

Low-Cost Early Prediction of Retinopathy of Prematurity Using Arterial Blood Gas Analysis: A Retrospective Cohort Study

Chaoqin Yu

Xiangya Hospital Central South University

Qianyu Hao

Hunan Diantou Education Technology Co., Ltd.

Moyan Liu

Xiangya Hospital Central South University

Junfeng Mao

Xiangya Hospital Central South University

Yuanyuan Chen

The Central Hospital of Shaoyang

Xiao Chen

The First People's Hospital of Changde

Junlong Tang

Hunan Diantou Education Technology Co., Ltd.

Xiaobo Xia

Xiangya Hospital Central South University

Huizhuo Xu

Xiangya Hospital Central South University

Wei Zheng

`zhengwei inxy@csu.edu.cn`

Xiangya Hospital Central South University

Research Article

Keywords: retinopathy of prematurity, machine learning, prediction, arterial blood gas analysis

Posted Date: October 19th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10591562/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Objective The objective was to develop a cost-efficient, easily implementable, and scalable predictive model for retinopathy of prematurity, ensuring effective prediction while constraining both the costs associated with data acquisition and the complexity of clinical deployment.

Setting Neonatal units of three medical institutions in Hunan, China

Patients Neonates born during the period from 2022 to 2025. Data were retrospectively collected from three medical institutions: Xiangya Hospital of Central South University, The Central Hospital of Shaoyang, and The first people's hospital of Changde city.

Main outcome measures Accuracy, Sensitivity, Specificity, PPV, NPV, AUC, F1 Score

Methods A cohort study was conducted using 584 neonatal records from Xiangya Hospital of Central South University. Correlations between blood gas parameters and ROP were first analyzed. Three routine postnatal monitoring variables (gestational age, birth weight, and arterial blood gas analysis results) obtained within 42 days after birth were selected via Pearson correlation analysis. Ten machine learning algorithms were trained and compared to identify the optimal model with the simplest data requirements and best predictive performance. Model effectiveness was subsequently evaluated through external validation at two independent medical centers.

Results Among ten machine learning models, XGBoost demonstrated the optimal predictive performance. During XGBoost model training, excessive inclusion of blood gas analysis features introduced noise and compromised model efficacy. Restricting input to the top 10 most clinically relevant features yielded the best performance, with an AUROC of 0.8512, accuracy of 0.82, sensitivity of 0.9513, and negative predictive value of 0.9129. External validation further confirmed the model's generalization robustness: across two independent validation cohorts, sensitivity was 0.8333 and 0.9167, and AUROC was 0.8576 and 0.7319, respectively.

Conclusions By relying solely on blood gas parameters and minimal clinical features, this model facilitates early ROP detection.

1 Introduction

Retinopathy of Prematurity (ROP) is a devastating neurovascular disease of the retina that can lead to vision deficits or even blindness and vary widely in both severity and prognosis^[1-2]. In 2019, the global age-standardized rates of prevalence per 100,000 population was 86.4 for vision loss, specifically, 31.6 for blindness due to ROP among people younger than 20 years^[3]. ROP has become one of the leading global causes of blindness in neonates^[4]. Early treatment improved visual acuity outcomes at 6 years of age^[5]. Currently, the clinical application of RetCam for neonatal ophthalmic screening and Artificial Intelligence (AI)-based retinal image analysis techniques^[6-8] are progressively becoming the primary screening modalities.

Despite tailored screening guidelines, only a minority of screened infants ultimately require treatment^[9-11], current screening criteria may result in numerous unnecessary examinations^[12-13], and accurately predicting which patients require intervention remains an unsolved clinical problem. Machine learning (ML)^[14] has

emerged as a promising tool for advancing the prediction of ROP by identifying subtle patterns of clinical deterioration within large-scale datasets. For instance, PW-ROP^[15] developed using an Asian cohort, incorporated only gestational age (GA) and birth weight (BW) and demonstrated excellent performance in internal validation involving 1,043 neonates. WINROP^[16] relies on longitudinal monitoring of weekly weight and serum insulin-like growth factor-1 (IGF-1) levels postnatally, providing warnings 5–7 weeks before treatment is needed and achieving 100% sensitivity in a Swedish cohort. However, this model requires repeated blood sampling for IGF-1 measurement, limiting its clinical adoption due to procedural complexity and cost. The applicability of most existing models is restricted to infants with a gestational age of < 32 weeks^[16–21]. However, emerging evidence indicates that ROP may also occur in more mature infants^[22], highlighting a gap in extending model applicability. Previous research confirms that low GA and low BW are the most significant risk factors for ROP^[23–25]. Hyperoxia may induce retinal vasoconstriction and obliteration, leading to hypoxic tissue injury^[26–28]. Variations in multiple arterial blood gas analysis (ABGA) parameters—including partial pressure of carbon dioxide (PCO₂), oxygen saturation, oxygen fluctuations^[29–34], lactate levels^[35], and lower pH values—are closely linked to severe ROP^[36–37]. These findings suggest that ABGA may have a potential predictive value for ROP.

Collectively, ABGA is routinely monitored in infants and offers high standardization and accessibility. This study proposes an ROP risk prediction model using machine learning algorithms, incorporating multi-timepoint ABGA parameters alongside GA and BW. Our objective is to achieve precise ROP risk prediction, simplify data acquisition, provide a critical time window for clinical intervention, and ultimately reduce long-term visual impairment in high-risk infants.

2 Results

2.1 Baseline characteristics

The final training cohort comprised 584 infants (1,168 eyes; 5,422 arterial blood gas analyses), of whom 196 (33.56%) developed retinopathy of prematurity (ROP) during hospitalization, including 45 cases (7.71%) with plus disease. External validation set 1 included 74 infants (148 eyes; 546 arterial blood gas analyses), with 31 (41.89%) developing ROP and 23 (31.08%) exhibiting plus disease. External validation set 2 included 35 infants (70 eyes; 309 arterial blood gas analyses), of whom 12 (34.29%) developed ROP during their hospital stay. The baseline characteristics of the ROP and non-ROP groups within the training cohort are summarized in Table 1. There was no statistically significant difference in gender distribution between the groups ($p > 0.05$). Infants in the ROP group generally had a lower gestational age and a lower birth weight. Significant differences in arterial blood gas analysis within the first 42 days of life were observed between the two groups. Regarding vital signs, body temperature showed a slight but significant difference (36.5 [36.3, 37.0] vs. 36.7 [36.5, 37.0]°C, $p < 0.05$). Core acid-base balance indicators differed significantly between ROP and non-ROP groups, including pH (7.3 [7.2, 7.3] vs. 7.3 [7.3, 7.4], $p < 0.001$), pCO₂ (46.6 [39.0, 53.9] vs. 40.8 [34.0, 49.0] mmHg, $p < 0.001$), BEecf (-3.0 [-5.3, -1.8] vs. -1.8 [-1.8, -1.8] mmol/L, $p < 0.001$), BE(B) (-2.4 [-5.4, 1.7] vs. -1.7 [-1.7, -1.7] mmol/L, $p < 0.001$) and HCO₃⁻ (21.9 [20.1, 23.5] vs. 20.8 [18.8, 23.0] mmol/L, $p < 0.001$). Furthermore, electrolytes used to calculate the anion gap $AG = Na^{+} + K^{+} - (Cl^{-} + HCO_{3}^{-})$ for identifying types of metabolic acidosis were notably different: Na⁺ (133.0 [131.0, 135.0] vs. 136.0 [134.0, 138.0] mmol/L, $p < 0.001$), K⁺ (4.5 [4.1, 4.9] vs. 4.3 [4.1, 4.7] mmol/L, $p <$

