Am J Cancer Res 2025;15(1):99-112
www.ajcr.us /ISSN:2156-6976/ajcr0160713

Original Article

Establishment and validation of a prognostic risk early-warning model for retinoblastoma based on XGBoost

Feng Wang[1], Jian Wang[2]

[1]*Department of Ophthalmology, Changzhi People's Hospital, Changzhi 046000, Shanxi, China;* [2]*Department of Radiology, Shanxi Provincial People's Hospital, Taiyuan 030012, Shanxi, China*

Received September 24, 2024; Accepted December 20, 2024; Epub January 15, 2025; Published January 30, 2025

Abstract: Retinoblastoma (RB) is the most common intraocular malignancy in children, and early detection and treatment are crucial for improving patient outcomes. Conventional treatments, such as enucleation and radiotherapy, have limitations in fully addressing prognosis. This study aimed to establish and validate an early-warning prognostic model for RB based on the XGBoost algorithm to improve the prediction accuracy of the 5-year survival rate in children. A retrospective analysis was conducted on 320 children with RB treated at Changzhi People's Hospital between February 2012 and April 2019. The patients were randomly divided into a training group (n=224) and a validation group (n=96). Clinical data, including age, gender, tumor characteristics, and tumor marker levels, were collected. Prognostic factors were analyzed using XGBoost and Cox regression models, and model performance was evaluated using various statistical methods. No significant differences were observed in baseline data between the two sets (P>0.05). Cox regression analysis identified tumor diameter (P=0.032), IIRC stage (P<0.001), and NSE (P=0.016) as independent prognostic factors. The XGBoost model achieved an area under the curve (AUC) of 0.951 in the training group, significantly higher than the Cox model (P=0.001), while in the validation group, the XGBoost model's AUC was 0.902, with no significant difference compared to the Cox model (P=0.117). The XGBoost model demonstrated high accuracy and clinical utility in predicting the 5-year survival of children with RB. Decision curve analysis (DCA) and calibration curves further confirmed that the XGBoost model offers higher clinical net benefits and superior calibration ability across various thresholds.

Keywords: Retinoblastoma, XGBoost, prognostic evaluation, cox regression, risk early-warning model

Introduction

Retinoblastoma (RB), one of the most prevalent intraocular malignancies in children, accounts for 3% of all childhood tumors [1]. It primarily affects children under the age of 5, typically manifested as strabismus and visual impairments [2]. The onset of RB is primarily attributed to biallelic loss-of-function mutations of the RB1 gene, leading to retinal cell cycle dysregulation and uncontrolled cell proliferation [3]. Early detection and prompt treatment are crucial for improving patient prognosis. Currently, ultrasound and magnetic resonance imaging (MRI) are the primary methods for diagnosing RB. Ultrasound scanning is a fast, non-invasive, and relatively inexpensive imaging technique that can offer detailed images of the eye's internal structure, allowing for assessment of tumor size and location [3]. MRI offers superior soft tissue contrast, aiding in determining tumor extend, especially in evaluating extraocular spread [4]. These imaging modalities provide valuable information before treatment, helping to optimize the treatment plan.

Therapeutic approaches for RB include ophthalmectomy, external radiotherapy (ERT), radioactive plaque therapy, cryotherapy, photocoagulation, and chemotherapy [5]. Ophthalmectomy, or enucleation, is typically utilized in advanced cases or when vision preservation is unfeasible, making it a radical treatment option [6]. ERT and radiation plaque therapy utilize radiation to target and destroy cancer cells. The latter involves implanting a radioactive source

https://doi.org/10.62347/WHUQ1208

directly into or near the eye, maximizing protection of surrounding normal tissue [7]. Cryotherapy and photocoagulation eradicate tumor cells by local freezing or laser coagulation, generally applicable to smaller tumors or serve as adjunctive treatments [8]. Chemotherapy, administered via intravenous or intra-arterial injection, plays a role in systemic treatment. The intra-arterial approach, in particular, increases drug concentration at the tumor site while minimizing systemic side effects [9]. Nevertheless, these treatment modalities cannot completely prevent disease progression. A portion of patients still encounter extraocular metastasis following treatment, significantly influencing their survival rate.

