International Journal of Women's Health

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Development and Validation of a Prenatal Prediction Model for Neonatal Hyperbilirubinemia Based on Maternal Factors: Revealing Heterogeneous Risk Profiles Across Gestational Age Subgroups

Jiawen Chen, Jin Wang, Jiandong Chen, Liangzhao Wu, Donglian Xiong

Neonatal Intensive Care Unit, Longyan First Affiliated Hospital of Fujian Medical University, Longyan, 364000, Fujian, People's Republic of China

Correspondence: Donglian Xiong, Neonatal Intensive Care Unit, Longyan First Affiliated Hospital of Fujian Medical University, No. 105, Jiuyi North Road, Xinluo District, Longyan, Fujian, 364000, People's Republic of China, Email 18558908900@163.com

Background: Neonatal hyperbilirubinemia is common, and current risk assessment depends largely on postnatal monitoring. A prediction model using only antenatal maternal factors could enable earlier identification, especially if it accounts for differences across gestational age (GA) subgroups.

Methods: This single-center case-control study enrolled 1967 subjects. After screening for significant differences in baseline characteristics and excluding multicollinearity, independent predictors were identified via univariate and multivariate logistic regression. Several machine learning models were developed and evaluated based on their area under the curve (AUC), sensitivity, specificity, accuracy, etc. Interaction analysis was performed, and subgroup-specific models were built for early-term and full-term subgroups. The SHAP analysis was employed to rank feature importance.

Results: Key independent antenatal predictors included GA, mode of delivery, maternal hypothyroidism, infection, mean corpuscular hemoglobin concentration (MCHC), alkaline phosphatase (AKP), albumin (ALB), and total bilirubin (TBIL). A significant interaction was found between GA and ALB (*P* for interaction=0.004). Logistic regression was selected as the optimal model (AUC: 0.661 [0.617–0.704]) for its generalizability and clinical utility. Subgroup analysis revealed distinct risk factor patterns; for instance, the SHAP analysis indicated that ALB and mode of delivery were the two most significant predictive factors in the full-term group, while mode of delivery and platelets were the key factors in the early-term group.

Conclusion: This study developed a predictive model for neonatal hyperbilirubinemia using an exploratory analysis of routine antenatal maternal factors. We demonstrated that GA significantly modified risk profiles, necessitating stratified assessment. However, the clinical relevance of certain non-traditional laboratory predictors requires further validation.

Keywords: machine learning, neonatal hyperbilirubinemia, maternal factors, predictive model

Introduction

Neonatal hyperbilirubinemia arises from a metabolic imbalance characterized by excessive bilirubin production that surpasses the hepatic and intestinal clearance capacity.[1,2] It affects 60% to 80% of newborns and is the leading cause of postnatal hospitalization in this population.[3] Without timely monitoring and intervention, elevated serum bilirubin levels may lead to acute bilirubin encephalopathy and kernicterus—permanent forms of neurological damage collectively known as chronic bilirubin encephalopathy—manifesting as cerebral palsy, hearing loss, ophthalmoplegia, dental enamel hypoplasia, and neurodevelopmental impairments.[4–7] Given the severe and irreversible nature of these outcomes, developing an early predictive model for neonatal hyperbilirubinemia is of critical clinical importance.

https://doi.org/10.2147/IJWH.S584032
Received: 5 December 2025
Accepted: 24 March 2026
Published: 27 April 2026

While postnatal bilirubin monitoring and transcutaneous bilirubinometry are widely adopted and effective tools,[8–10] they represent reactive rather than proactive strategies. A more preventive approach would involve identifying high-risk infants before birth to enable earlier clinical alertness and targeted interventions. Maternal health status significantly influences fetal development,[11] with prenatal factors such as nutritional status, blood type, specific metabolic comorbidities, and intrapartum complications—including premature rupture of membranes—known to impact neonatal outcomes.[12–16] Despite this, studies on predictive models using antenatal maternal clinical data remain limited,[17,18] and most existing models fail to account for potential heterogeneity in risk profiles across different gestational age (GA) subgroups. In addition, it is well established that premature infants are at high risk for hyperbilirubinemia; consequently, the risk of this condition in full-term infants may be underappreciated.[19] Furthermore, while certain maternal clinical conditions are recognized risk factors, the predictive value of broad, routine maternal laboratory parameters—such as hematological, liver, and renal indices—remains largely underexplored. Incorporating these widely accessible markers could help determine whether subtle physiological shifts act as early, indirect indicators of risk.

Therefore, this study aims to develop and validate a prenatal prediction model for neonatal hyperbilirubinemia. By utilizing a comprehensive set of antenatal maternal factors—including an exploratory analysis of routine laboratory indices—we specifically emphasize exploring the differential risk patterns associated with varying GA.

Methods

Data Source and Study Participants

This study employed a case-control design in which the grouping criterion was determined by the clinical outcomes of the newborns. Clinical data were retrospectively collected from parturients who delivered at the Obstetrics Department of Longyan First Affiliated Hospital of Fujian Medical University, from January 2020 to December 2024. The study was approved by the hospital ethics committee (approval number: LYREC2025-k176-01). As a retrospective study, all identifiable personal information was removed, and thus the ethics committee waived informed consent.

Hyperbilirubinemia Group

Initially, 3,188 mothers who delivered between January 2020 and December 2024 and had neonates diagnosed with neonatal hyperbilirubinemia were identified for retrospective analysis. A two-step exclusion process was applied. First, neonates with any recognized major secondary causes of hyperbilirubinemia were excluded. These causes included hemolytic disease (eg., ABO/Rh incompatibility, G6PD deficiency), prematurity or low birth weight, major perinatal complications (such as intracranial hemorrhage, respiratory distress syndrome, birth asphyxia), polycythemia, hepatobiliary system abnormalities, and major congenital malformations. This step yielded 1,400 neonates with non-secondary hyperbilirubinemia. Subsequently, mothers of these remaining neonates were excluded if they had specific maternal conditions: autoimmune diseases, hematological diseases, major antepartum or postpartum hemorrhage, or other severe diseases with organ dysfunction. After applying these criteria and excluding cases with missing critical data, 977 mothers were ultimately included in the hyperbilirubinemia group.

Healthy Control Group

The healthy control group was derived from mothers who delivered healthy neonates during the same period (initial pool: n=15,309). The same maternal exclusion criteria used in the second step for the case group were applied: presence of autoimmune diseases, hematological diseases, major antepartum or postpartum hemorrhage, other severe diseases with organ dysfunction, or missing critical data. Following screening, 990 eligible subjects were ultimately included in the healthy control group, based on an initial random selection at an approximate 1:1 ratio.

Diagnosis of Neonatal Hyperbilirubinemia

This study utilized the Bhutani nomogram[20] to assess neonatal hyperbilirubinemia risk. Hour-specific total serum bilirubin levels were measured before discharge and plotted on the percentile curve according to the infant's postnatal age. Newborns were categorized as low-risk (< 40th percentile), intermediate-risk (40th–95th percentile), or high-risk (≥ 95th percentile).

Those falling within the high-risk zone were diagnosed with hyperbilirubinemia. All blood analyses in this study were performed using the Cobas m511 fully automated analyzer (Roche, USA), with reagents supplied by the manufacturer.

Clinical Data of the Parturients

Clinical and laboratory data were extracted from the electronic medical record system for all parturients, including maternal age, body mass index (BMI), blood type, GA, delivery mode, multiparity, gestational diabetes mellitus (GDM), hypertensive disorders of pregnancy (HDP), hypothyroidism, infection, placental abruption, premature rupture of membranes (PROM), and pre-delivery hematological and biochemical parameters. The latter included white blood cell (WBC), lymphocytes, neutrophils, platelets (PLT), red blood cell (RBC), hemoglobin (HGB), hematocrit (HCT), mean corpuscular hemoglobin concentration (MCHC), mean corpuscular volume (MCV), albumin (ALB), prealbumin (PALB), alkaline phosphatase (AKP), total bilirubin (TBIL), gamma-glutamyl transferase (GGT), total bile acids (TBA), creatinine, and blood urea nitrogen (BUN). Mode of delivery was categorized as cesarean section or vaginal delivery. Multiparity was defined as three or more pregnancies with delivery events. GA was classified as early term (37–-38 weeks), full term (39–40 weeks), and late term (≥41 weeks). Diagnoses of HDP and GDM were established according to the clinical practice guidelines issued by the Chinese Medical Doctor Association.[21,22]

Statistical Analysis

Data are summarized as follows: continuous variables are presented as median (P_{25}-P_{75}), and categorical variables are expressed as numbers and percentages. Group comparisons were performed using the Mann–Whitney *U*-test for continuous variables and the Chi-square test for categorical variables. All statistical analyses were conducted using RStudio (with R version 4.2.1) and SPSS version 26.0. A two-sided P-value less than 0.05 was considered statistically significant.

First, we performed predictor screening through a two-step regression approach: univariate logistic regression was initially conducted to identify potentially significant variables, which were then entered into multivariate logistic regression analysis; only those variables that remained statistically significant in the multivariate model were retained as final predictors. All data were randomly split into a training set (70%) and a validation set (30%). The model was subsequently trained and evaluated 10 times using resampling techniques to enhance the robustness and reliability of performance estimates. We evaluated five machine learning models—Extreme Gradient Boosting (XGBoost), Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), and k-Nearest Neighbors (KNN)—to determine the optimal model. Model performance was assessed using multiple metrics, including area under the curve (AUC), sensitivity, specificity, accuracy, negative predictive value (NPV), positive predictive value (PPV), and F1 score. Visual diagnostic tools—including receiver operating characteristic (ROC) curves, calibration plots, decision curve analysis (DCA), and precision-recall (PR) curves—were generated to comprehensively evaluate predictive capability and clinical utility. SHAP (SHapley Additive exPlanations), a game theory-based method for interpreting machine learning outputs,[23] was applied after model selection to quantify the contribution of each feature to individual predictions, thereby improving model interpretability and transparency. Subsequently, interaction analysis was conducted to assess whether GA modified the effects of identified predictors. Finally, subgroup analyses were carried out stratifying by GA categories (full term and early term), with re-evaluation of predictors, model performance, and feature importance using SHAP to visualize contributions within each subgroup.