0.001), and Cl^- (105.0 [103.0, 108.0] vs. 109.0 [106.0, 112.0] mmol/L, $p < 0.001$). Indicators reflecting oxygenation and oxygen-carrying capacity also exhibited significant disparities, such as alveolar partial pressure of oxygen (PaO_2) (89.5 [67.0, 107.5] vs. 101.9 [84.0, 112.8] mmHg, $p < 0.001$), fraction of inspired oxygen (FiO_2) (21.0 [21.0, 37.1] vs. 21.0 [21.0, 21.0]%, $p < 0.001$), total hemoglobin (ctHb) (15.6 [14.2, 17.2] vs. 16.4 [15.1, 18.1] g/dL, $p < 0.001$), fO_2Hb (93.7 [90.8, 95.1]% vs. 94.5 [91.3, 96.1]%, $p < 0.001$), fCOHb (1.7 [1.3, 2.1]% vs. 1.1 [0.9, 1.6]%, $p < 0.001$), fMetHb (1.1 [0.8, 1.3]% vs. 0.6 [0.5, 0.9]%, $p < 0.001$), and Hct (46.0 [41.0, 51.0]% vs. 49.7 [45.5, 54.3]%, $p < 0.001$). Additionally, metabolic and perfusion markers varied significantly, including lactate (2.0 [1.3, 3.1] vs. 2.4 [1.5, 3.6] mmol/L, $p < 0.01$) and glucose (3.6 [2.6, 4.7] vs. 3.9 [3.0, 5.2] mmol/L, $p < 0.05$). Total bilirubin (tBil) levels were significantly different between the cohorts (34.0 [34.0, 40.0] vs. 36.0 [34.0, 53.0] $\mu\text{mol/L}$, $p < 0.001$). Finally, no significant differences were observed between the two groups across several blood gas parameters, including fractional carboxyhemoglobin (fHHb), ionized calcium (Ca^{2+}), oxygen saturation (sO_2) and anion gap (AG).

Table 1
Baseline characteristics of the study population

Parameters	ROP	Non-ROP	p-value
GA [M(IQR)]	201.0 (192.0, 213.0)	243.5 (216.0, 270.0)	< 0.001
BW [M(IQR)]	1080.0 (900.0, 1342.5)	2140.0 (1370.0, 3182.5)	< 0.001
Gender [F(n):M(n)]	102:94	169:219	0.0638
T [M(IQR)]	36.5 (36.3, 37.0)	36.7 (36.5, 37.0)	< 0.05
pH [M(IQR)]	7.3 (7.2, 7.3)	7.3 (7.3, 7.4)	< 0.001
pCO ₂ [M(IQR)]	46.6 (39.0, 53.9)	40.8 (34.0, 49.0)	< 0.001
HCO ₃ ⁻ [M(IQR)]	21.9 (20.1, 23.5)	20.8 (18.8, 23.0)	< 0.001
K ⁺ [M(IQR)]	4.5 (4.1, 4.9)	4.3 (4.1, 4.7)	< 0.001
Cl ⁻ [M(IQR)]	105.0 (103.0, 108.0)	109.0 (106.0, 112.0)	< 0.001
Na ⁺ [M(IQR)]	133.0 (131.0, 135.0)	136.0 (134.0, 138.0)	< 0.001
ctHb [M(IQR)]	15.6 (14.2, 17.2)	16.4 (15.1, 18.1)	< 0.001
FO ₂ Hb [M(IQR)]	93.7 (90.8, 95.1)	94.5 (91.3, 96.1)	< 0.001
fCOHb [M(IQR)]	1.7 (1.3, 2.1)	1.1 (0.9, 1.6)	< 0.001
fMetHb [M(IQR)]	1.1 (0.8, 1.3)	0.6 (0.5, 0.9)	< 0.001
Hct [M(IQR)]	46.0 (41.0, 51.0)	49.7 (45.5, 54.3)	< 0.001
Lac [M(IQR)]	2.0 (1.3, 3.1)	2.4 (1.5, 3.6)	< 0.01
Glu [M(IQR)]	3.6 (2.6, 4.7)	3.9 (3.0, 5.2)	< 0.05
tBil [M(IQR)]	34.0 (34.0, 40.0)	36.0 (34.0, 53.0)	< 0.001
fHHb [M(IQR)]	3.3 (2.1, 6.1)	3.3 (1.9, 6.8)	0.8502
FiO ₂ [M(IQR)]	21.0 (21.0, 37.1)	21.0 (21.0, 21.0)	< 0.001
Ca ²⁺ [M(IQR)]	1.3 (1.3, 1.4)	1.4 (1.3, 1.4)	0.1217
BEecf [M(IQR)]	-3.0 (-5.3, -1.8)	-1.8 (-1.8, -1.8)	< 0.001
BE(B) [M(IQR)]	-2.4 (-5.4, -1.7)	-1.7 (-1.7, -1.7)	< 0.001
AG [M(IQR)]	10.0 (9.0, 11.0)	10.0 (10.0, 10.0)	0.8706

Continuous variables are shown as median [IQR], and categorical variables as counts (percentages). fHHb: Fraction of reduced hemoglobin (deoxyhemoglobin), FiO₂: Fraction of inspired oxygen, Ca²⁺: Ionized calcium, AG: Anion gap, BEecf: Base excess in extracellular fluid, HCO₃⁻: Bicarbonate ion, pCO₂: Partial pressure of carbon dioxide, Cl⁻: Chloride ion, BE(B): Standard base excess, T: Temperature, ctHb: Total hemoglobin concentration, sO₂: Oxygen saturation, Hct: Hematocrit, fMetHb: Fraction of methemoglobin, fCOHb: Fraction of carboxyhemoglobin, fO₂Hb: Fraction of oxyhemoglobin, PaO₂: Arterial partial pressure of oxygen, K⁺: Potassium ion, pH: Negative logarithm of hydrogen ion activity, Lac: Lactate, Glu: Glucose, Na⁺: Sodium ion, tBil: Total bilirubin

Parameters	ROP	Non-ROP	p-value
sO ₂ [M(IQR)]	96.3 (93.3, 97.9)	96.6 (93.1, 98.1)	0.5435
PaO ₂ [M(IQR)]	89.5 (67.0, 107.5)	101.9 (84.0, 112.8)	< 0.001
Continuous variables are shown as median [IQR], and categorical variables as counts (percentages). fHHb: Fraction of reduced hemoglobin (deoxyhemoglobin), FiO ₂ : Fraction of inspired oxygen, Ca ²⁺ : Ionized calcium, AG: Anion gap, BEecf: Base excess in extracellular fluid, HCO ₃ ⁻ : Bicarbonate ion, pCO ₂ : Partial pressure of carbon dioxide, Cl ⁻ : Chloride ion, BE(B): Standard base excess, T: Temperature, ctHb: Total hemoglobin concentration, sO ₂ : Oxygen saturation, Hct: Hematocrit, fMetHb: Fraction of methemoglobin, fCOHb: Fraction of carboxyhemoglobin, fO ₂ Hb: Fraction of oxyhemoglobin, PaO ₂ : Arterial partial pressure of oxygen, K ⁺ : Potassium ion, pH: Negative logarithm of hydrogen ion activity, Lac: Lactate, Glu: Glucose, Na ⁺ : Sodium ion, tBil: Total bilirubin			

2.2 Feature selection

Feature selection began with gestational age (GA), birth weight (BW), postmenstrual age (PMA), and core arterial blood gas analysis (ABGA) parameters associated with ROP. Pearson correlation coefficient (Fig. 1) and SHAP analysis (Fig. 2) were used for feature importance screening and contribution analysis. Variables were ranked by importance, and iteratively incorporated into ROP prediction models to identify the optimal balance between performance and parsimony. The top ten features yielded the best predictive performance. Three feature sets were constructed: a four-feature set (gender, GA, BW, PMA); a ten-feature set (GA, BW, PMA, Hct, ctHb, HCO₃⁻, fCOHb, Lac, fMetHb, pCO₂); and an all-feature set comprising all 28 variables. These sets formed the structural basis for subsequent model development and validation.