In recent years, the Cox regression model has been widely used in the prognostic evaluation of RB. It analyzes patient survival data to identify key factors influencing the survival rate with considerable accuracy [10]. However, the Cox regression model has some limitations when dealing with complex and high-dimensional data. With the development of machine learning techniques, algorithms like eXtreme Gradient Boosting (XGBoost) are increasingly being applied in medical predictions [11]. XGBoost builds multiple weak classifiers and gradually optimizes the model, improving prediction accuracy and robustness [12]. This machine learning method is extensively used in the medical field, especially in disease prognosis assessment. As an integrated learning algorithm, XGBoost enhances prediction performance by constructing and optimizing multiple weak classifiers [13]. Compared to the traditional Cox regression model, XGBoost excels at handling complex and high-dimensional data. Recently, XGBoost has been widely applied in various medical prediction tasks, such as cancer prognosis evaluation and disease risk prediction [14, 15]. This study aims to develop a prognostic risk early-warning model for RB based on the XGBoost algorithm and validate its effectiveness.

Methods and materials

This retrospective study included 320 children with RB who received treatment at Changzhi People's Hospital from February 2012 to April 2019. The patients were divided into a training set (n=224) and a validation set (n=96) in a 7:3 ratio. This study was approved by the Ethics Committee of Changzhi People's Hospital. The patient's informed consent was waived due to the retrospective nature of this study.

Inclusion and exclusion criteria

Inclusion criteria: Children aged 0-6 years with a pathologically confirmed diagnosis of intraocular RB [16]; International Intraocular Retinoblastoma Classification (IIRC) staging of C to E; Availability of complete medical history.

Exclusion criteria: Children with other malignancies, or with heart, liver, kidney, and other vital organ dysfunction, as well as blood and/or immune system diseases; Eyes with clinical high-risk factors, including anterior chamber infiltration, iris neovascularization, suspected retrobulbar optic nerve metastasis, or extraocular tumor invasion; Children with an estimated survival time of less than 6 months.

Clinical data acquisition

Clinical data of children were collected through the hospital's management system, including age, sex, tumor diameter (≥20 mm or <20 mm), affected side (unilateral or bilateral), growth pattern (intraocular or extraocular), IIRC stage (C, D, or E) [17], histological classification (differentiated or undifferentiated), tumor residual (present or absent), brain metastasis (present or absent), choroidal infiltration (present or absent), and the pre-treatment levels of biomarkers such as carbohydrate antigen (CA) 199 (CA199), CA153, and neuron-specific enolase (NSE).

Detection of tumor markers

Peripheral blood (5 ml) were collected one day before treatment, and the serum was separated by centrifugation. Tumor markers, CA199, CA153, and NSE, were then measured using Myeri automatic chemiluminescence immunoassay analyzer (CL-1200i). All reagents were supplied by the manufacturer.

Follow-up

Patients were enrolled in the follow-up program after completing their treatment. During the first year, follow-up visits were scheduled every three months. In the second year, follow-up fre-

quency was adjusted to every four months for the continuous tracking of the patient's health status. From the third to the fifth year, follow-up frequency was extended to every six months.

XGBoost model

The XGBoost model was constructed by first converting the target variable into binary labels, where 'P' labeled as 1 and the alternative as 0. The dataset was split into features and the target variable, and the predictor variables were formatted into a DMatrix optimized for XGBoost computation. Key model parameters were set as follows: the "gbtree" booster was used, a "binary" objective for binary classification was selected, AUC was chosen as the evaluation metric, the learning rate (eta) was set to 0.01, the maximum tree depth was 6, and both the subsample and colsample_bytree ratios were set to 0.8 to control overfitting. A 10-fold cross-validation was performed over 1000 boosting rounds, with early stopping applied after after 10 rounds if no improvement was observed. The best iteration was selected based on the highest AUC. The model was then retrained with the optimal number of rounds, and its performance was evaluated through a receiver operating characteristic (ROC) curve with the corresponding AUC.

Outcome measures

Primary outcome measures: 1. Cox regression was used to identify the prognostic factors affecting the 5-year survival of children. 2. The XGBoost machine learning algorithm was applied to identify key prognostic factors and establish a prediction model. 3. The risk scores for children in both the training and validation groups were calculated and compared.

Secondary outcome measures: 1. Baseline data for children in the training and validation groups were compared. 2. ROC curves were used to evaluate the performance of the XGBoost and Cox models in predicting 5-year survival. 3. Clinical net benefit and accuracy of the model were assessed using decision curve analysis (DCA) and calibration curves. 4. Net reclassification index (NRI) analysis was conducted to evaluate the effectiveness of the model in re-classification of children's survival risk.