Results

Baseline Characteristics

The baseline maternal characteristics of the overall cohort (n=1967) and the stratification by neonatal hyperbilirubinemia status are presented in Table 1. Significant differences in several key antenatal and intrapartum factors were observed between the normal (n=990) and hyperbilirubinemia (n=977) groups. The distribution of GA was markedly different ($P<0.001$), with the hyperbilirubinemia group comprising a lower proportion of full-term neonates (52.10% vs. 64.85%) and a higher proportion of early-term neonates (44.73% vs. 29.70%) compared to the normal group. Mode

Table 1 Baseline Characteristics

Variable	Overall (n = 1967)	Normal (n = 990)	Hyperbilirubinemia (n = 977)	*P*-value
Age, hours	30.000[27.000,33.000]	31.000[27.000,33.000]	30.000[27.000,33.000]	0.118
BMI, kg/m^2	26.500[24.700,28.600]	26.438[24.524,28.500]	26.600[24.900,28.700]	0.077
Blood type, n (%)				0.043
O	727(36.960)	398(40.202)	329(33.675)	
A	514(26.131)	252(25.455)	262(26.817)	
B	399(20.285)	190(19.192)	209(21.392)	
AB	111(5.643)	50(5.051)	61(6.244)	
Unknown	216(10.981)	100(10.101)	116(11.873)	
Multiparity, n (%)				0.001
No	1706(86.731)	834(84.242)	872(89.253)	
Yes	261(13.269)	156(15.758)	105(10.747)	
GA, n (%)				<0.001
Full term	1151(58.516)	642(64.848)	509(52.098)	
Early term	731(37.163)	294(29.697)	437(44.729)	
Late term	85(4.321)	54(5.455)	31(3.173)	
Mod of delivery, n (%)				<0.001
Cesarean section	847(43.060)	489(49.394)	358(36.643)	
Vaginal	1120(56.940)	501(50.606)	619(63.357)	
Placental abruption, n (%)				0.795
Yes	19(0.966)	9(0.909)	10(1.024)	
PROM, n (%)				0.001
Yes	462(23.488)	202(20.404)	260(26.612)	
HDP, n (%)				0.954
Yes	72(3.660)	36(3.636)	36(3.685)	
GDM, n (%)				0.826
Yes	481(24.453)	240(24.242)	241(24.667)	
Hypothyroidism, n (%)				0.010
Yes	102(5.186)	64(6.465)	38(3.889)	
Infection, n (%)				<0.001
Yes	276(14.032)	110(11.111)	166(16.991)	
WBC, $10^9/L$	8.900[7.530,10.430]	8.780[7.490,10.400]	8.990[7.580,10.490]	0.279
RBC, $10^{12}/L$	4.040[3.800,4.290]	4.030[3.800,4.280]	4.050[3.790,4.290]	0.559
HGB, g/L	122.000[113.000,129.000]	122.000[115.000,129.000]	121.000[112.000,129.000]	0.008
HCT, %	0.360[0.340,0.380]	0.360[0.340,0.380]	0.360[0.340,0.380]	0.339
MCV, fl	90.900[87.000,94.000]	91.100[87.400,94.000]	90.600[86.300,93.900]	0.097
MCHC, g/L	336.000[330.000,340.000]	336.000[330.000,341.000]	335.000[328.000,340.000]	0.001
AKP, IU/L	175.000[141.000,211.000]	181.000[148.000,215.000]	167.000[135.000,208.000]	<0.001
ALB, g/L	34.100[32.600,35.700]	33.900[32.400,35.400]	34.400[32.700,35.900]	<0.001
PALB, mg/dL	20.700[18.100,23.200]	20.700[18.100,23.100]	20.700[18.100,23.400]	0.809
Lymphocyte, $10^9/L$	1.560[1.280,1.860]	1.570[1.280,1.860]	1.550[1.290,1.860]	0.956

(*Continued*)

Table 1 (Continued).

Variable	Overall (n = 1967)	Normal (n = 990)	Hyperbilirubinemia (n = 977)	*P*-value
Neutrophils, 10^9/L	6.550[5.480,7.980]	6.530[5.470,7.900]	6.570[5.500,8.050]	0.494
PLT, 10^9/L	209.000[178.000,248.000]	207.000[176.000,242.000]	212.000[179.000,252.000]	0.035
TBIL, μmol/L	7.500[6.200,9.400]	7.300[6.000,8.900]	7.800[6.400,10.000]	<0.001
GGT, IU/L	10.000[8.000,14.000]	10.000[8.000,14.000]	10.000[8.000,14.000]	0.258
TBA, μmol/L	2.900[2.000,4.400]	3.000[2.100,4.300]	2.900[2.000,4.400]	0.524
Creatinine, μmol/L	45.000[40.000,51.000]	45.000[40.000,51.000]	45.000[40.000,51.000]	0.784
BUN, mmol/L	3.290[2.740,3.930]	3.370[2.810,3.960]	3.200[2.690,3.870]	0.002

Note: Median [P_{25},P_{75}] for continuous variables and counts (percentage) for categorical variables.
Abbreviations: BMI, body mass index; GA, gestational age; PROM, premature rupture of membrane; HDP, hypertensive disorders of pregnancy; GDM, gestational diabetes mellitus; WBC, white blood cell; RBC, red blood cell; HGB, hemoglobin; HCT, hematocrit; MCV, mean corpuscular volume; MCHC, mean corpuscular hemoglobin concentration; AKP, alkaline phosphatase; ALB, albumin; PALB, prealbumin; PLT, platelets; TBIL, total bilirubin; GGT, gamma-glutamyl transpeptidase; TBA, total bile acids; BUN, blood urea nitrogen.

of delivery also differed significantly (P<0.001), with vaginal delivery being more prevalent in the hyperbilirubinemia group (63.36% vs. 50.61%). Furthermore, the hyperbilirubinemia group had a higher incidence of PROM (26.61% vs. 20.40%, P=0.001) and a lower incidence of maternal hypothyroidism (3.89% vs. 6.47%, P=0.010) and multiparity (10.75% vs. 15.76%, P=0.001). Regarding maternal blood type, the proportion of type O was lower in the hyperbilirubinemia group (33.68% vs. 40.20%, P=0.043). Maternal infection was more common in mothers whose neonates developed hyperbilirubinemia (16.99% vs. 11.11%, P<0.001). Several maternal laboratory parameters showed significant differences. Notably, lower levels of HGB (P=0.008), MCHC (P=0.001), and AKP (P<0.001), along with higher ALB (P<0.001) and PLT (P=0.035), were observed in the hyperbilirubinemia group. In contrast, no significant differences were observed for other variables such as maternal age, BMI, HDP, GDM, placental abruption, WBC, RBC, HCT, MCV, PALB, lymphocyte count, neutrophil count, GGT, TBA, or creatinine.

Predictor Screening, Model Selection, and Model Explanation in the Overall Population

To identify predictors for constructing the prediction model, we first selected variables showing significant differences in baseline characteristics—after confirming the absence of multicollinearity (all variance inflation factors [VIF] < 5)—for univariate logistic regression analysis. Variables that achieved statistical significance in univariate analysis were then advanced to multivariate logistic regression (Table 2). This process identified MCHC, AKP, ALB, TBIL, GA, multiparity, mode of delivery, hypothyroidism, and infection as independent predictors of neonatal hyperbilirubinemia (all $P < 0.05$).

We developed multiple predictive models using XGBoost, LR, RF, SVM, and KNN algorithms, and systematically compared their performance to select the optimal model. In the training set, both XGBoost and RF achieved perfect AUC values of 1.000 (95% CI: 1.000–1.000), suggesting potential overfitting (Figure 1A). However, in the validation set, LR demonstrated superior generalizability with the highest AUC of 0.661 (95% CI: 0.617–0.704) (Figure 1B). The forest plot further confirmed LR's consistent performance, yielding an AUC of 0.661 with zero standard error (Figure 1C). Although XGBoost showed the best calibration (Figure 1D), LR exhibited higher average precision (AP = 0.628) in the validation set (Figure 1E). As detailed in Table 3, LR also achieved the highest NPV (0.627), accuracy (0.628), and F1 score (0.636) among all models in the validation cohort. DCA revealed that LR provided the greatest net clinical benefit across low to moderate decision thresholds (Figure 1F). Taken together, these results support the selection of LR as the most balanced and clinically useful model.

To improve model transparency and interpretability, we applied the SHAP algorithm to elucidate feature contributions to model predictions. As shown in Figure 2A and B, the mode of delivery and GA emerged as the two most influential features, followed by TBIL and ALB, highlighting their critical roles in the model's decision logic.