2.3 Predictive performance

The predictive performance of ten machine learning algorithms was systematically evaluated across three distinct feature sets, with results presented in Table 2 and Fig. 3: a "ten-feature set" (comprising 10 key variables), a "four-feature set" (limited to core indicators), and an "all-feature set" (incorporating all available variables). Gradient-boosted decision trees, specifically XGBoost and CatBoost, demonstrated superior overall performance, particularly within the ten-feature set. In this configuration, XGBoost achieved the highest performance metrics, with an AUROC of 0.8512, sensitivity of 0.9513, and an F1 score of 0.8600, indicating exceptional discriminative power. CatBoost performed comparably (AUROC 0.8321, sensitivity 0.8020, F1 0.8000), exhibiting strong robustness and generalizability. Notably, the inclusion of additional variables in the all-feature set did not yield significant performance gains over the ten-feature set. For instance, the AUROC for XGBoost decreased to 0.8044; while sensitivity remained high (0.9664), specificity dropped markedly to 0.4589, resulting in a lower positive predictive value (PPV) of 0.6973. This suggests that redundant or noisy features may introduce overfitting risks, potentially undermining model stability and clinical utility. Consequently, the ten-feature set offered the optimal balance between predictive accuracy and parsimony. The four-feature set maintained competitive efficacy despite its reduced dimensionality. CatBoost achieved an AUROC of 0.7685 (sensitivity 0.7534, F1 0.74), while the Multi-layer Perceptron (MLP) reached an AUROC of 0.7566 and sensitivity of 0.8305. These results indicate that a minimal subset of highly relevant clinical variables is sufficient for developing viable early-warning tools, particularly for resource-limited settings or real-time decision-making where model deployability is paramount. Regarding model architecture, gradient-boosting algorithms consistently outperformed traditional methods such as KNN, Naive Bayes, and Logistic Regression, especially in capturing non-linear interactions. However, Logistic Regression maintained respectable

performance on the 10-feature set (AUROC 0.7626, sensitivity 0.797), offering high interpretability suitable for clinical pathway integration. Across all configurations, models exhibited high sensitivity (up to 0.9513) and negative predictive value (NPV) (up to 0.9129). This high NPV underscores the models' strength in accurately identifying low-risk individuals and minimizing missed diagnoses—a critical requirement for Retinopathy of Prematurity screening.

Table 2
Performance metrics of machine learning models across feature sets

Set	Model	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC	F1 Score
all-feature set	XGB	0.7400	0.9664	0.4589	0.6973	0.9138	0.8044	0.8100
	CB	0.7700	0.8171	0.7121	0.7855	0.7511	0.8008	0.8000
	MLP	0.7193	0.8255	0.5823	0.7182	0.7212	0.7818	0.7700
	DT	0.7400	0.7433	0.7468	0.7911	0.6928	0.7822	0.7700
	RF	0.6900	0.6443	0.7446	0.7649	0.6187	0.7645	0.7000
	SVM	0.6900	0.7265	0.6450	0.7253	0.6464	0.7698	0.6900
	LR	0.7000	0.7836	0.5844	0.7086	0.6767	0.7595	0.7400
	NB	0.7200	0.9044	0.4827	0.6928	0.7964	0.7437	0.7800
	KNN	0.7100	0.7785	0.6147	0.7227	0.6827	0.7850	0.7500
	LGB	0.7000	0.7752	0.6104	0.7196	0.6779	0.7321	0.7500
ten-feature set	XGB	0.8200	0.9513	0.6580	0.7821	0.9129	0.8512	0.8600
	CB	0.7800	0.8020	0.7403	0.7993	0.7435	0.8321	0.8000
	MLP	0.7344	0.8456	0.5909	0.7273	0.7479	0.8100	0.7800
	DT	0.7300	0.7064	0.7532	0.7869	0.6654	0.7765	0.7300
	RF	0.7200	0.7131	0.7403	0.7798	0.6667	0.8017	0.7400
	SVM	0.7400	0.9564	0.4697	0.6994	0.8930	0.7703	0.7200
	LR	0.7000	0.7970	0.5779	0.7090	0.6881	0.7626	0.7000
	NB	0.7000	0.8809	0.4740	0.6836	0.7552	0.7563	0.6900
	KNN	0.7000	0.8070	0.5606	0.7032	0.6925	0.7475	0.6900
	LGB	0.6900	0.7466	0.6061	0.7097	0.6497	0.7345	0.7300
four-feature set	XGB	0.7400	0.9228	0.4978	0.7033	0.8333	0.7253	0.8000
	CB	0.7400	0.7534	0.7273	0.7809	0.6957	0.7685	0.7400
	MLP	0.7100	0.8305	0.5541	0.7061	0.7171	0.7566	0.7000
	DT	0.7100	0.7785	0.6277	0.7296	0.6872	0.7560	0.7100
	RF	0.6900	0.7198	0.6494	0.7259	0.6424	0.7200	0.6900
XGB eXtreme Gradient Boosting, CB Categorical Boosting, MLP Multilayer Perceptron, DT Decision Tree, RF Random Forest, SVM Support Vector Machine, LR Logistic Regression, NB Naive Bayes, KNN K-Nearest Neighbors, LGB Light Gradient Boosting Machine, PPV Positive Predictive Value, NPV Negative Predictive Value, AUC Area Under the Curve								

Set	Model	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC	F1 Score
	SVM	0.6900	0.7265	0.6450	0.7253	0.6464	0.7698	0.6900
	LR	0.6000	0.7000	0.4700	0.6300	0.5500	0.6942	0.5900
	NB	0.6500	0.8389	0.4091	0.6468	0.6632	0.6796	0.6300
	KNN	0.6200	0.6426	0.5887	0.6684	0.5608	0.6815	0.6200
	LGB	0.6600	0.7349	0.5584	0.6822	0.6202	0.6847	0.6600
XGB eXtreme Gradient Boosting, CB Categorical Boosting, MLP Multilayer Perceptron, DT Decision Tree, RF Random Forest, SVM Support Vector Machine, LR Logistic Regression, NB Naive Bayes, KNN K-Nearest Neighbors, LGB Light Gradient Boosting Machine, PPV Positive Predictive Value, NPV Negative Predictive Value, AUC Area Under the Curve								

In summary, the XGBoost model trained on the 10-feature set emerged as the optimal configuration, achieving an AUROC of 0.8512, sensitivity of 0.9513, specificity of 0.658, and an F1 score of 0.86. Two inference strategies were implemented during the model development phase: mean aggregation, which averages the probabilistic outputs across multiple data entries to inform decision-making, and max aggregation, which utilizes the maximum value across entries for final prediction. In the external validation phase, we maintained these two aggregation approaches to ensure consistency with the development cohort. The results demonstrated robust performance across diverse datasets: in External Validation Set 1 (n = 74), the model achieved a sensitivity of 0.8333 and an AUROC of 0.8576; in External Validation Set 2 (n = 35), the sensitivity was 0.9167 with an AUROC of 0.7319 (Table 3). The performance drop in External Validation Set 2 may be attributed to the smaller sample size (n = 35) and potential distributional differences across centers. This combination validates the effectiveness of a "less is more" feature selection strategy in clinical modeling. Given its high sensitivity and NPV, this model is a primary candidate for early ROP risk assessment, supporting the clinical objective of early detection and intervention.

