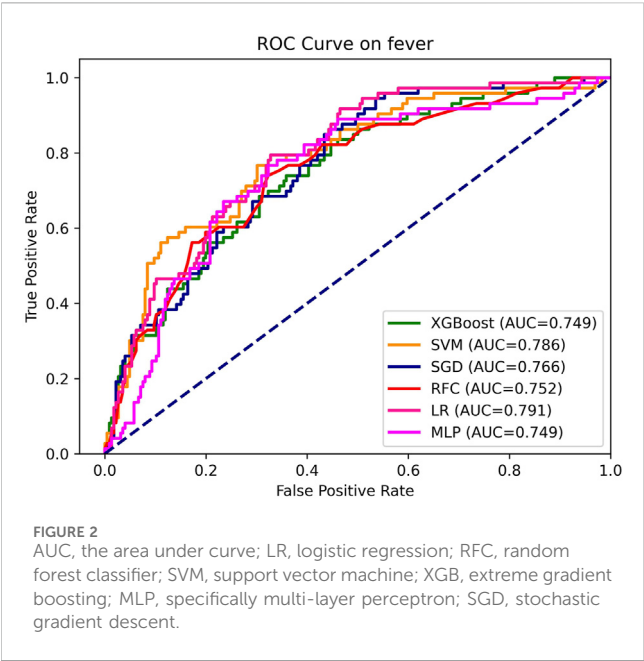
Variable	Estimated value
Gestational week	0.111526212
Intrapartum hemorrhage	0.000872844376
Neutrophil percentage	0.104583290
First stage of labor	0.279716503
Primiparous woman	0.0348427924
Amniotic fluid contamination during membranes rupture	0.0480524449
Artificial rupture of membranes	0.0326044562
Chorioamnionitis	0.173171166
Acute chorioamnionitis	0.00000000000000637527975
Pre-analgesic prophylactic antimicrobials	0.00446981574
Post-analgesic prophylactic antimicrobials	0.0715522209
Post-analgesic antimicrobials	0.205517042
Pre-analgesic oxytocin	-0.0575105519
Post-analgesic oxytocin	0.0540198947
Dinoprostone Suppositories	0.0161213136

(LASSO, The least absolute shrinkage and selection operator.).

The LR model sacrificed some accuracy to achieve improved AUC and sensitivity. After careful consideration, we opted for the predictive model based on LR analysis for three key reasons. ERMF has a prevalence of approximately 20% (Fusi et al., 1989), and the dataset suffers from severe imbalance. It is important to note that a predictive model can achieve high accuracy $[(TP + TN)/(TP + TN + FP + FN)]$ if it simply predicts that all patients do not have ERMF. Hence, when comparing different models, accuracy is not the primary concern. Second, for ERMF patients, sensitivity holds greater importance than accuracy. The development of ERMF in a woman can lead to severe adverse pregnancy outcomes. Third, screening these variables is more straightforward when visualized

in a nomogram. All variables are interpretable and quantifiable, eliminating the “black box” nature of machine learning. This outcome is unsurprising, as previous evidence has demonstrated that machine learning does not outperform logistic regression in clinical prediction models (Christodoulou et al., 2019).

Our study revealed an elevated neutrophils percentage and an extended first stage of labor, both of which are associated with ERMF, which is consistent with previous studies (Zhao et al., 2022; Curtin et al., 2015). This may be due to the fact that extended labor may deplete physical resources, potentially compromising immunity and elevating the risk of infection (Goetzl et al., 2001). Also, previous studies have confirmed a higher risk of fever in women with



placental neutrophilic inflammation (Sultan and Segal, 2020). Furthermore, instances of epidural fever in women opting for labor analgesia are often linked to chorioamnionitis and placental inflammation (Sharma, 2000). Caution is warranted when interpreting the relationship between ERMF and elevated neutrophil percentages in this study since these neutrophil percentages represent the most recent tests conducted before PCEA. Previous study have demonstrated a lack of significant correlation between the development of puerperal fever and the levels of inflammatory factors during pregnancy (Arce et al., 2019), which contradicts the findings of the present study. Consequently, further studies are required to investigate the relationships among maternal fever, maternal inflammation levels, and potential sources of inflammation. It is yet to be determined whether fever during labor is linked to prenatal inflammation levels or if it is triggered by events during labor that amplify the maternal inflammatory response. Additionally, our study identified amniotic fluid contamination during membrane rupture, artificial membrane rupture, and post-analgesic antimicrobials as independent risk factors for ERMF, consistent with previous research on risk

factors for intrapartum fever (Jiang et al., 2021; Maayan-Metzger et al., 2006; Kim et al., 2021). Interestingly, the use of oxytocin before analgesia acted as a protective factor against ERMF, while its use after the initiation of analgesia emerged as a risk factor. Previous studies have indeed linked oxytocin use to an elevated risk of ERMF (Chang et al., 2023). However, given the scarcity of research regarding the timing of oxytocin use, further investigation is needed to fully understand its impact on ERMF.