Table 2 Univariate and Multivariate Logistic Regression Analyses of Prenatal Maternal Factors for Predicting Neonatal Hyperbilirubinemia

Variables	Univariate		Multivariate	
	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
HGB	0.988 (0.980, 0.995)	0.001	0.992 (0.982, 1.002)	0.108
MCHC	0.981 (0.972, 0.991)	<0.001	0.984 (0.972, 0.996)	0.010
AKP	0.998 (0.996, 0.999)	0.003	0.998 (0.997, 1.000)	0.044
ALB	1.068 (1.031, 1.108)	<0.001	1.081 (1.038, 1.126)	<0.001
PLT	1.002 (1.000, 1.003)	0.035	1.001 (0.999, 1.003)	0.163
TBIL	1.080 (1.047, 1.114)	<0.001	1.089 (1.053, 1.126)	<0.001
BUN	0.876 (0.797, 0.963)	0.006	0.916 (0.829, 1.012)	0.086
Multiparity (Yes)	0.644 (0.494, 0.839)	0.001	0.553 (0.414, 0.736)	<0.001
Mod of delivery (Vaginal)	1.688 (1.409, 2.021)	<0.001	1.863 (1.519, 2.288)	<0.001
Hypothyroidism (Yes)	0.586 (0.388, 0.884)	0.011	0.591 (0.376, 0.917)	0.020
Infection (Yes)	1.637 (1.264, 2.122)	<0.001	1.719 (1.308, 2.267)	<0.001
PROM (Yes)	1.415 (1.147, 1.745)	0.001	1.189 (0.940, 1.505)	0.149
GA				
Full term	Ref		Ref	
Early term	1.875 (1.553, 2.263)	<0.001	2.065 (1.682, 2.539)	<0.001
Late term	0.724 (0.459, 1.143)	0.166	0.720 (0.444, 1.149)	0.174
Blood type				
O	Ref		Ref	
A	1.258 (1.003, 1.577)	0.047	1.224 (0.961, 1.561)	0.101
AB	1.476 (0.988, 2.205)	0.057	1.448 (0.944, 2.229)	0.091
B	1.331 (1.042, 1.700)	0.022	1.244 (0.957, 1.616)	0.103
Unknown	1.403 (1.035, 1.903)	0.029	1.169 (0.833, 1.642)	0.367

Abbreviations: OR, odds ratio; CI, confidence interval; HGB, hemoglobin; MCHC, mean corpuscular hemoglobin concentration; AKP, alkaline phosphatase; ALB, albumin; PLT, platelets; TBIL, total bilirubin; BUN, blood urea nitrogen; PROM, premature rupture of membrane; GA, gestational age.

Interaction Analysis

Notably, while multivariate logistic regression indicated that GA had the largest odds ratio—suggesting that early-term group are at significantly higher risk of hyperbilirubinemia compared to full-term group—the SHAP importance ranking identified mode of delivery as the top predictor. Nevertheless, GA remained the second most important feature, underscoring its substantial contribution to risk prediction. Given that GA is a well-established marker of neonatal maturity and a core determinant of bilirubin metabolism,[24] we hypothesize that it may act not only as a direct risk factor but also as an effect modifier that modulates the impact of other clinical variables. The subsequent interaction analysis revealed a significant interaction between GA and ALB, with an interaction odds ratio of 0.912 (95% CI: 0.856–0.971; $P = 0.004$), indicating that GA may modify the effect of ALB on the risk of neonatal hyperbilirubinemia (Table 4).

Predictor Screening, Model Selection, and Model Explanation in the Full-Term Group

The interaction analysis provided statistical evidence that GA modified the effect of at least one key predictor, ALB, suggesting potential heterogeneity in the underlying risk profile across GA groups. To account for this variability, we performed stratified analyses and developed separate prediction models within each GA subgroup. The late-term group was excluded from further analysis due to insufficient sample size, which would compromise the stability and reliability of the results.

In the full-term group, the same variable selection procedure was applied to identify predictive factors, and MCHC, AKP, ALB, TBIL, multiparity, mode of delivery, and infection were retained as significant predictors (Table 5, all $P < 0.05$). In the training set, both XGBoost and RF models exhibited overfitting, whereas in the validation set, LR achieved the highest AUC of 0.626 (95% CI: 0.568–0.685). The forest plot further confirmed superior performance of the LR model,

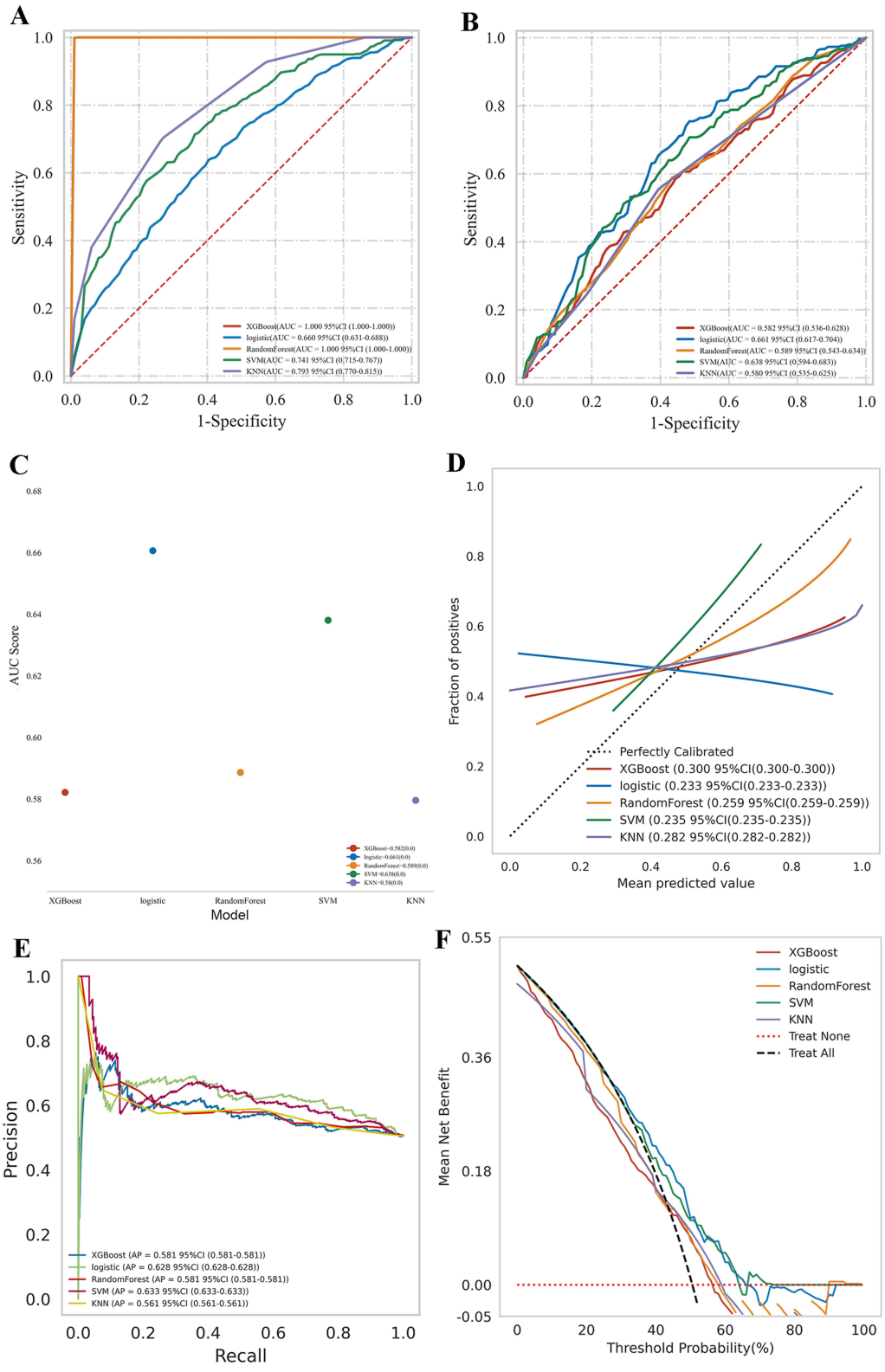

Figure 1 Comprehensive performance evaluation of machine learning models for predicting neonatal hyperbilirubinemia. (**A**) ROC curves for the training sets. (**B**) ROC curves for the validation sets. (**C**) Forest plot showing the AUC scores of each model. (**D**) Calibration curves for the validation set across all models. (**E**) PR curves for the validation set across all models. (**F**) DCA for the validation set across all models.

Abbreviations: XGBoost, Extreme Gradient Boosting; RF, Random Forest; SVM, Support Vector Machine; KNN, K-Nearest Neighbors; ROC, receiver operator characteristic curve; AUC, area under the curve; DCA, decision curve analysis; PR, precision-recall.

https://doi.org/10.2147/IJWH.S584032

Table 3 Predictive Performance of the ML Models on Training and Validation Sets

Dataset	Model	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	Accuracy	F1 score
Training	XGBoost	1.000 (1.000, 1.000)	0.996	0.999	0.999	0.996	0.997	0.997
	LR	0.660 (0.631, 0.688)	0.639	0.597	0.608	0.628	0.618	0.623
	RF	1.000 (1.000, 1.000)	0.999	0.991	0.991	0.999	0.995	0.995
	SVM	0.741 (0.715, 0.767)	0.577	0.778	0.717	0.653	0.678	0.639
	KNN	0.793 (0.770, 0.815)	0.702	0.732	0.719	0.715	0.717	0.710
Validation	XGBoost	0.582 (0.536, 0.628)	0.572	0.560	0.570	0.562	0.566	0.571
	LR	0.661 (0.617, 0.704)	0.643	0.612	0.628	0.627	0.628	0.636
	RF	0.589 (0.543, 0.634)	0.579	0.570	0.579	0.570	0.575	0.579
	SVM	0.638 (0.594, 0.683)	0.515	0.701	0.638	0.586	0.607	0.570
	KNN	0.580 (0.535, 0.625)	0.556	0.605	0.589	0.571	0.580	0.572

Abbreviations: XGBoost, Extreme Gradient Boosting; LR, logistic regression; RF, Random Forest; SVM, Support Vector Machine; KNN, K-Nearest Neighbors; ML, machine learning; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value.

with an AUC of 0.626 and a standard error of 0 (Figure 3). For calibration, XGBoost showed the closest agreement between predicted and observed outcomes in the validation set. Moreover, LR yielded the highest average precision (AP = 0.587), indicating strong discriminative ability for positive cases (Figure 3). As detailed in Table 6, LR also demonstrated the best performance across multiple validation metrics, including sensitivity (0.487), PPV (0.553), NPV (0.596), accuracy (0.578), and F1 score (0.518). DCA further revealed that LR provided the greatest net clinical benefit within the majority of threshold probabilities (Figure 3). Taken together, these results support the selection of LR as the optimal predictive model for this subgroup.