Table 3 Performance metrics of external validation sets								
Set	Inference strategy	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC	F1 Score
external validation set 1	Mean aggregation	0.8108	0.8333	0.7955	0.7353	0.8750	0.8576	0.7821
	Max aggregation	0.7973	0.8333	0.7727	0.7143	0.8718	0.8523	/
external validation set 2	Mean aggregation	0.6857	0.7500	0.6522	0.5294	0.8333	0.6812	0.6027
	Max aggregation	0.6571	0.9167	0.5217	0.5000	0.9231	0.7319	/
PPV Positive Predictive Value, NPV Negative Predictive Value, AUC Area Under the Curve								

3 Discussion

The prevalence of ROP in our dataset was 33.56%. Recognizing that many primary care facilities lack specialized fundus imaging equipment, we intentionally selected routine clinical parameters to enhance applicability in resource-limited settings. This study is the first to propose a comprehensive blood gas analysis profile, integrated with gestational age and birth weight, as a predictive framework for ROP. Our machine learning model achieved high diagnostic accuracy (AUROC 0.8512, sensitivity 0.9513) using admission-level features and a single routine monitoring parameter, enabling immediate application for early risk stratification. This model features a parsimonious architecture with robust performance, achieving an accuracy of 0.8200 in ROP prediction, thereby substantiating its potential as a reliable clinical screening tool. Existing ROP prediction models exhibit substantial heterogeneity in feature selection and predictive performance. Furthermore, a considerable number of these models rely on real-time weight gain data or dynamically fluctuating laboratory biomarkers, thereby constraining their generalizability and clinical applicability in primary care settings. For instance, the ROPScore^[17] calculation incorporates risk factors such as weight gain rate through six weeks post-birth, blood transfusions, and the use of mechanical ventilation. Similarly, the algorithm developed by Chatarina Löfqvist et al.^[16] requires weekly monitoring of weight gain alongside dynamic fluctuations in IGF-I and IGFBP-3 levels from birth until discharge. Although this algorithm provides early warnings 5 to 7 weeks prior to treatment (ACC 0.848, sensitivity 1.00, specificity 0.836), it necessitates additional laboratory testing and increases the diagnostic burden on primary care facilities. Furthermore, the narrow gestational age range of existing models—most of which are applicable only to preterm infants with a GA below 32 weeks—limits their utility in more mature neonates. For example, while ROPScore^[17] has not been validated in infants with a GA > 31 weeks. Similarly, WINROP^[16] excludes infants ≥ 32 weeks, and CHOP ROP^[19] was established based on screening criteria of BW < 1500 g or GA < 30 weeks. However, recent studies indicate that ROP can be diagnosed in more mature infants^[22]; consequently, our model does not exclude infants above 32 weeks GA.

SHAP analysis elucidated the model's decision-making process, revealing patterns that refine established clinical knowledge. Integration of blood gas parameters within the first 42 days enhanced predictive performance. Consistent with prior evidence, pCO₂ demonstrated high predictive value, while the model additionally identified significant associations between ROP and fCOHb, fO₂Hb, fMetHb, Hct, and ctHb. These findings align with the fundamental pathophysiology of ROP, driven by an imbalance between retinal vascular development and oxygen metabolism. In mammals, oxygen transport is optimized at the systemic level through the tight regulation of ventilation and red cell mass.^[38] Complementing this macro-regulation, internal erythrocyte metabolic adaptation is equally critical: the modulation of oxygen affinity mediated by oxyhemoglobin (fO₂Hb) and 2,3-bisphosphoglycerate (2,3-DPG) represents a core mechanism for hypoxic adaptation. Carboxyhemoglobin (fCOHb), possessing an affinity for hemoglobin approximately 250 times greater than that of oxygen, competitively inhibits oxygenation and exacerbates tissue hypoxia. Furthermore, elevated fCOHb concentrations can directly inhibit mitochondrial cytochrome c oxidase activity, disrupting cellular energy metabolism and accelerating ischemic injury.^[39] Beyond oxygen affinity modulation, the accumulation of methemoglobin (fMetHb) and the resulting disequilibrium between oxidative stress and antioxidant defense constitute another metabolic pillar of ROP-related vascular injury. The auto-oxidation of hemoglobin, involving the conversion of Fe²⁺ to Fe³⁺, is accompanied by the generation of superoxide anions, which trigger an oxidative stress cascade. The accumulation of fMetHb promotes apoptosis in retinal vascular endothelial cells and induces pathological neovascularization. In summary, parameters such as Hct, ctHb, fCOHb, fO₂Hb, and fMetHb are not isolated clinical indicators; rather, they are interconnected through multiple

pathways, including gas transport efficiency, glycolytic regulation, oxidative stress response, and energy metabolism. Together, they reflect the erythrocyte metabolic status and the degree of tissue hypoxia in preterm infants. In this machine learning-based ROP prediction models, integrating these metrics not only improves the detection sensitivity for hypoxia-hyperoxia injury but also helps elucidate the intrinsic links between individualized metabolic phenotypes and disease susceptibility, providing a theoretical foundation for precision oxygen therapy. The capacity to capture these intricate relationships, which are often challenging to delineate using traditional scoring systems, represents a key advantage of the present computational approach.

It is crucial to emphasize that this model is designed as a risk stratification tool based on gestational age, birth weight, and blood gas analysis within the first 42 days of life, rather than as a real-time diagnostic tool. The model achieved an Accuracy of 0.8200 and an AUC of 0.8512, demonstrating robust performance. Although the specificity was 0.6580, this remains a recognized challenge when predicting low-prevalence clinical events. For context, Eckert et al. reported specificities of 0.26 for any-stage ROP and 0.56 for severe ROP, respectively. Furthermore, the specificity of the CHOP^[18] and DIGIROP^[21] models reached only 0.53 and 0.48. Collectively, these findings suggest that our model demonstrates significant efficacy in identifying ROP risk.

Initiating treatment based solely on the model's flag would be premature and would risk significant overtreatment alongside subsequent iatrogenic visual impairment. Rather, it should trigger a state of heightened ophthalmic surveillance. For the high-risk cohort, a lower threshold for the initial screening and the implementation of intensified monitoring may be a more reasonable approach. This could include an earlier onset of the first screening, more frequent assessments via wide-field digital retinal imaging, and the integration of ultra-widefield OCT-A to quantify the area of the foveal avascular zone. Finally, the model may help individualize the decision to prematurely enter the 'intervention window.' Currently, indications for laser therapy or anti-VEGF treatment in ROP vary, and the decision often requires clinical judgment. In these borderline cases, our model could provide data-driven support to expedite the timing of intervention, thereby mitigating visual loss. In the common scenario of a preterm infant exhibiting ambiguous retinal imaging abnormalities, the model provides objective data that can reinforce a decision to either continue observation or initiate immediate therapy. The model can serve as an adjunctive decision-making tool during clinical diagnosis.

Future efforts will focus on assembling extensive datasets to conduct prospective validation studies, while making the model publicly accessible to enable broader external validation and seamless integration into routine clinical practice. Plans are underway to host the model and its documentation on an online mini-program, enabling healthcare providers and researchers worldwide to access and utilize it. By fostering transparency and collaboration, this study aims to refine the model further through external validation and user feedback, ultimately enhancing its utility in diverse clinical settings.

4 Methods

4.1 Data source and study design

This study employed a retrospective cohort design, enrolling infants born at the Neonatal Intensive Care Unit (NICU) of Xiangya Hospital, Central South University, between October 1, 2022, and June 30, 2025, who subsequently underwent screening for ROP. All included infants were registered in the Xiangya Hospital

electronic medical record system. Additionally, data from two other medical centers in Hunan Province—The First People’s Hospital of Changde City and The Central Hospital of Shaoyang—were included for external validation. For each infant, data collection encompassed essential clinical variables—including GA, BW, gender, and date of birth—as well as all arterial blood gas analysis results obtained within the first 42 days of life and detailed ROP screening outcomes. ROP examinations were performed by ophthalmologists with specialized training in neonatal retinal assessment. Data abstraction was carried out by trained personnel using standardized case report forms, and all entries were input into the database following a uniform protocol. All patients included in this study received care in accordance with national clinical guidelines and established institutional pathways. To protect patient privacy and ensure data confidentiality, all collected information was fully de-identified prior to analysis.