To the best of our knowledge, this study is the first to use a machine model for predicting ERMF and to contrast it with a traditional regression model using extensive data on pregnancy history, clinical assessments, and pregnancy biochemical variables. However, there are certain limitations. Firstly, it was a retrospective study with inherent shortcomings, including inadequate documentation of maternal and fetal information by healthcare providers, resulting in incomplete data collection. The study's sample size is small, and it is a single-center study. To enhance the study's robustness, a larger, multi-center study should be considered in the future. Despite ERMF having an incidence of approximately 20%, maternal fever has significant adverse implications for maternal safety and neonatal wellbeing. Women with maternal fever are at an increased likelihood of receiving antibiotics, undergoing cesarean sections, and experiencing low Apgar scores, respiratory distress, hypotonia, neonatal convulsions, and neonatal encephalopathy, all of which have grave consequences (Dior et al., 2016; Towers et al., 2017). Hence, there is merit in developing predictive models for pregnant women with ERMF. Prospective studies could offer additional variables for prediction and enable a more precise assessment of ERMF. Subsequent prospective studies are necessary. Secondly, this study lacked external validation of the model, which may limit the broad application of the results. Therefore, external validation will be carried out in a multicenter prospective study. Lastly, the performance of several models in our study was only moderate, with an AUC of <80% for all models. Consequently, the validity of using these predictive models remains a subject of debate.

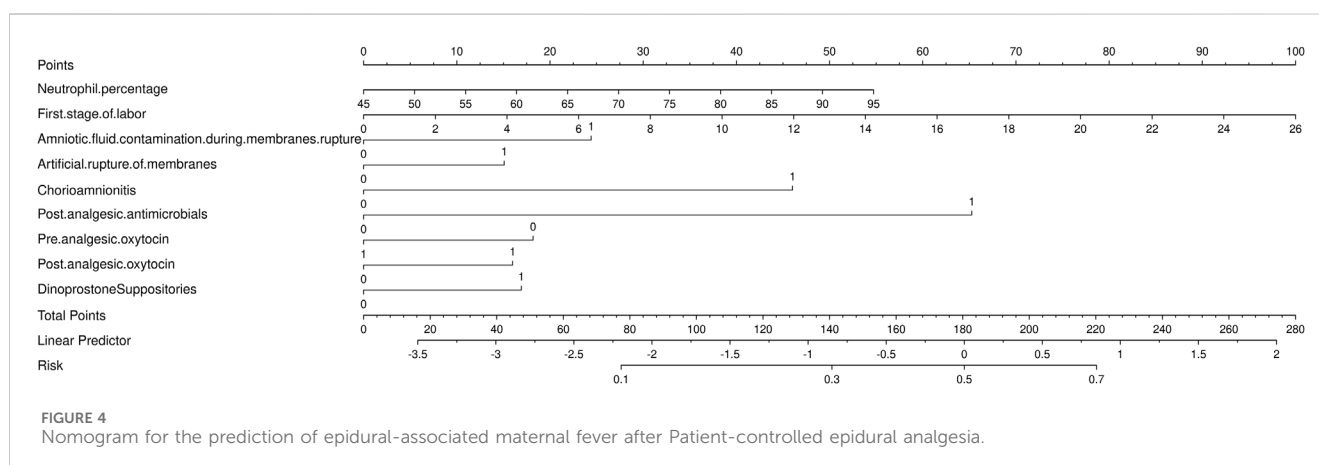
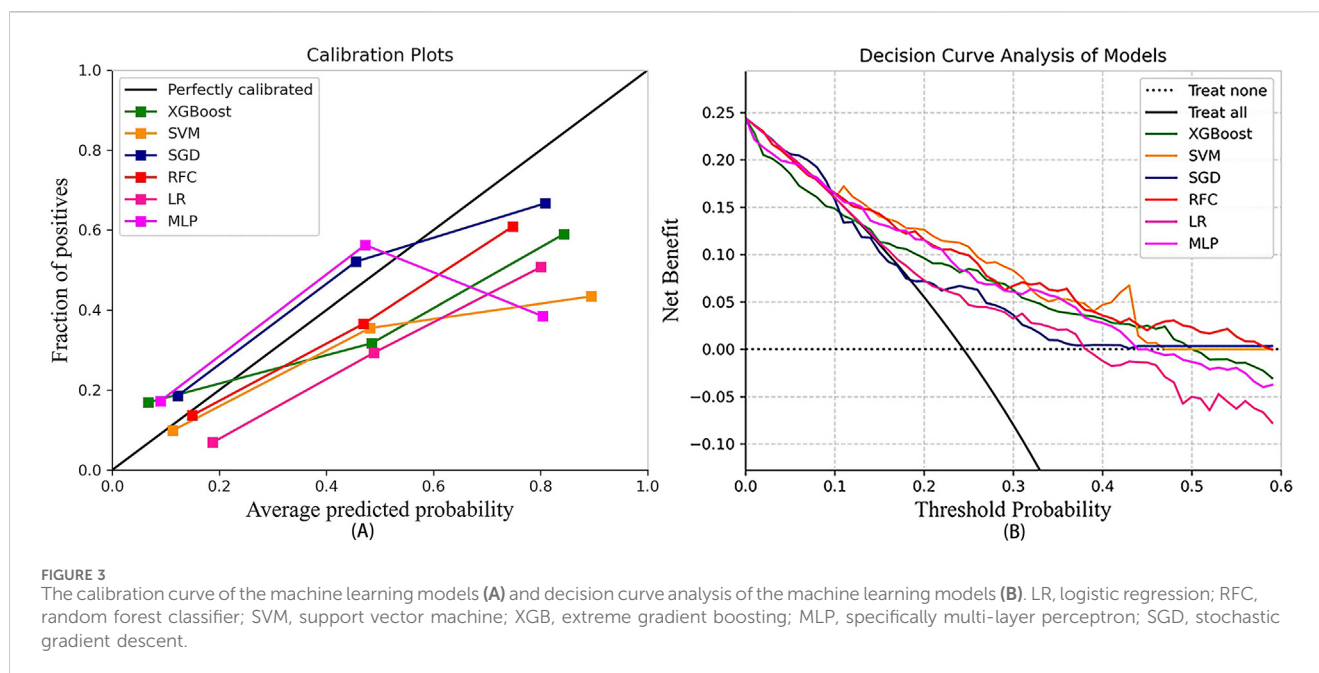
5 Conclusion