SHAP analysis indicates that ALB and mode of delivery are the two most influential features in the prediction model, with TBIL ranking third, followed by infection, MCHC, AKP, and multiparity (Figure 4).

Predictor Screening, Model Selection, and Model Explanation in the Early-Term Group

In the early-term group, mode of delivery, multiparity, TBIL, PLT, and BUN were identified as the final predictors (Table 7, all $P < 0.05$). As shown in Figure 5, both XGBoost and RF exhibited overfitting in the training set; in the validation set, LR achieved the highest AUC of 0.606 (95% CI: 0.528–0.684). The forest plot further confirmed superior performance of the LR model, with an AUC of 0.607 and a standard error of 0. For calibration, XGBoost demonstrated the closest agreement between predicted and observed outcomes in the validation set. Moreover, LR yielded the highest average precision (AP = 0.720), indicating strong discriminative ability for positive cases. DCA revealed that LR provided the greatest net clinical benefit across moderate to high decision thresholds. As detailed in Table 8, LR also showed the best performance in terms of specificity (0.662), PPV (0.690), NPV (0.414), and accuracy (0.526) on the validation set. Taken together, these results support the selection of LR as the optimal predictive model for this subgroup. SHAP analysis indicated that mode of delivery was the most influential feature, followed by PLT, TBIL, multiparity, and BUN (Figure 6).

Discussion

In this study, we developed and validated a predictive model for neonatal hyperbilirubinemia based on prenatal maternal factors. First, we identified a set of independent maternal clinical and laboratory characteristics significantly associated with the risk of neonatal hyperbilirubinemia. Second, we demonstrated that GA not only functioned as a strong independent predictor but also acted as an effect modifier, exhibiting a significant interaction with maternal ALB levels. This interaction suggests heterogeneous risk patterns across GA subgroups. Consequently, we constructed and validated subgroup-specific models in both the full-term and early-term populations. Across all analyses, LR consistently emerged as the most generalizable and clinically applicable algorithm. Finally, SHAP analysis revealed the relative importance of features in model decision-making: mode of delivery and GA were the top contributors in the overall population; ALB and mode of delivery ranked highest in the full-term group; and mode of delivery together with PLT were the most influential features in the early-term group.

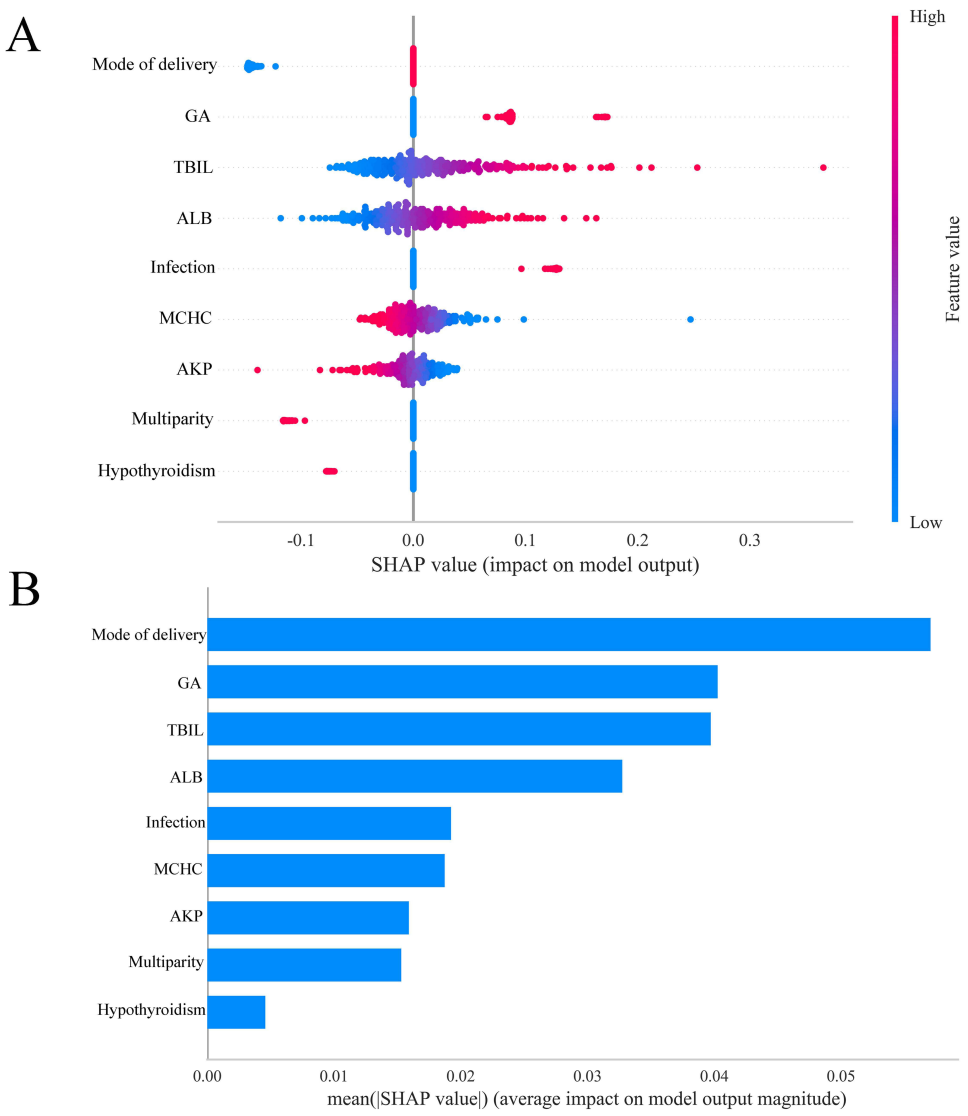

Figure 2 Visually interpret the best machine learning models using SHAP. (**A**) SHAP summary plot. (**B**) SHAP importance plot.
Abbreviations: MCHC, mean corpuscular hemoglobin concentration; AKP, alkaline phosphatase; ALB, albumin; TBIL, total bilirubin; GA, gestational age.

The predictors identified in this study are not only statistically significant but also deeply congruent with the pathophysiology of neonatal hyperbilirubinemia, while revealing potential novel risk patterns. Among the risk factors, the elevated level of maternal TBIL serves as the most direct biomarker for the disease, reinforcing the concept of

Table 4 Analysis of Interaction Effects Between GA and Other Risk Factors

Interaction Term	OR (95% CI)	*P*-value
GA * Mod of delivery	0.984 (0.717, 1.350)	0.919
GA * Infection	0.640 (0.408, 1.003)	0.051
GA * ALB	0.912 (0.856, 0.971)	0.004
GA * MCHC	1.000 (0.983, 1.017)	0.987
GA * AKP	0.999 (0.997, 1.002)	0.531
GA * TBIL	0.963 (0.913, 1.016)	0.167
GA * Multiparity	1.071 (0.660, 1.736)	0.782
GA * Hypothyroidism	0.549 (0.272, 1.107)	0.094

Abbreviations: OR, odds ratio; CI, confidence interval; MCHC, mean corpuscular hemoglobin concentration; AKP, alkaline phosphatase; ALB, albumin; TBIL, total bilirubin; GA, gestational age. The symbol "*" denotes the interaction term between the two variables.

https://doi.org/10.2147/IJWH.S584032

Table 5 Univariate and Multivariate Logistic Regression Analyses of Risk Factors for Neonatal Hyperbilirubinemia in the Full-Term Group

Variables	Univariate		Multivariate	
	OR (95% CI)	***P*-value**	**OR (95% CI)**	***P*-value**
MCHC	0.981 (0.969, 0.994)	0.003	0.977 (0.964, 0.990)	0.001
AKP	0.998 (0.996, 1.000)	0.035	0.998 (0.996, 1.000)	0.030
ALB	1.128 (1.072, 1.187)	<0.001	1.116 (1.058, 1.178)	<0.001
TBIL	1.103 (1.057, 1.150)	<0.001	1.109 (1.061, 1.162)	<0.001
Multiparity (Yes)	0.550 (0.364, 0.832)	0.005	0.517 (0.332, 0.791)	0.003
Mod of delivery (Vaginal)	1.754 (1.378, 2.232)	<0.001	1.922 (1.477, 2.507)	<0.001
Infection (Yes)	1.966 (1.414, 2.734)	<0.001	2.053 (1.459, 2.905)	<0.001
Blood type				
O	ref		ref	
A	1.385 (1.032, 1.861)	0.030	1.287 (0.942, 1.758)	0.113
AB	1.315 (0.752, 2.301)	0.337	1.238 (0.688, 2.216)	0.472
B	1.221 (0.880, 1.694)	0.231	1.154 (0.817, 1.628)	0.416
Unknown	1.472 (0.995, 2.178)	0.053	1.260 (0.822, 1.934)	0.289
Age	0.982 (0.957, 1.008)	0.182	-	-
BMI	1.015 (0.978, 1.053)	0.437	-	-
WBC	1.027 (0.985, 1.071)	0.215	-	-
RBC	0.990 (0.752, 1.302)	0.941	-	-
HGB	0.992 (0.983, 1.002)	0.111	-	-
HCT	1.054 (0.671, 1.656)	0.819	-	-
MCV	0.996 (0.979, 1.012)	0.609	-	-
PALB	0.996 (0.969, 1.025)	0.796	-	-
Lymphocyte	0.950 (0.740, 1.219)	0.687	-	-
Neutrophils	1.017 (0.964, 1.074)	0.529	-	-
PLT	1.000 (0.998, 1.002)	0.877	-	-
GGT	0.995 (0.981, 1.010)	0.547	-	-
TBA	0.993 (0.961, 1.027)	0.698	-	-
Creatinine	1.002 (0.989, 1.015)	0.770	-	-
BUN	0.937 (0.830, 1.059)	0.299	-	-
HDP (Yes)	1.040 (0.508, 2.130)	0.915	-	-
GDM (Yes)	1.125 (0.861, 1.469)	0.389	-	-
Hypothyroidism (Yes)	0.732 (0.431, 1.242)	0.247	-	-
Placental abruption (Yes)	1.264 (0.405, 3.944)	0.686	-	-
PROM (Yes)	1.249 (0.932, 1.673)	0.137	-	-