4.2 Data preprocessing

The initial study population comprised a development cohort of 1,000 infants (partitioned into training and internal testing sets) and two independent external validation cohorts, consisting of 80 preterm infants in Validation Set 1 and 40 preterm infants in Validation Set 2. Inclusion criteria required infants to undergo screening for ROP at hospital, with their ROP status definitively confirmed during subsequent follow-up. Eligible infants were required to have comprehensive medical records and sufficient data across all key clinical variables. Infants were excluded if they presented with ocular pathologies that interfered with screening—such as congenital cataracts, glaucoma, intraocular tumors, extensive retinal hemorrhage, familial exudative vitreoretinopathy, or congenital corneal opacity—or if ROP status remained unknown due to hospital discharge or transfer prior to a definitive diagnosis. ROP was defined as the primary endpoint, encompassing the occurrence of any stage of ROP at any time point throughout the screening period. In cases of asymmetrical bilateral involvement, the eye with the more severe disease progression was utilized for the final statistical analysis. The diagnostic criteria were based on the International Classification of Retinopathy of Prematurity, Third Edition.^[40]

Following the application of the aforementioned criteria, the final study population consisted of 584 infants (1,168 eyes) for internal model development and testing, with 74 and 35 infants retained as External Validation Cohorts 1 and 2, respectively. Prior to model construction, systematic data cleaning and preprocessing were performed to enhance data quality and reduce feature redundancy. Based on clinical relevance and data integrity, blood gas parameters with high missing rates or low correlation with the study objectives—including $t\text{CO}_2$, $F(\text{HbF})$, $\text{PaO}_2/\text{FiO}_2$, $\text{P}(\text{A-a})\text{O}_2$, HCO_3std , $\text{P}_{50}(\text{t})$, $\text{PO}_2(\text{t})$, $\text{tO}_2(\text{a})$, RI , SBE , $\text{Ca}^{++}(7.4)$, $\text{PCO}_2(\text{t})$, $\text{pH}(\text{t})$, ABE , and $\text{PO}_2(\text{a/a})$ —were excluded to minimize noise interference. Redundant variables reflecting identical physiological meanings were merged; during this process, non-missing values from the original variables were prioritized to maximize information retention. For inspired oxygen concentration, missing values were imputed with the fraction of inspired oxygen in room air (21%) to ensure consistency. Regarding missing data management, numerical variables underwent type conversion, and missing values in continuous variables were addressed using median imputation to mitigate the impact of outliers on data distribution. Categorical variables, such as sex, were processed using numerical encoding. Furthermore, to prevent data leakage, stratified sampling was performed using patient IDs as unique identifiers. The cohort was partitioned into a training set (80%) and a testing set (20%), ensuring that all longitudinal blood gas records from a single patient remained within the same subset. This approach prevents information overlap between the training and testing phases, thereby ensuring the reliability of the model evaluation. Following these procedures, the final dataset was established

for subsequent model training and validation. All preprocessing was performed on anonymized data in compliance with national data protection regulations.

4.3 Machine learning workflow

Ten machine learning models (Logistic Regression, SVM, KNN, Decision Tree, Random Forest, Naive Bayes, MLP, LightGBM, CatBoost, and XGBoost) were evaluated and compared. Hyperparameters were optimized using the Optuna framework with AUROC as the objective function. XGBoost achieved the highest AUROC with stable sensitivity and specificity on the internal testing set, and was selected as the final model for external validation using independent datasets from two additional medical centers. Of 1,120 initially enrolled patients from three medical centers in Hunan Province, China, those with ocular diseases affecting screening, incomplete data within 42 days postpartum, or unknown ROP outcomes were excluded. Following preprocessing, 693 patients underwent feature selection and model development. The XGBoost model was trained on 584 patients and validated on 109 patients. The overall study workflow is illustrated in Fig. 4.

Abbreviations

ROP Retinopathy of Prematurity

GA Gestational Age

BW Birth Weight

PMA Postmenstrual Age

ABGA Arterial Blood Gas Analysis

PCO₂ Partial Pressure of Carbon Dioxide

PaO₂ Alveolar Partial Pressure of Oxygen

FiO₂ Fraction of Inspired Oxygen

HCO₃⁻ Bicarbonate

BE_{ecf} Base Excess Extracellular Fluid

BE(B) Base Excess Blood

AG Anion Gap

Na⁺ Sodium

K⁺ Potassium

Cl⁻ Chloride

Lac Lactate

Ca²⁺ ☐ Ionized Calcium ☐

ctHb ☐ Total Hemoglobin ☐

Hct ☐ Hematocrit ☐

fO₂Hb ☐ Fractional Oxyhemoglobin ☐

fCOHb ☐ Fractional Carboxyhemoglobin ☐

fMetHb ☐ Fractional Methemoglobin ☐

fHHb ☐ Fractional Deoxyhemoglobin ☐

sO₂ ☐ Oxygen Saturation ☐

tBil ☐ Total Bilirubin ☐

tCO₂ ☐ Total Carbon Dioxide ☐

F(HbF) ☐ Fetal Hemoglobin Fraction ☐

PaO₂/FiO₂ ☐ Oxygenation Index ☐

P(A-a)O₂ ☐ Alveolar-arterial Oxygen Difference ☐

HCO₃std ☐ Standard Bicarbonate ☐

P₅₀(t) ☐ Half-saturation Partial Pressure of Oxygen ☐

PO₂(t) ☐ Tissue Partial Pressure of Oxygen ☐

tO₂(a) ☐ Total Oxygen Content Arterial ☐

RI ☐ Respiratory Index ☐

SBE ☐ Standard Base Excess ☐

Ca⁺⁺(7.4) ☐ Ionized Calcium at pH 7.4 ☐

PCO₂(t) ☐ Tissue Partial Pressure of CO₂ ☐

pH(t) ☐ Tissue pH ☐

ABE ☐ Actual Base Excess ☐

PO₂(a/a) ☐ Arterial/Alveolar Oxygen Ratio ☐

PPV ☐ Positive Predictive Value ☐

NPV ☐ Negative Predictive Value ☐

AUC ¶Area Under the Curve¶

AUROC ¶Area Under the Receiver Operating Characteristic¶

F1 ¶F1 Score¶

ACC ¶Accuracy¶

AI ¶Artificial Intelligence¶

ML ¶Machine Learning¶

SHAP ¶SHapley Additive exPlanations¶

KNN ¶K-Nearest Neighbors¶

MLP ¶Multi-layer Perceptron¶

SVM ¶Support Vector Machine¶

LightGBM ¶Light Gradient Boosting Machine¶

XGBoost ¶eXtreme Gradient Boosting¶

CatBoost ¶Categorical Boosting¶

NICU ¶Neonatal Intensive Care Unit¶

IGF-1 ¶Insulin-like Growth Factor-1¶

IGFBP-3 ¶Insulin-like Growth Factor Binding Protein-3¶

OCT-A ¶Optical Coherence Tomography Angiography¶

VEGF ¶Vascular Endothelial Growth Factor¶

2,3-DPG ¶2,3-Bisphosphoglycerate.

Declarations

Ethics approval and consent to participate

This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. Guardians of the infant participants provided informed consent for their data to be recorded and used for scientific purposes. The study protocol was reviewed and approved by the Ethics Committee of Xiangya Hospital, Central South University (Reference Number: 2026030496) and registered with the Chinese Clinical Trial Registry (Reference Number: ChiCTR2500115438). As this was a retrospective study using anonymized data, it was conducted in full compliance with ethical regulations.

Consent for publication

Not Applicable.

Availability of data and materials

The data supporting this study were collected from three medical centers in China: the Neonatal Intensive Care Unit at Xiangya Hospital of Central South University, as well as The Central Hospital of Shaoyang and The first people's hospital of Changde city. The dataset is not publicly available owing to patient confidentiality and data protection requirements.

Competing Interests

The authors declare no competing interests.

Funding

The present study was supported by National Key Research and Development Program of China (No. 2024YFA1108700 and 2024YFA1108704)

Author contributions

Wei Zheng, Chaoqin Yu and Qianyu Hao contributed to study conceptualization and data analysis. Qianyu Hao, Chaoqin Yu and Junlong Tang provided methodological support and statistical analysis. Chaoqin Yu and Moyan Liu, on behalf of Xiangya Hospital, contributed to data collection. Yuanyuan Chen, on behalf of The Central Hospital of Shaoyang, contributed to data collection. Xiao Chen, on behalf of The first people's hospital of Changde city, contributed to data collection. Wei Zheng and Junfeng Mao supervised the project. Huizhuo Xu and Xiaobo Xia contributed to retinopathy of prematurity image interpretation and diagnostic validation. Chaoqin Yu wrote the initial manuscript draft, Wei Zheng and Qianyu Hao were responsible for manuscript revision. All authors critically reviewed and approved the final manuscript.