In summary, we developed six prediction models for ERMF using LR, RF, SVM, XGBoost, GBM, and MLP. Although the studied

TABLE 3 Discrimination and calibration of each machine learning algorithms on the testing set.

Model	AUC	Accuracy	Precision	Sensitivity	F1 score	Intercept	Slope	Brier
LR	0.791	0.70	0.44	0.79	0.57	−0.062	0.715	0.193
SVM	0.786	0.79	0.57	0.59	0.58	0.085	0.428	0.155
RFC	0.752	0.69	0.43	0.74	0.54	0.012	0.787	0.156
XGBoost	0.749	0.69	0.41	0.68	0.52	0.109	0.537	0.176
SGD	0.766	0.74	0.47	0.63	0.54	0.120	0.867	0.184
MLP	0.749	0.71	0.44	0.70	0.54	0.231	0.306	0.184

(AUC, area under curve of receiver operating characteristic; LR, logistic regression; RFC, random forest classifier; SVM, support vector machine; XGB, extreme gradient boosting; MLP, specifically multi-layer perceptron; SGD, stochastic gradient descent).



machine learning models can predict ERMF, they do not surpass the current gold standard logistic regression, which exhibits superior AUC, sensitivity, and interpretability. Additionally, while predictive models enable preoperative risk assessment and treatment decisions, they still necessitate clinical expertise. Furthermore, identifying predictors of specific factors, such as medication considerations and comorbid conditions, in predictive models could enhance perioperative care. This, in turn, enables timely interventions by physicians and nurses, ultimately leading to improved surgical outcomes.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Medical Ethics Committee of the Affiliated Drum Tower Hospital of Nanjing University Medical School (approval number 2022-106-04). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