Abbreviations: OR, odds ratio; CI, confidence interval; BMI, body mass index; GA, gestational age; PROM, premature rupture of membrane; HDP, hypertensive disorders of pregnancy; GDM, gestational diabetes mellitus; WBC, white blood cell; RBC, red blood cell; HGB, hemoglobin; HCT, hematocrit; MCV, mean corpuscular volume; MCHC, mean corpuscular hemoglobin concentration; AKP, alkaline phosphatase; ALB, albumin; PALB, prealbumin; PLT, platelets; TBIL, total bilirubin; GGT, gamma-glutamyl transpeptidase; TBA, total bile acids; BUN, blood urea nitrogen.

metabolic continuity in bilirubin handling from the fetal to the neonatal period.[25] Vaginal delivery as a significant risk factor suggests mechanisms that may extend beyond simple erythrocyte destruction from birth trauma, potentially involving an inflammatory state induced by labor stress or differences in gut microbiome colonization—which directly affects enterohepatic circulation of bilirubin—associated with different delivery modes.[26,27] Maternal infection as a defined risk factor further underscores the central role of inflammation in exacerbating hyperbilirubinemia.[28] The high risk associated with early-term GA highlights the critical importance of the final weeks of pregnancy for the maturation of hepatic enzyme systems (eg., UGT1A1), indicating that even slight prematurity can result in insufficient metabolic capacity.[24,29–31] Regarding protective factors, the risk reduction linked to multiparity suggests that previous pregnancies may induce an adaptive change in maternal immunity or placental function.[30,32] The protective effect

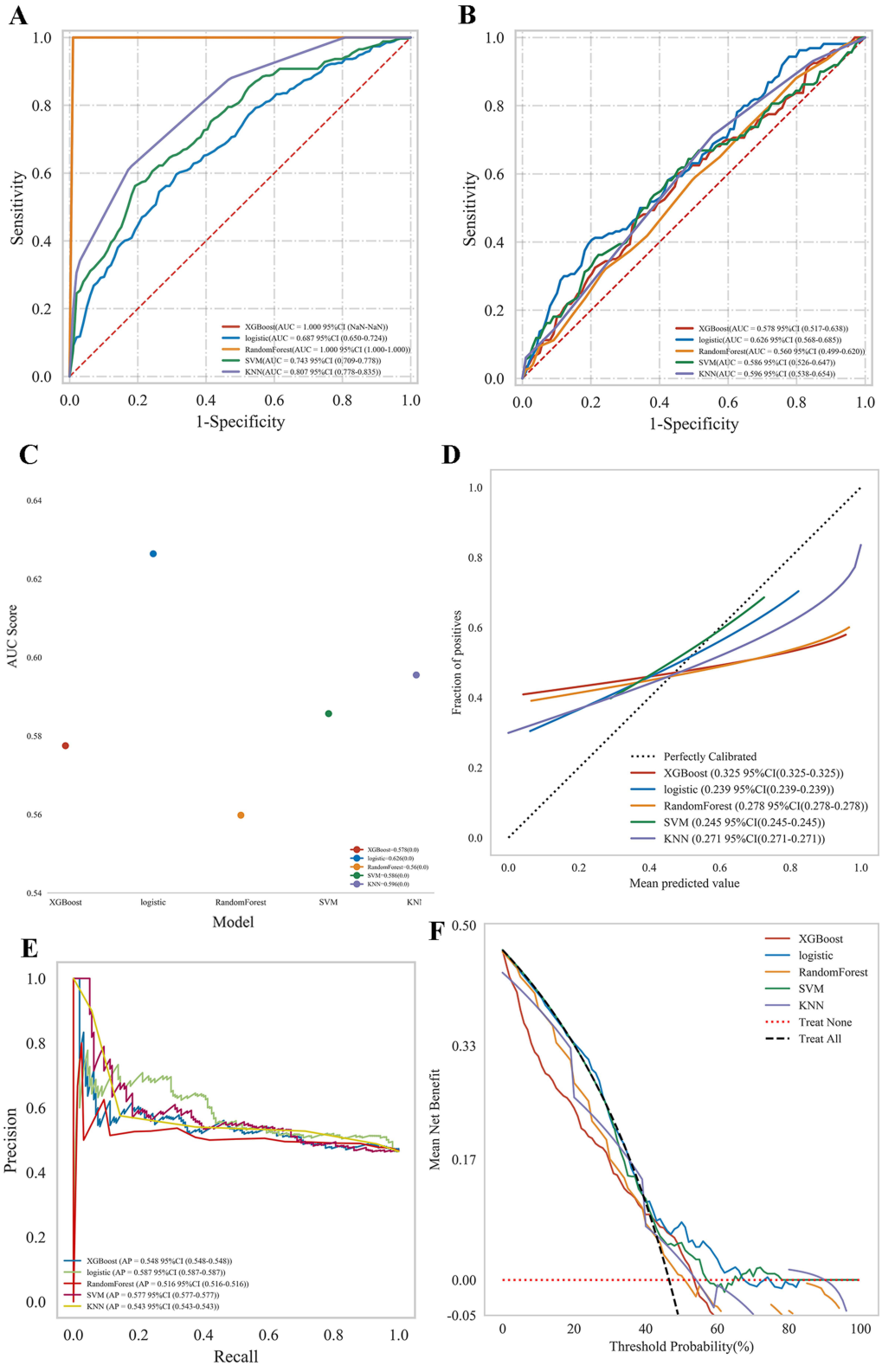

Figure 3 Comprehensive performance evaluation of machine learning models for predicting neonatal hyperbilirubinemia in the full-term group. (**A**) ROC curves for the training sets. (**B**) ROC curves for the validation sets. (**C**) Forest plot showing the AUC scores of each model. (**D**) Calibration curves for the validation set across all models. (**E**) PR curves for the validation set across all models. (**F**) DCA for the validation set across all models.

Abbreviations: XGBoost, Extreme Gradient Boosting; RF, Random Forest; SVM, Support Vector Machine; KNN, K-Nearest Neighbors; ROC, receiver operator characteristic curve; AUC, area under the curve; DCA, decision curve analysis; PR, precision-recall.

https://doi.org/10.2147/IJWH.S584032

Table 6 Predictive Performance of the ML Models in the Full-Term Group on Training and Validation Sets

Dataset	Model	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	Accuracy	F1 score
Training	XGBoost	1.000 (NaN, NaN)	1.000	1.000	1.000	1.000	1.000	1.000
	LR	0.687 (0.650, 0.724)	0.602	0.683	0.592	0.692	0.648	0.597
	RF	1.000 (1.000, 1.000)	0.994	1.000	1.000	0.996	0.998	0.997
	SVM	0.743 (0.709, 0.778)	0.565	0.806	0.690	0.708	0.702	0.621
	KNN	0.807 (0.778, 0.835)	0.614	0.826	0.729	0.737	0.734	0.667
Validation	XGBoost	0.578 (0.517, 0.638)	0.350	0.723	0.523	0.561	0.549	0.419
	LR	0.626 (0.568, 0.685)	0.487	0.658	0.553	0.596	0.578	0.518
	RF	0.560 (0.499, 0.620)	0.375	0.685	0.508	0.558	0.541	0.432
	SVM	0.586 (0.526, 0.647)	0.456	0.674	0.549	0.588	0.573	0.498
	KNN	0.596 (0.538, 0.654)	0.375	0.723	0.541	0.571	0.561	0.443

Abbreviations: XGBoost, Extreme Gradient Boosting; LR, logistic regression; RF, Random Forest; SVM, Support Vector Machine; KNN, K-Nearest Neighbors; ML, machine learning; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value.

associated with maternal hypothyroidism is a particularly intriguing finding that requires cautious interpretation. While we hypothesize that this might be related to more intensive medical monitoring and management received by these women throughout the perinatal period, this remains speculative. Our current dataset lacks specific data on the frequency of prenatal visits or targeted management protocols to substantiate this claim, highlighting the need for future studies to