Acknowledgements

The authors thank the editor and reviewers for their constructive comments, which have substantially improved the quality of this manuscript.

Code availability

The custom code used for all stages of data analysis in this study, including data preprocessing, feature selection, model training, hyperparameter tuning, and generation of figures, is available from the corresponding author upon reasonable request.

Patient and Public Involvement

Patients or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

References

1. Hellström A, Smith LE, Damman O. Retinopathy of prematurity. *Lancet*. 2013;382(9902):1445–57. 10.1016/S0140-6736(13)60178-6.
2. Epidemiology of blindness in children. *Arch Dis Child*. 2017;102(10):995. 10.1136/archdischild-2016-310532corr1.
3. Wang S, Liu J, Zhang X, et al. Global, regional and national burden of retinopathy of prematurity among childhood and adolescent: a spatiotemporal analysis based on the Global Burden of Disease Study 2019. *BMJ Paediatr Open*. 2024;8(1):e002267. 10.1136/bmjpo-2023-002267. Published 2024 Jan 6.
4. Sabri K, Ells AL, Lee EY, Dutta S, Vinekar A. Retinopathy of Prematurity: A Global Perspective and Recent Developments. *Pediatrics*. 2022;150(3):e2021053924. 10.1542/peds.2021-053924.
5. Early Treatment for Retinopathy of Prematurity Cooperative Group, Good WV, Hardy RJ, et al. Final visual acuity results in the early treatment for retinopathy of prematurity study. *Arch Ophthalmol*. 2010;128(6):663–71. 10.1001/archophthalmol.2010.72.
6. Lin JY, Kang EY, Banker AS et al. Comparison of RetCam and Smartphone-Based Photography for Retinopathy of Prematurity Screening. *Diagnostics (Basel)*. 2022;12(4):945. Published 2022 Apr 10. 10.3390/diagnostics12040945
7. Kiran Yenice E, Kara C, Erdaş ÇB. Automated detection of type 1 ROP, type 2 ROP and A-ROP based on deep learning. *Eye (Lond)*. 2024;38(13):2644–8. 10.1038/s41433-024-03184-0.
8. Mao J, Luo Y, Liu L, et al. Automated diagnosis and quantitative analysis of plus disease in retinopathy of prematurity based on deep convolutional neural networks. *Acta Ophthalmol*. 2020;98(3):e339–45. 10.1111/aos.14264.
9. Early Treatment For Retinopathy Of Prematurity Cooperative Group. Revised indications for the treatment of retinopathy of prematurity: results of the early treatment for retinopathy of prematurity randomized trial. *Arch Ophthalmol*. 2003;121(12):1684–94. 10.1001/archopht.121.12.1684.
10. Fierston WM, American Academy of Pediatrics Section on Ophthalmology; American Academy of Ophthalmology; American Association for Pediatric Ophthalmology and Strabismus; American Association of Certified Orthoptists. Screening Examination of Premature Infants for Retinopathy of Prematurity. *Pediatrics*. 2018;142(6):e20183061. *Pediatrics*. 2019;143(3):e20183810. 10.1542/peds.2018-3810
11. Palmer EA, Flynn JT, Hardy RJ, et al. Incidence and Early Course of Retinopathy of Prematurity. *Ophthalmology*. 2020;127(4S):S84–96. 10.1016/j.ophtha.2020.01.034.
12. Chiang MF, Arons RR, Flynn JT, Starren JB. Incidence of retinopathy of prematurity from 1996 to 2000: analysis of a comprehensive New York state patient database. *Ophthalmology*. 2004;111(7):1317–25. 10.1016/j.ophtha.2003.10.030.
13. Fielder AR, Haines L, Scrivener R, Wilkinson AR, Pollock JI. Royal Colleges of Ophthalmologists and Paediatrics and Child Health, and the British Association of Perinatal Medicine. Retinopathy of prematurity in the UK II: audit of national guidelines for screening and treatment. *Eye (Lond)*. 2002;16(3):285–91. 10.1038/sj.eye.6700131.
14. Hutchinson AK, Melia M, Yang MB, VanderVeen DK, Wilson LB, Lambert SR. Clinical Models and Algorithms for the Prediction of Retinopathy of Prematurity: A Report by the American Academy of

- Ophthalmology. Ophthalmology. 2016;123(4):804–16. 10.1016/j.ophtha.2015.11.003.
15. Lu LPL, Yip WWK, Lok JYC, et al. Prediction model to predict type 1 retinopathy of prematurity using gestational age and birth weight (PW-ROP). *Br J Ophthalmol*. 2023;107(7):1007–11. 10.1136/bjophthalmol-2021-320670.
 16. Löfqvist C, Hansen-Pupp I, Andersson E, et al. Validation of a new retinopathy of prematurity screening method monitoring longitudinal postnatal weight and insulinlike growth factor I. *Arch Ophthalmol*. 2009;127(5):622–7. 10.1001/archophthalmol.2009.69.
 17. Eckert GU, Fortes Filho JB, Maia M, Procianoy RS. A predictive score for retinopathy of prematurity in very low birth weight preterm infants. *Eye (Lond)*. 2012;26(3):400–6. 10.1038/eye.2011.334.
 18. Binenbaum G, Ying GS, Quinn GE, et al. The CHOP postnatal weight gain, birth weight, and gestational age retinopathy of prematurity risk model. *Arch Ophthalmol*. 2012;130(12):1560–5. 10.1001/archophthalmol.2012.2524.
 19. Binenbaum G, Bell EF, Donohue P, et al. Development of Modified Screening Criteria for Retinopathy of Prematurity: Primary Results From the Postnatal Growth and Retinopathy of Prematurity Study. *JAMA Ophthalmol*. 2018;136(9):1034–40. 10.1001/jamaophthalmol.2018.2753.
 20. Binenbaum G, Tomlinson LA, de Alba Campomanes AG, et al. Validation of the Postnatal Growth and Retinopathy of Prematurity Screening Criteria. *JAMA Ophthalmol*. 2020;138(1):31–7. 10.1001/jamaophthalmol.2019.4517.
 21. Pivodic A, Hård AL, Löfqvist C, et al. Individual Risk Prediction for Sight-Threatening Retinopathy of Prematurity Using Birth Characteristics. *JAMA Ophthalmol*. 2020;138(1):21–9. 10.1001/jamaophthalmol.2019.4502.
 22. Blencowe H, Lawn JE, Vazquez T, Fielder A, Gilbert C. Preterm-associated visual impairment and estimates of retinopathy of prematurity at regional and global levels for 2010. *Pediatr Res*. 2013;74(1):35–49. 10.1038/pr.2013.205.
 23. Fortes Filho JB, Bonomo PP, Maia M, Procianoy RS. Weight gain measured at 6 weeks after birth as a predictor for severe retinopathy of prematurity: study with 317 very low birth weight preterm babies. *Graefes Arch Clin Exp Ophthalmol*. 2009;247(6):831–6. 10.1007/s00417-008-1012-3.
 24. Kim J, Jin JY, Kim SS. Postnatal weight gain in the first two weeks as a predicting factor of severe retinopathy of prematurity requiring treatment. *Korean J Pediatr*. 2015;58(2):52–9. 10.3345/kjp.2015.58.2.52.
 25. Chen M, Citil A, McCabe F, et al. Infection, oxygen, and immaturity: interacting risk factors for retinopathy of prematurity. *Neonatology*. 2011;99(2):125–32. 10.1159/000312821.
 26. Shi W, Zhu L, He X, Wang S, Wang C. Combined indicator assists in early recognition of retinopathy of prematurity. *Sci Rep*. 2025;15(1):8048. 10.1038/s41598-025-92205-3. Published 2025 Mar 7.
 27. Ashton N. The pathogenesis of retrolental fibroplasia. *Ophthalmology*. 1979;86(10):1695–9. 10.1016/s0161-6420(79)35330-1.
 28. Biglan AW, Brown DR, Reynolds JD, Milley JR. Risk factors associated with retrolental fibroplasia. *Ophthalmology*. 1984;91(12):1504–11. 10.1016/s0161-6420(84)34104-5.
 29. Di Fiore JM, Bloom JN, Orge F, et al. A higher incidence of intermittent hypoxemic episodes is associated with severe retinopathy of prematurity. *J Pediatr*. 2010;157(1):69–73. 10.1016/j.jpeds.2010.01.046.