XG: Writing – review and editing. HZ: Writing – review and editing, Writing – original draft. HM: Writing – review and editing, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study was funded by Jiangsu Pharmaceutical Association - HENGRUI hospital pharmacy fund scientific research project (No: H202107); Nanjing Pharmaceutical Association - CHANGZHOU SIYAO hospital pharmacy scientific research fund grant project (No: 2021YX004); National Natural Science Foundation of China (No: 81903870); Aid project of Jiangsu Ningai Medical Development and Medical Aid Foundation (NDYGN2024044).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Arce, D. Y., Bellavia, A., Cantonwine, D. E., Napoli, O. J., Meeker, J. D., James-Todd, T., et al. (2019). Average and time-specific maternal prenatal inflammatory biomarkers and the risk of labor epidural associated fever. *PLoS One* 14 (11), e0222958. doi:10.1371/journal.pone.0222958
- Banaei, M., Roozbeh, N., Darsareh, F., Mehrnough, V., Farashah, M. S. V., and Montazeri, F. (2025). Utilizing machine learning to predict the risk factors of episiotomy in parturient women. *AJOG Glob. Rep.* 5 (1), 100420. doi:10.1016/j.xagr.2024.100420
- Chang, X. Y., Wang, L. Z., Xia, F., and Zhang, Y. F. (2023). Factors associated with epidural-related maternal fever in low-risk term women: a systematic review. *Int. J. Obstet. Anesth.* 56, 103915. doi:10.1016/j.ijoa.2023.103915
- Christodoulou, E., Ma, J., Collins, G. S., Steyerberg, E. W., Verbakel, J. Y., and Van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J. Clin. Epidemiol.* 110, 12–22. doi:10.1016/j.jclinepi.2019.02.004
- Curtin, W. M., Katzman, P. J., Florescue, H., Metlay, L. A., and Ural, S. H. (2015). Intrapartum fever, epidural analgesia and histologic chorioamnionitis. *J. Perinatol.* 35 (6), 396–400. doi:10.1038/jp.2014.235
- Darsareh, F., Ranjbar, A., Farashah, M. V., Mehrnough, V., Shekari, M., and Jahromi, M. S. (2023). Application of machine learning to identify risk factors of birth asphyxia. *BMC Pregnancy Childbirth* 23 (1), 156. doi:10.1186/s12884-023-05486-9
- Del Arroyo, A. G., Sanchez, J., Patel, S., Phillips, S., Reyes, A., Cubillos, C., et al. (2019). Role of leucocyte caspase-1 activity in epidural-related maternal fever: a single-centre, observational, mechanistic cohort study. *Br. J. Anaesth.* 122 (1), 92–102. doi:10.1016/j.bja.2018.09.024
- Dior, U. P., Kogan, L., Eventov-Friedman, S., Gil, M., Bahar, R., Ergaz, Z., et al. (2016). Very high intrapartum fever in term pregnancies and adverse obstetric and neonatal outcomes. *Neonatology* 109 (1), 62–68. doi:10.1159/000440938
- Fusi, L., Steer, P. J., Maresh, M. J., and Beard, R. W. (1989). Maternal pyrexia associated with the use of epidural analgesia in labour. *Lancet* 1 (8649), 1250–1252. doi:10.1016/s0140-6736(89)92341-6
- Goecks, J., Jalili, V., Heiser, L. M., and Gray, J. W. (2020). How machine learning will transform biomedicine. *Cell* 181 (1), 92–101. doi:10.1016/j.cell.2020.03.022
- Goetzl, L., Cohen, A., Frigoletto, F., Ringer, S. A., Lang, J. M., and Lieberman, E. (2001). Maternal epidural use and neonatal sepsis evaluation in afebrile mothers. *Pediatrics* 108 (5), 1099–1102. doi:10.1542/peds.108.5.1099
- Impey, L. W. M., Greenwood, C. E. L., Black, R. S., Yeh, P. S. Y., Sheil, O., and Doyle, P. (2008). The relationship between intrapartum maternal fever and neonatal acidosis as risk factors for neonatal encephalopathy. *Am. J. Obstet. Gynecol.* 198 (1), 49.e1–e6. doi:10.1016/j.ajog.2007.06.011
- Jiang, Z., Hu, X., Zeng, H., Wang, X., Tan, C., Ni, C., et al. (2021). Nomogram for perinatal prediction of intrapartum fever: a retrospective case-control study. *BMC Pregnancy Childbirth* 21 (1), 445. doi:10.1186/s12884-021-03891-6
- Kim, S. Y., Hong, S. Y., Kwon, D. Y., Park, H., Choi, S. J., Oh, S. Y., et al. (2021). Degree of intrapartum fever and associated factors: three group analysis of no fever, borderline and overt fever. *J. Obstet. Gynaecol. Res.* 47 (3), 1153–1163. doi:10.1111/jog.14651
- Lai, Z., Chen, X., Zhang, J., Kong, H., and Wen, J. (2023). Maximal margin support vector machine for feature representation and classification. *IEEE Trans. Cybern.* 53, 6700–6713. doi:10.1109/TCYB.2022.3232800
- Maayan-Metzger, A., Mazkereth, R., Shani, A., and Kuint, J. (2006). Risk factors for maternal intrapartum fever and short-term neonatal outcome. *Fetal Pediatr. Pathol.* 25 (3), 169–177. doi:10.1080/15513810600908461
- Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *J. Thorac. Oncol.* 5 (9), 1315–1316. doi:10.1097/JTO.0b013e3181ec173d
- Nguyen, B., Morell, C., and De Baets, B. (2020). Scalable large-margin distance metric learning using stochastic gradient descent. *IEEE Trans. Cybern.* 50 (3), 1072–1083. doi:10.1109/TCYB.2018.2881417
- Nwanosike, E. M., Conway, B. R., Merchant, H. A., and Hasan, S. S. (2022). Potential applications and performance of machine learning techniques and algorithms in clinical practice: a systematic review. *Int. J. Med. Inf.* 159, 104679. doi:10.1016/j.ijmedinf.2021.104679
- Oosterhoff, J. H. F., Savelberg, ABMC, Karhade, A. V., Gravesteijn, B. Y., Doornberg, J. N., Schwab, J. H., et al. (2022). Development and internal validation of a clinical prediction model using machine learning algorithms for 90 day and 2 year mortality in femoral neck fracture patients aged 65 years or above. *Eur. J. Trauma Emerg. Surg.* 48 (6), 4669–4682. doi:10.1007/s00068-022-01981-4
- Ren, J., Wang, T., Yang, B., Jiang, L., Xu, L., Geng, X., et al. (2021). Risk factors and safety analyses for intrapartum fever in pregnant women receiving epidural analgesia during labor. *Med. Sci. Monit.* 27, e929283. doi:10.12659/MSM.929283
- Rigatti, S. J. (2017). Random forest. *J. Insur Med.* 47 (1), 31–39. doi:10.17849/insm-47-01-31-39.1
- Roozbeh, N., Montazeri, F., Farashah, M. V., Mehrnough, V., and Darsareh, F. (2025). Proposing a machine learning-based model for predicting nonreassuring fetal heart. *Sci. Rep.* 15 (1), 7812. doi:10.1038/s41598-025-92810-2
- Seijmonsbergen-Schermers, A. E., van den Akker, T., Rydahl, E., Beeckman, K., Bogaerts, A., Binf, L., et al. (2020). Variations in use of childbirth interventions in 13 high-income countries: a multinational cross-sectional study. *PLoS Med.* 17 (5), e1003103. doi:10.1371/journal.pmed.1003103
- Sharma, S. K. (2000). Epidural analgesia during labor and maternal fever. *Curr. Opin. Anaesthesiol.* 13 (3), 257–260. doi:10.1097/00001503-200006000-00003
- Sharma, S. K., Rogers, B. B., Alexander, J. M., McIntire, D. D., and Leveno, K. J. (2014). A randomized trial of the effects of antibiotic prophylaxis on epidural-related fever in labor. *Anesth. Analg.* 118 (3), 604–610. doi:10.1213/ANE.0b013e3182a5d539