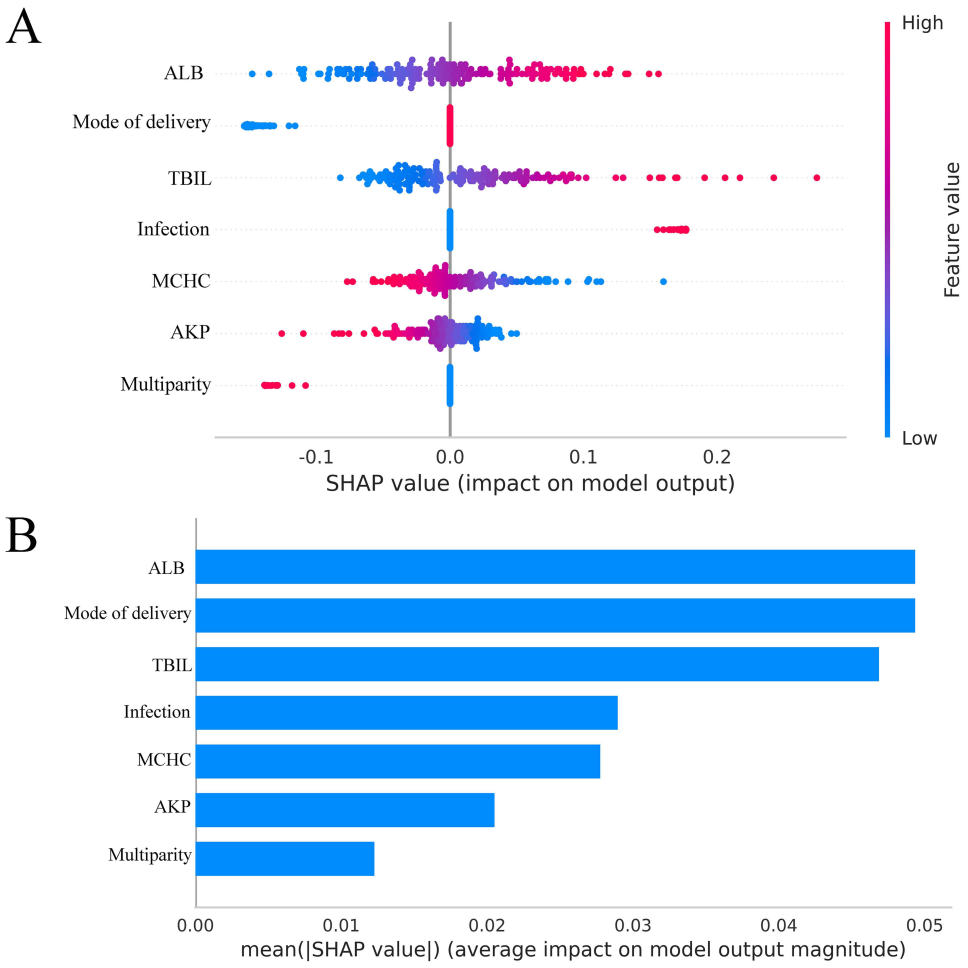

Figure 4 Visually interpret the best machine learning models using SHAP in the full-term group. (**A**) SHAP summary plot. (**B**) SHAP importance plot.
Abbreviations: MCHC, mean corpuscular hemoglobin concentration; AKP, alkaline phosphatase; ALB, albumin; TBIL, total bilirubin.

https://doi.org/10.2147/IJWH.S584032

Table 7 Univariate and Multivariate Logistic Regression Analyses of Risk Factors for Neonatal Hyperbilirubinemia in the early-term Group

Variables	Univariate		Multivariate	
	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
HGB	0.982 (0.970, 0.995)	0.006	0.987 (0.971, 1.002)	0.090
MCV	0.978 (0.958, 0.999)	0.039	0.993 (0.968, 1.018)	0.592
MCHC	0.980 (0.964, 0.996)	0.015	0.991 (0.971, 1.010)	0.334
PLT	1.004 (1.001, 1.006)	0.010	1.003 (1.001, 1.006)	0.026
TBIL	1.051 (1.001, 1.104)	0.045	1.063 (1.010, 1.122)	0.022
BUN	0.829 (0.708, 0.971)	0.020	0.828 (0.702, 0.974)	0.023
Multiparity (Yes)	0.579 (0.399, 0.839)	0.004	0.580 (0.391, 0.861)	0.007
Mod of delivery (Vaginal)	1.842 (1.366, 2.485)	<0.001	1.911 (1.395, 2.627)	<0.001
Age	0.973 (0.941, 1.007)	0.114	-	-
BMI	1.028 (0.982, 1.076)	0.244	-	-
WBC	0.985 (0.930, 1.044)	0.616	-	-
RBC	1.058 (0.744, 1.504)	0.754	-	-
HCT	0.019 (0.000, 2.269)	0.104	-	-
AKP	1.000 (0.997, 1.003)	0.994	-	-
ALB	1.008 (0.956, 1.063)	0.756	-	-
PALB	0.976 (0.942, 1.010)	0.167	-	-
Lymphocyte	1.100 (0.835, 1.450)	0.497	-	-
Neutrophils	0.995 (0.937, 1.056)	0.866	-	-
GGT	0.992 (0.975, 1.010)	0.401	-	-
TBA	0.985 (0.956, 1.016)	0.348	-	-
CRE	0.998 (0.980, 1.015)	0.789	-	-
HDP (Yes)	0.813 (0.428, 1.543)	0.527	-	-
GDM (Yes)	0.818 (0.584, 1.145)	0.241	-	-
Hypothyroidism (Yes)	0.514 (0.257, 1.030)	0.060	-	-
Infection (Yes)	1.482 (0.934, 2.354)	0.095	-	-
Placental abruption (Yes)	0.671 (0.134, 3.345)	0.626	-	-
PROM (Yes)	1.338 (0.969, 1.847)	0.077	-	-
Blood type				
O	Ref.	-	-	-
A	1.062 (0.727, 1.552)	0.755	-	-
AB	1.268 (0.681, 2.359)	0.454	-	-
B	1.393 (0.931, 2.085)	0.107	-	-
Unknown	1.554 (0.891, 2.712)	0.121	-	-

Abbreviations: OR, odds ratio; CI, confidence interval; BMI, body mass index; GA, gestational age; PROM, premature rupture of membrane; HDP, hypertensive disorders of pregnancy; GDM, gestational diabetes mellitus; WBC, white blood cell; RBC, red blood cell; HGB, hemoglobin; HCT, hematocrit; MCV, mean corpuscular volume; MCHC, mean corpuscular hemoglobin concentration; AKP, alkaline phosphatase; ALB, albumin; PALB, prealbumin; PLT, platelets; TBIL, total bilirubin; GGT, gamma-glutamyl transpeptidase; TBA, total bile acids; BUN, blood urea nitrogen.

further investigate the underlying mechanisms.[33,34] Furthermore, the predictive associations of routine laboratory parameters like MCHC and AKP should be interpreted strictly as exploratory observations rather than definitive pathophysiological mechanisms. Given their limited traditional association with neonatal hyperbilirubinemia, these findings might reflect dataset-specific correlations or serve as hypothesis-generating signals regarding specific erythrocyte characteristics or placental functional states, which require substantial corroboration in future studies. Finally, the SHAP analysis offers insights beyond traditional OR values by quantifying feature contribution to the model's decision-making. It visually demonstrates that factors like delivery mode and GA should be assigned higher priority in clinical risk assessment, thereby significantly enhancing the model's interpretability and potential for clinical translation.

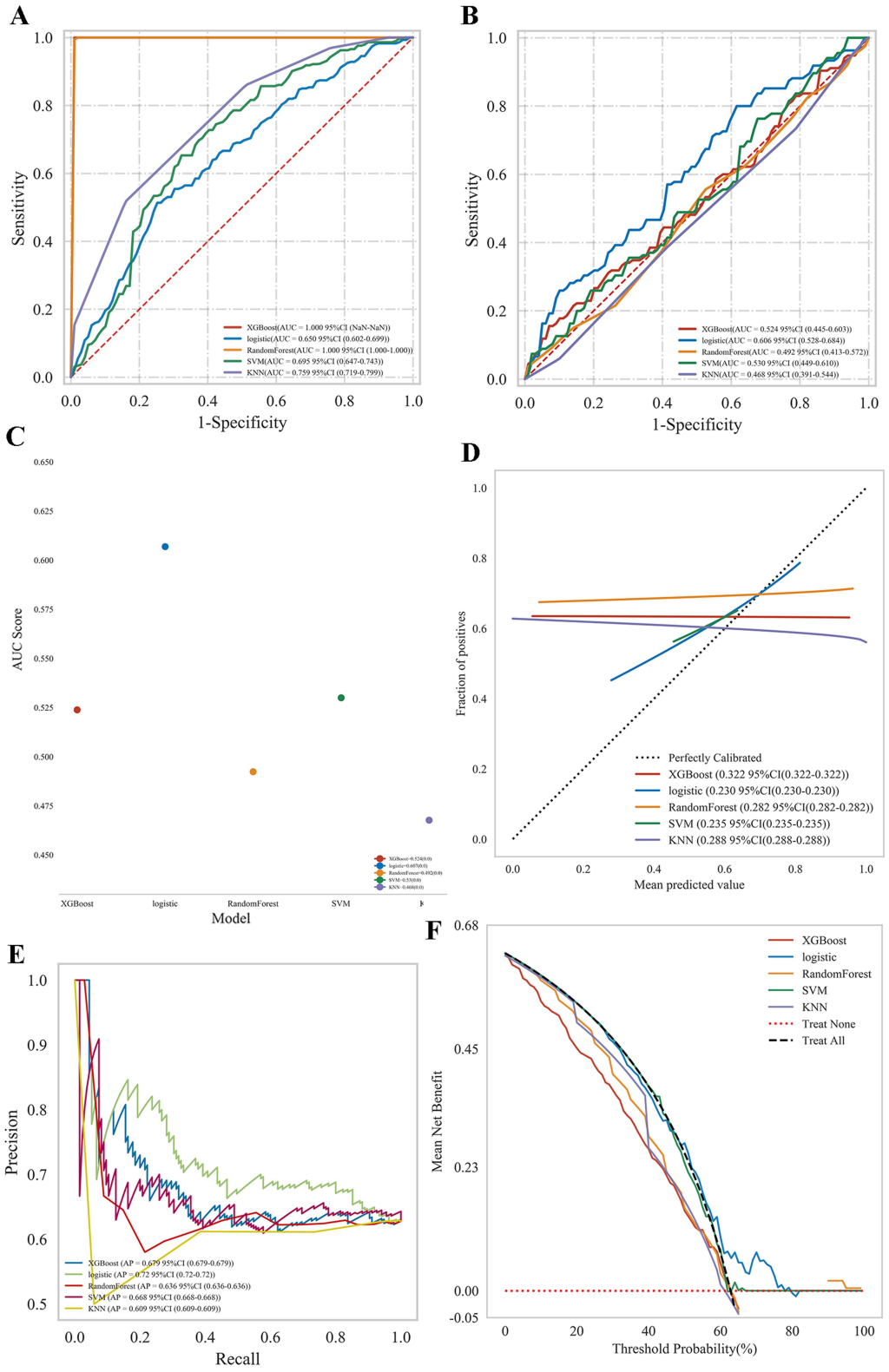

Figure 5 Comprehensive performance evaluation of machine learning models for predicting neonatal hyperbilirubinemia in the early-term group. (**A**) ROC curves for the training sets. (**B**) ROC curves for the validation sets. (**C**) Forest plot showing the AUC scores of each model. (**D**) Calibration curves for the validation set across all models. (**E**) PR curves for the validation set across all models. (**F**) DCA for the validation set across all models.