30. Saugstad OD. Oxidative stress in the newborn—a 30-year perspective. *Biol Neonate*. 2005;88(3):228–36. 10.1159/000087586.
31. Askie LM, Darlow BA, Finer N, et al. Association Between Oxygen Saturation Targeting and Death or Disability in Extremely Preterm Infants in the Neonatal Oxygenation Prospective Meta-analysis Collaboration. *JAMA*. 2018;319(21):2190–201. 10.1001/jama.2018.5725.
32. York JR, Landers S, Kirby RS, Arbogast PG, Penn JS. Arterial oxygen fluctuation and retinopathy of prematurity in very-low-birth-weight infants. *J Perinatol*. 2004;24(2):82–7. 10.1038/sj.jp.7211040.
33. Gantz MG, Carlo WA, Finer NN, et al. Achieved oxygen saturations and retinopathy of prematurity in extreme preterms. *Arch Dis Child Fetal Neonatal Ed*. 2020;105(2):138–44. 10.1136/archdischild-2018-316464.
34. Kubota H, Fukushima Y, Kawasaki R, et al. Continuous oxygen saturation and risk of retinopathy of prematurity in a Japanese cohort. *Br J Ophthalmol*. 2024;108(9):1275–80. 10.1136/bjo-2023-324225. Published 2024 Aug 22.
35. Zhao C, Sun Z, Chen H, Li K, Sun H. The impact of blood lactic acid levels on retinopathy of prematurity morbidity. *BMC Pediatr*. 2024;24(1):152. 10.1186/s12887-024-04571-y. Published 2024 Feb 29.
36. Ying GS, Quinn GE, Wade KC, et al. Predictors for the development of referral-warranted retinopathy of prematurity in the telemedicine approaches to evaluating acute-phase retinopathy of prematurity (e-ROP) study. *JAMA Ophthalmol*. 2015;133(3):304–11. 10.1001/jamaophthalmol.2014.5185.
37. Trzcionkowska K, Groenendaal F, Andriessen P, et al. Risk Factors for Retinopathy of Prematurity in the Netherlands: A Comparison of Two Cohorts. *Neonatology*. 2021;118(4):462–9. 10.1159/000517247.
38. Bunn HF, Poyton RO. Oxygen sensing and molecular adaptation to hypoxia. *Physiol Rev*. 1996;76(3):839–85. 10.1152/physrev.1996.76.3.839.
39. Piantadosi CA. Carbon monoxide, reactive oxygen signaling, and oxidative stress. *Free Radic Biol Med*. 2008;45(5):562–9. 10.1016/j.freeradbiomed.2008.05.013.
40. Chiang MF, Quinn GE, Fielder AR, et al. International Classification of Retinopathy of Prematurity, Third Edition. *Ophthalmology*. 2021;128(10):e51–68. 10.1016/j.ophtha.2021.05.031.

Figures

Correlation Matrix with ROP

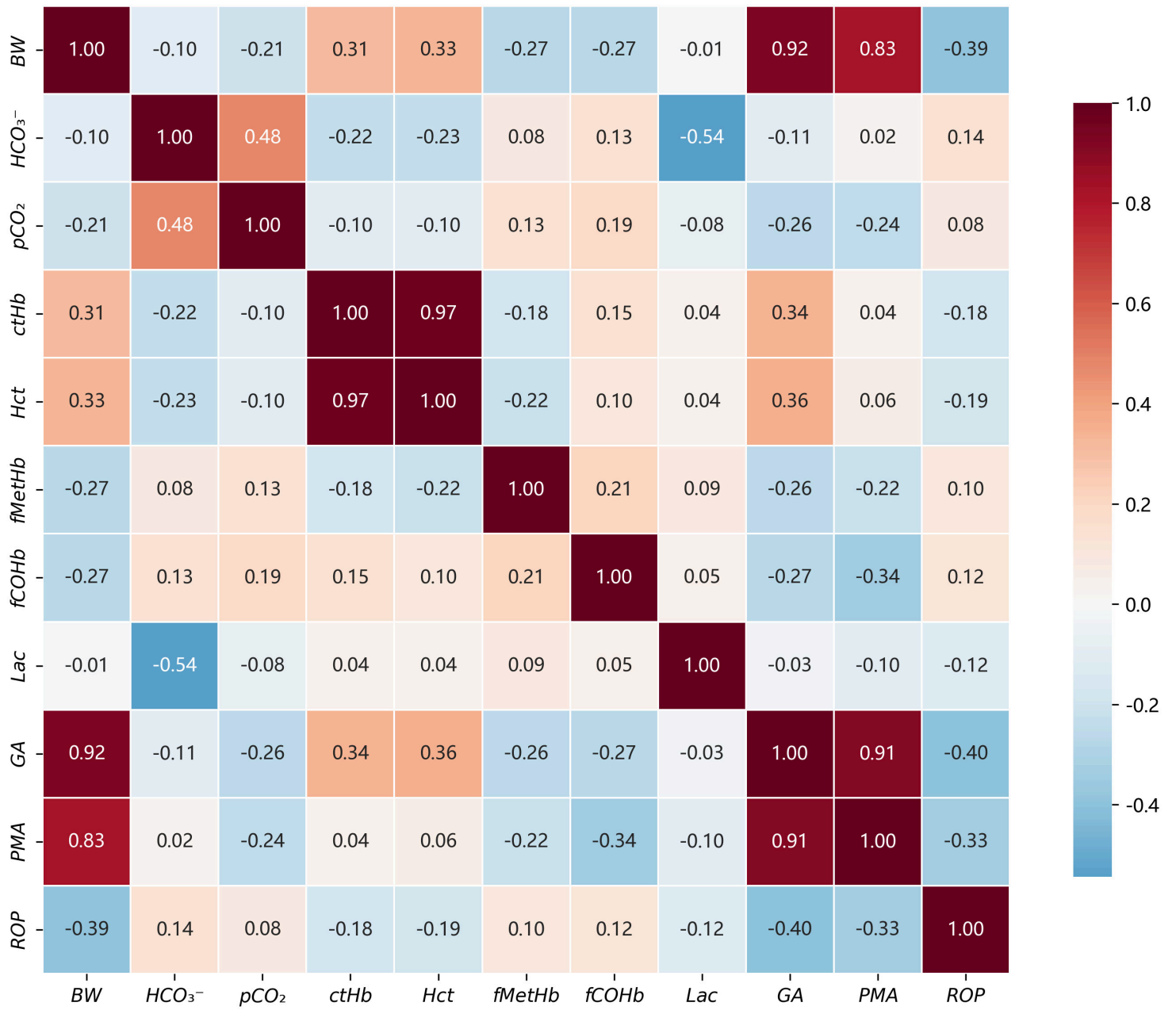


Figure 1

Legend not included with this version.