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphar.2025.1614770/full#supplementary-material>

- Sultan, P., David, A. L., Fernando, R., and Ackland, G. L. (2016). Inflammation and epidural-related maternal fever: proposed mechanisms. *Anesth. Analg.* 122 (5), 1546–1553. doi:10.1213/ANE.0000000000001195
- Sultan, P., and Segal, S. (2020). Epidural-related maternal fever: still a hot topic, but what are the burning issues? *Anesth. Analg.* 130 (2), 318–320. doi:10.1213/ANE.0000000000004576
- Towers, C. V., Yates, A., Zite, N., Smith, C., Chernicky, L., and Howard, B. (2017). Incidence of fever in labor and risk of neonatal sepsis. *Am. J. Obstet. Gynecol.* 216 (6), 596.e1–596. doi:10.1016/j.ajog.2017.02.022
- Wang, H., Yang, Z., Wei, S., Xia, L., Li, Y., Wu, X., et al. (2023). Perinatal outcomes and risk factors for epidural analgesia-associated intrapartum maternal fever: a retrospective study. *J. Matern. Fetal Neonatal Med.* 36 (1), 2179383. doi:10.1080/14767058.2023.2179383
- Yin, H., and Hu, R. (2022). Risk factors of maternal intrapartum fever and the effect of fever duration on neonatal morbidity in different temperature. *J. Obstet. Gynaecol. Res.* 48 (10), 2522–2527. doi:10.1111/jog.15362
- Zhao, B., Li, B., Wang, Q., and Song, X. (2022). 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.* 35 (25), 5354–5362. doi:10.1080/14767058.2021.1879042