Abbreviations: XGBoost, Extreme Gradient Boosting; RF, Random Forest; SVM, Support Vector Machine; KNN, K-Nearest Neighbors; ROC, receiver operator characteristic curve; AUC, area under the curve; DCA, decision curve analysis; PR, precision-recall.

Table 8 Predictive Performance of the ML Models in the Early-Term Group on Training and Validation Sets

Dataset	Model	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	Accuracy	F1 score
Training	XGBoost	1.000 (NaN, NaN)	1.000	1.000	1.000	1.000	1.000	1.000
	LR	0.650 (0.602, 0.699)	0.514	0.749	0.744	0.520	0.611	0.608
	RF	1.000 (1.000, 1.000)	1.000	0.986	0.990	1.000	0.994	0.995
	SVM	0.695 (0.647, 0.743)	0.653	0.686	0.747	0.582	0.667	0.697
	KNN	0.759 (0.719, 0.799)	0.517	0.841	0.822	0.551	0.651	0.635
Validation	XGBoost	0.524 (0.445, 0.603)	0.511	0.487	0.627	0.371	0.502	0.563
	LR	0.606 (0.528, 0.684)	0.444	0.662	0.690	0.414	0.526	0.541
	RF	0.492 (0.413, 0.572)	0.622	0.362	0.622	0.362	0.526	0.622
	SVM	0.530 (0.449, 0.610)	0.489	0.525	0.635	0.378	0.502	0.552
	KNN	0.468 (0.391, 0.544)	0.385	0.588	0.612	0.362	0.460	0.473

Abbreviations: XGBoost, Extreme Gradient Boosting; LR, logistic regression; RF, Random Forest; SVM, Support Vector Machine; KNN, K-Nearest Neighbors; ML, machine learning; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value.

The interaction analysis revealed a significant interactive effect between GA and ALB, which carries important clinical implications. It indicates that the influence of maternal ALB levels on the risk of neonatal hyperbilirubinemia is not fixed but is modulated by GA. For instance, the same high maternal ALB level may confer a very high risk in full-term group, whereas in early-term group, the risk might be negligible. This finding strongly supports the necessity for personalized, stratified risk assessment. The in-depth understanding of this interaction also provides the key rationale for interpreting the subsequent subgroup analysis: precisely because GA is a critical effect modifier, a "one-size-fits-all" prediction model has limitations. The results of the subgroup analysis confirmed this heterogeneity. In the full-term subgroup, ALB replaced the mode of delivery as the most contributory predictor, but its positive association with risk

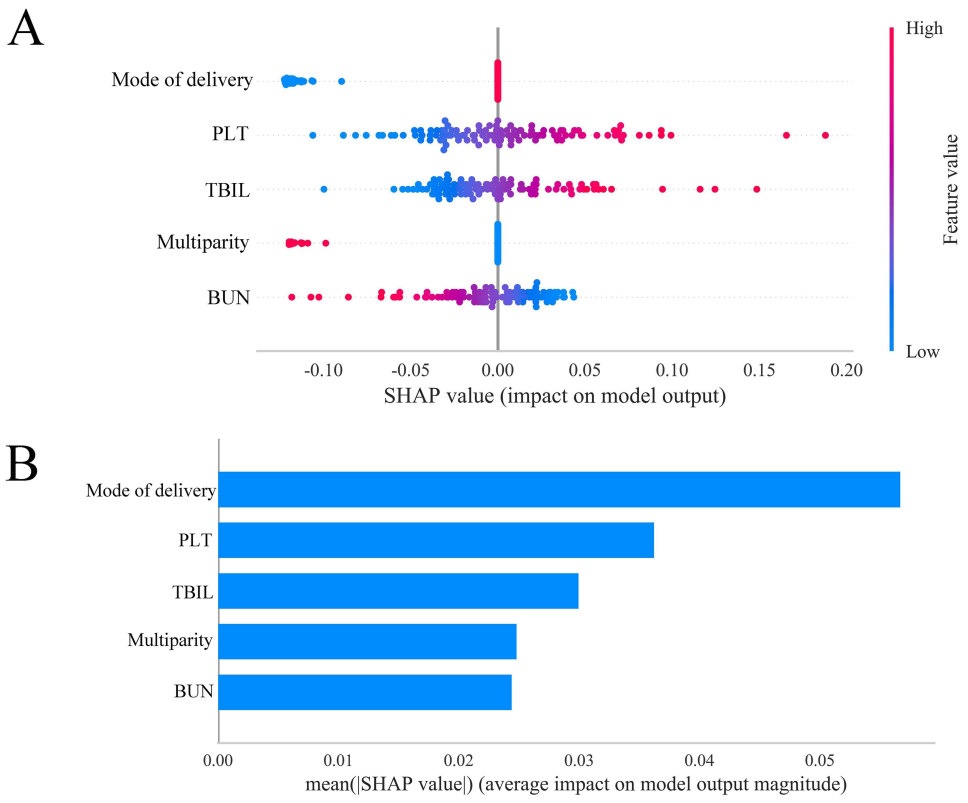

Figure 6 Visually interpret the best machine learning models using SHAP in the early-term group. (**A**) SHAP summary plot. (**B**) SHAP importance plot.
Abbreviations: PLT, platelets; TBIL, total bilirubin; BUN, blood urea nitrogen.

https://doi.org/10.2147/IJWH.S584032

(OR > 1) suggests there may be complex mechanisms beyond traditional nutritional explanations, potentially related to bilirubin transport or distribution.[12] Conversely, in the early-term subgroup, the spectrum of risk factors changed fundamentally, with indicators more directly linked to hemolysis and inflammatory potential, such as platelets, being selected, while factors significant in the overall model like ALB were not. This highlights the distinct pathophysiological context of this subpopulation.

In model development, LR proved to be the optimal choice in this study due to its stability and interpretability. Compared to complex models like RF and XGBoost, which overfitting on the training set, LR demonstrated superior generalizability. This is likely because the number of features included was limited and their relationships were approximately linear, making the simplicity of LR an advantage that avoided the overfitting risks common with "black-box" models, thereby enhancing its clinical utility. Although the model's discriminative ability is considered moderate, it is crucial to recognize that its core value lies in integrating multiple easily accessible routine prenatal indicators to provide a comprehensive risk assessment tool. DCA confirmed that this model yields a positive net clinical benefit across a wide range of threshold probabilities. Consequently, the most promising application of this model is as a prenatal screening tool to identify high-risk neonates. This could guide the implementation of targeted postnatal monitoring protocols, enabling early warning and intervention, and ultimately providing decision support for preventing severe hyperbilirubinemia.

This study has several limitations. First, as a single-center retrospective study, it is susceptible to selection bias, and the generalizability of the findings may be limited. Second, although the model demonstrated clinical utility, its discriminative ability remains limited, indicating room for improvement and emphasizing its role primarily as an early screening tool rather than a definitive diagnostic criterion. Third, despite including numerous clinical variables, some potentially important confounding factors were unmeasured, such as neonatal feeding patterns and reticulocyte counts. Fourth, the sample size for the late-term group was small, potentially compromising the stability and accuracy of the model in this specific population; as a result, this subgroup was not analyzed, which may limit the comprehensiveness of the heterogeneity analysis across gestational ages. Furthermore, our predictor selection strategy relied on the screening of a broad panel of routine maternal laboratory parameters without strictly predefined clinical justifications. Consequently, the identification of non-traditional risk factors (such as MCHC, AKP, PLT, and BUN) is inherently exploratory. We acknowledge the risk of data dredging associated with this approach and recognize that these data-driven associations may lack direct biological causality or deep clinical interpretability. While SHAP analysis enhances model transparency, it does not mitigate the potential for spurious predictor selection. Future studies employing penalized regression techniques or restricting candidate variables exclusively to clinically justified predictors are warranted to validate these exploratory markers. Finally, the model has currently only undergone internal validation; the lack of validation using an independent external dataset remains a key issue to be addressed in future research.

Conclusion

This study developed a preliminary, hypothesis-generating prediction model for neonatal hyperbilirubinemia based on antenatal maternal factors, demonstrating the critical modifying role of GA in risk stratification. Given the modest predictive performance and the exploratory inclusion of certain routine laboratory parameters, this model is not yet ready for direct clinical implementation. Therefore, distinct predictive models tailored to different GA subgroups should be viewed as an initial step. Rigorous external validation through multicenter, large-scale prospective studies, along with the refinement of clinically justified predictors, is crucially warranted to confirm its potential role and applicability in clinical practice.