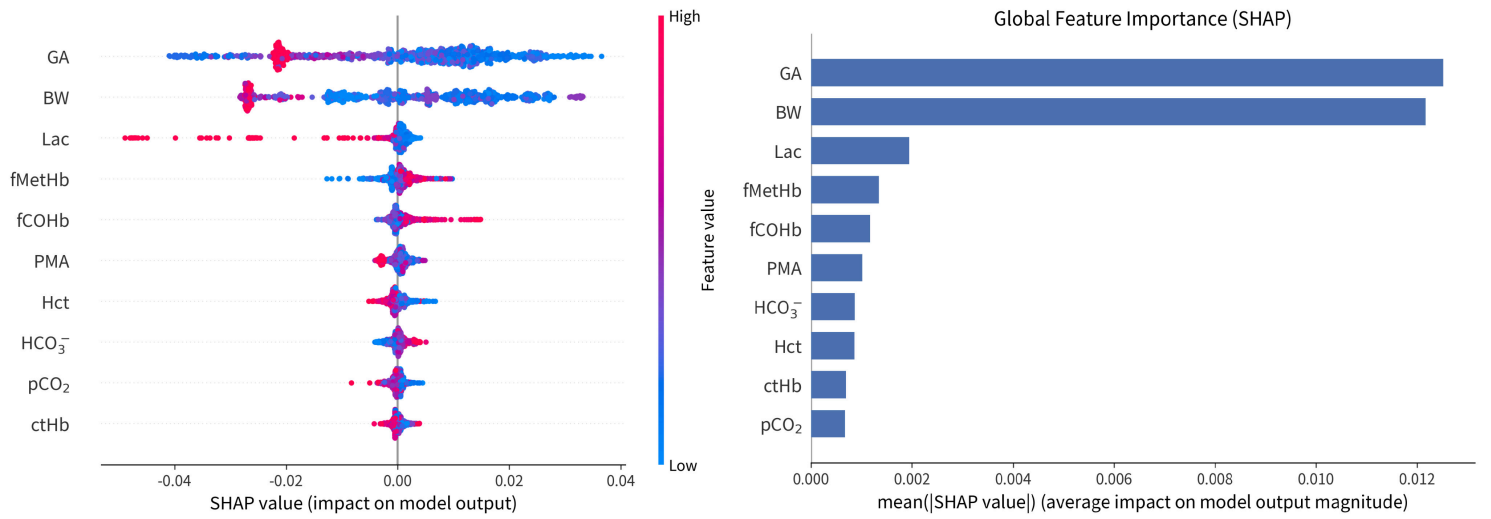


Figure 2

Legend not included with this version.

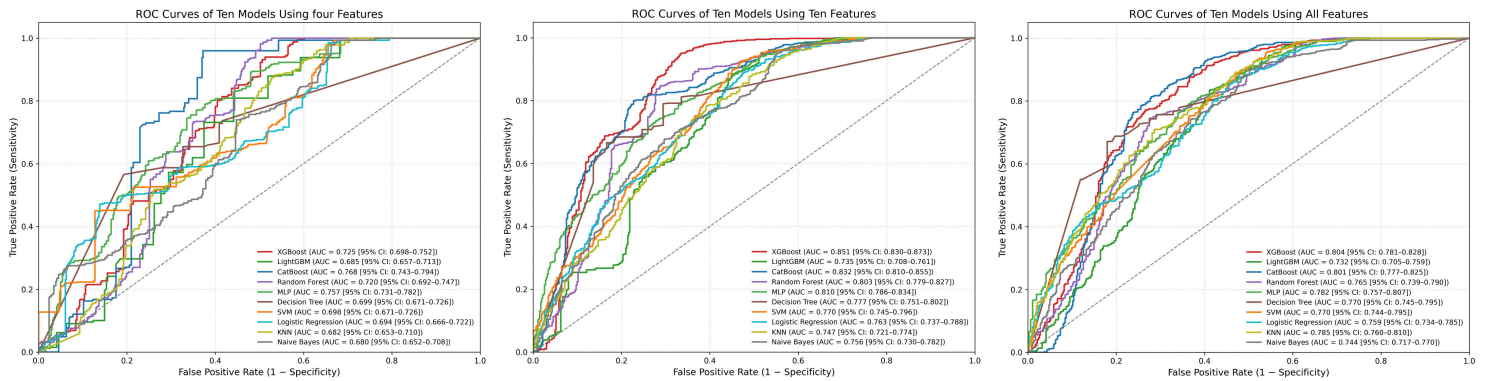


Figure 3

Legend not included with this version.

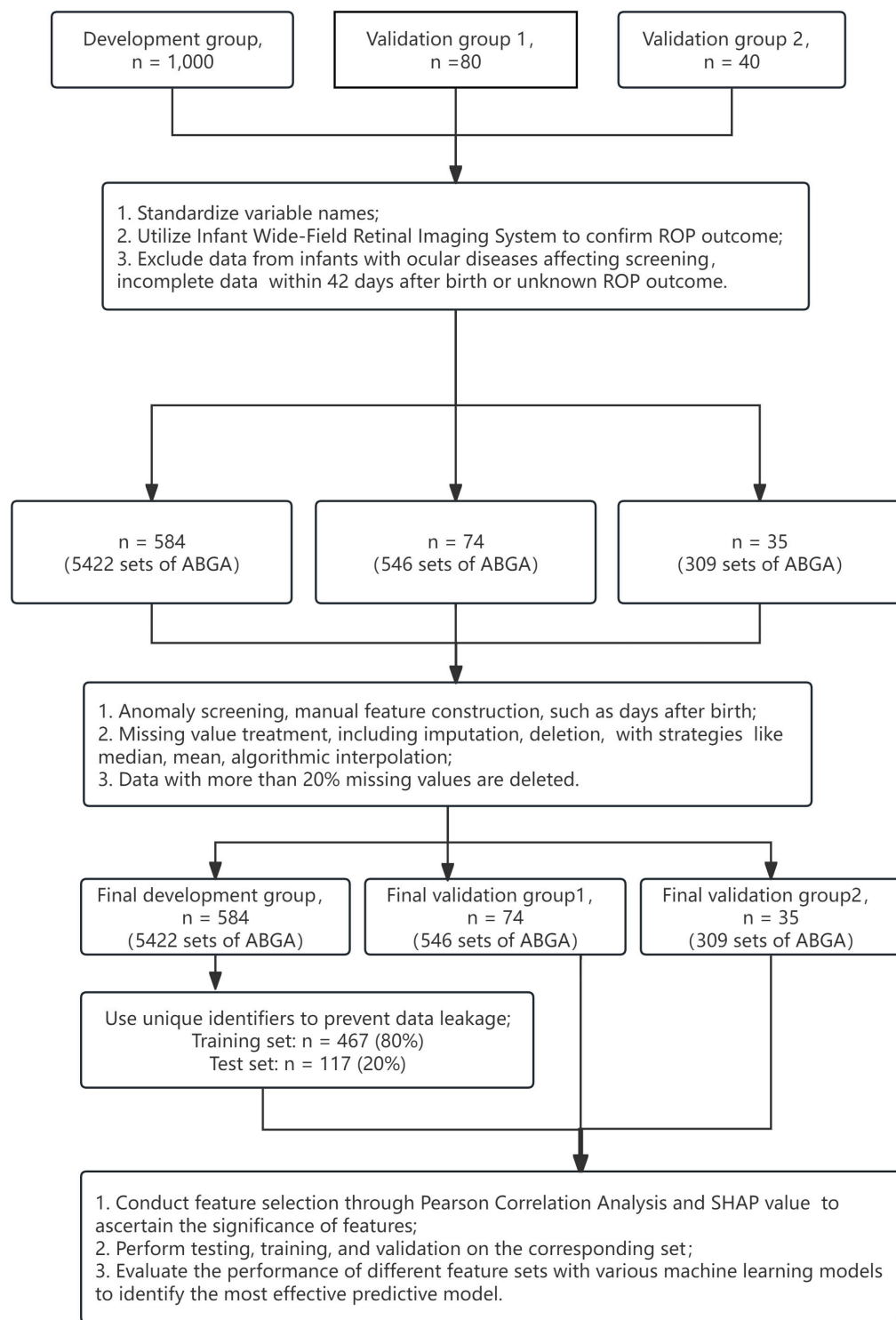


Figure 4

Legend not included with this version.

Supplementary Files

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

- [TRIPODCLUSTERChecklist.pdf](#)