Data Sharing Statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate

The study was approved by the Ethics Committee of Longyan First Affiliated Hospital of Fujian Medical University (Ethical approval number: LYREC2025-k176-01). All patients who were familiar with the contents and processes of the study and able to complete all the scheduled study processes signed the informed consent. Our study complies with the Declaration of Helsinki.

Funding

This work was supported by the following grants: Startup Fund for scientific research, Fujian Medical University of 2023 (Grant number: 2023QH1359).

Disclosure

The authors have conflicts of interest to disclose.

References

1. Kaplan M, Muraca M, Hammerman C, et al. Imbalance between production and conjugation of bilirubin: a fundamental concept in the mechanism of neonatal jaundice. *Pediatrics*. 2002;110(4):e47. doi:10.1542/peds.110.4.e47
2. Olusanya BO, Kaplan M, Hansen TWR. Neonatal hyperbilirubinaemia: a global perspective. *Lancet Child Adolesc Health*. 2018;2(8):610–620. doi:10.1016/S2352-4642(18)30139-1
3. American Academy of Pediatrics Subcommittee on H. Management of hyperbilirubinemia in the newborn infant 35 or more weeks of gestation. *Pediatrics*. 2004;114(1):297–316. doi:10.1542/peds.114.1.297.
4. Par EJ, Hughes CA, DeRico P. Neonatal Hyperbilirubinemia: evaluation and Treatment. *Am Fam Physician*. 2023;107(5):525–534.
5. Sgro M, Campbell DM, Kandasamy S, Shah V. Incidence of chronic bilirubin encephalopathy in Canada, 2007-2008. *Pediatrics*. 2012;130(4): e886–90. doi:10.1542/peds.2012-0253
6. Alken J, Hakansson S, Ekeus C, Gustafson P, Norman M. Rates of extreme neonatal hyperbilirubinemia and kernicterus in children and adherence to national guidelines for screening, diagnosis, and treatment in Sweden. *JAMA Netw Open*. 2019;2(3):e190858. doi:10.1001/jamanetworkopen.2019.0858
7. Donneborg ML, Hansen BM, Vandborg PK, Rodrigo-Domingo M, Ebbesen F. Extreme neonatal hyperbilirubinemia and kernicterus spectrum disorder in Denmark during the years 2000-2015. *J Perinatol*. 2020;40(2):194–202. doi:10.1038/s41372-019-0566-8
8. Varvarigou A, Fouzas S, Skylogianni E, Mantagou L, Bougioukou D, Mantagos S. Transcutaneous bilirubin nomogram for prediction of significant neonatal hyperbilirubinemia. *Pediatrics*. 2009;124(4):1052–1059. doi:10.1542/peds.2008-2322
9. Yu ZB, Han SP, Chen C. Bilirubin nomograms for identification of neonatal hyperbilirubinemia in healthy term and late-preterm infants: a systematic review and meta-analysis. *World J Pediatr*. 2014;10(3):211–218. doi:10.1007/s12519-014-0495-8
10. Casnocha Lucanova L, Zibolenova J, Matasova K, Zibolen M, Zibolen M. The use of transcutaneous bilirubin nomograms for the prevention of bilirubin neurotoxicity in the neonates. *Front Public Health*. 2023;11:1212667. doi:10.3389/fpubh.2023.1212667
11. Navaee M, Kashanian M, Kabir A, Zamaninour N, Chamari M, Pazouki A. Maternal and fetal/neonatal outcomes in pregnancy, delivery and postpartum following bariatric surgery and comparison with pregnant women with obesity: a study protocol for a prospective cohort. *Reprod Health*. 2024;21(1):8. doi:10.1186/s12978-023-01736-3
12. Wei H, Chang X, Ji R, Tang Y. Association between late pregnancy A/G ratio and the risk of neonatal admission for neonatal hyperbilirubinemia. *BMC Pregnancy Childbirth*. 2025;25(1):563. doi:10.1186/s12884-025-07706-w
13. Bahado-Singh RO, Dashe J, Deren O, Daftary G, Copel JA, Ehrenkranz RA. Prenatal prediction of neonatal outcome in the extremely low-birth-weight infant. *Am J Obstet Gynecol*. 1998;178(3):462–468. doi:10.1016/s0002-9378(98)70421-1
14. Gulbiniene V, Balciuniene G, Petroniene J, et al. The significance of epidermal growth factor in noninvasively obtained amniotic fluid predicting respiratory outcomes of preterm neonates. *Int J Mol Sci*. 2022;23(6). doi:10.3390/ijms23062978
15. Betts KS, Kisely S, Alati R. Predicting neonatal respiratory distress syndrome and hypoglycaemia prior to discharge: leveraging health administrative data and machine learning. *J Biomed Inform*. 2021;114:103651. doi:10.1016/j.jbi.2020.103651
16. Jiang N, Qian L, Lin G, et al. Maternal blood parameters and risk of neonatal pathological jaundice: a retrospective study. *Sci Rep*. 2023;13 (1):2627. doi:10.1038/s41598-023-28254-3
17. Deng H, Zhou Y, Wang L, Zhang C. Ensemble learning for the early prediction of neonatal jaundice with genetic features. *BMC Med Inform Decis Mak*. 2021;21(1):338. doi:10.1186/s12911-021-01701-9
18. Hulzebos CV, Vitek L, Coda Zabetta CD, et al. Screening methods for neonatal hyperbilirubinemia: benefits, limitations, requirements, and novel developments. *Pediatr Res*. 2021;90(2):272–276. doi:10.1038/s41390-021-01543-1
19. Wang ML, Dorer DJ, Fleming MP, Catlin EA. Clinical outcomes of near-term infants. *Pediatrics*. 2004;114(2):372–376. doi:10.1542/peds.114.2.372
20. Bhutani VK, Johnson L, Sivieri EM. Predictive ability of a predischarge hour-specific serum bilirubin for subsequent significant hyperbilirubinemia in healthy term and near-term newborns. *Pediatrics*. 1999;103(1):6–14. doi:10.1542/peds.103.1.6
21. Guidelines for the Diagnosis and Management of Gestational Diabetes Mellitus. *Chin J Obstet Gynecol*. 2014;49(8):561–569. doi:10.3760/cma.j.issn.0529-567X.2014.08.001
22. Guidelines for the Diagnosis and Management of Hypertensive Disorders of Pregnancy. *Chin J Obstet Gynecol*. 2015;50(10):721–728. doi:10.3760/cma.j.issn.0529-567x.2015.10.001

23. Liu J, Jiang W, Yu Y, et al. Applying machine learning to predict bowel preparation adequacy in elderly patients for colonoscopy: development and validation of a web-based prediction tool. *Ann Med*. 2025;57(1):2474172. doi:10.1080/07853890.2025.2474172
24. Castillo A, Grogan TR, Wegrzyn GH, Ly KV, Walker VP, Calkins KL. Umbilical cord blood bilirubins, gestational age, and maternal race predict neonatal hyperbilirubinemia. *PLoS One*. 2018;13(6):e0197888. doi:10.1371/journal.pone.0197888
25. Knudsen A, Lebech M. Maternal bilirubin, cord bilirubin, and placenta function at delivery and the development of jaundice in mature newborns. *Acta Obstet Gynecol Scand*. 1989;68(8):719–724. doi:10.3109/00016348909006145
26. Chang PF, Lin YC, Liu K, Yeh SJ, Ni YH. Risk of hyperbilirubinemia in breast-fed infants. *J Pediatr*. 2011;159(4):561–565. doi:10.1016/j.jpeds.2011.03.042
27. Zhou S, Wang Z, He F, et al. Association of serum bilirubin in newborns affected by jaundice with gut microbiota dysbiosis. *J Nutr Biochem*. 2019;63:54–61. doi:10.1016/j.jnutbio.2018.09.016
28. Norman M, Naver L, Soderling J, et al. Association of maternal SARS-CoV-2 infection in pregnancy with neonatal outcomes. *JAMA*. 2021;325(20):2076–2086. doi:10.1001/jama.2021.5775
29. Mehrad-Majd H, Haerian MS, Akhtari J, Ravanshad Y, Azarfar A, Mamouri G. Effects of Gly71Arg mutation in UGT1A1 gene on neonatal hyperbilirubinemia: a systematic review and meta-analysis. *J Matern Fetal Neonatal Med*. 2019;32(10):1575–1585. doi:10.1080/14767058.2017.1410789
30. Norman M, Aberg K, Holmsten K, Weibel V, Ekeus C. Predicting nonhemolytic neonatal hyperbilirubinemia. *Pediatrics*. 2015;136(6):1087–1094. doi:10.1542/peds.2015-2001
31. Muchowski KE. Evaluation and treatment of neonatal hyperbilirubinemia. *Am Fam Physician*. 2014;89(11):873–878.
32. Lloyd DJ, Belgaumkar TK, Scott KE, Wort AJ, Aterman K, Krause VW. Prevention of group-B beta-haemolytic streptococcal septicaemia in low-birth-weight neonates by penicillin administered within two hours of birth. *Lancet*. 1979;1(8118):713–715. doi:10.1016/s0140-6736(79)91160-7
33. Hussain S, Asghar I, Sabir MU, Chattha MN, Tarar SH, Mushtaq R. Prevalence and pattern of congenital malformations among neonates in the neonatal unit of a teaching hospital. *J Pak Med Assoc*. 2014;64(6):629–634.
34. Kiran Z, Sheikh A, Humayun KN, Islam N. Neonatal outcomes and congenital anomalies in pregnancies affected by hypothyroidism. *Ann Med*. 2021;53(1):1560–1568. doi:10.1080/07853890.2021.1970798