Statistical analyses

Baseline characteristics were analyzed using SPSS version 26.0. Continuous variables were first tested for normality using the Kolmogorov-Smirnov (KS) test. If the data followed a normal distribution, the independent sample t-test was applied; otherwise, the rank-sum test (Wilcoxon test) was used for non-normally distributed data. Categorical variables were analyzed using the chi-square test.

Further statistical analyses were performed using R version 4.3.2. The following packages were used: the 'stats' package for basic statistics and chi-square tests, the 'survival' package for Cox regression analysis, the 'xgboost' package for XGBoost model construction, and the 'pROC' package for ROC curve drawing and area under the curve (AUC) calculation. The 'rms' package was employed for DCA, the 'Hmisc' or calibrate package to drawing calibration curves, the 'survMisc' package for NRI analysis, and the 'coin' package for performing the Delong test. A *P*-value of less than 0.05 was considered statistically significant.

Results

Comparison of baseline data between training and validation groups

The comparison of baseline data between the training group (n=224) and the validation group (n=96) revealed no significant differences across all variables. The age distribution was similar between the two groups (P=0.379), as was the gender distribution (P=0.643). Tumor diameter (≥20 mm or <20 mm) showed no significant difference (P=0.329), and the affected side (unilateral or bilateral) was also comparable (P=0.532). Additionally, no significant differences were observed in tumor growth patterns (intraocular vs. extraocular, P=0.526), IIRC stage (C, D, E; P=0.687), or histological classification (differentiated vs. undifferentiated, P=0.417). Tumor residual (P=0.456), brain metastasis (P=0.402), and choroidal infiltration (P=0.313) were evenly distributed between the groups. Finally, serum levels of tumor markers CA199 (P=0.978), CA153 (P=0.865), and NSE (P=0.738) also showed no significant differences, indicating that the two groups were well-matched in terms of baseline characteristics (**Table 1**).

Table 1. Baseline data of children in the training and validation groups

Variable	Total	Training group (n=224)	Validation group (n=96)	P
Age				
<1	130 (40.62%)	90 (40.18%)	40 (41.67%)	0.379
1-4	151 (47.19%)	103 (45.98%)	48 (50%)	
>4	39 (12.19%)	31 (13.84%)	8 (8.33%)	
Sex				
Male	157 (49.06%)	108 (48.21%)	49 (51.04%)	0.643
Female	163 (50.94%)	116 (51.79%)	47 (48.96%)	
Tumor diameter				
≥20 mm	160 (50%)	108 (48.21%)	52 (54.17%)	0.329
<20 mm	160 (50%)	116 (51.79%)	44 (45.83%)	
Affected side				
Unilateral	195 (60.94%)	139 (62.05%)	56 (58.33%)	0.532
Bilateral	125 (39.06%)	85 (37.95%)	40 (41.67%)	
Growth pattern				
Intraocular growth	158 (49.38%)	108 (48.21%)	50 (52.08%)	0.526
Extraocular growth	162 (50.62%)	116 (51.79%)	46 (47.92%)	
IIRC staging				
C	28 (8.75%)	18 (8.04%)	10 (10.42%)	
D	200 (62.5%)	143 (63.84%)	57 (59.38%)	0.687
E	92 (28.75%)	63 (28.12%)	29 (30.21%)	
Histological typing				
Differentiated	181 (56.56%)	130 (58.04%)	51 (53.12%)	0.417
Undifferentiated	139 (43.44%)	94 (41.96%)	45 (46.88%)	
Tumor residual				
With	130 (40.62%)	94 (41.96%)	36 (37.5%)	0.456
Without	190 (59.38%)	130 (58.04%)	60 (62.5%)	
Brain metastasis				
With	61 (19.06%)	40 (17.86%)	21 (21.88%)	0.402
Without	259 (80.94%)	184 (82.14%)	75 (78.12%)	
Choroidal infiltration				
With	75 (23.44%)	49 (21.88%)	26 (27.08%)	0.313
Without	245 (76.56%)	175 (78.12%)	70 (72.92%)	
CA199 (U/mL)		24.90±8.71	24.92±7.68	0.978
CA153 (U/mL)		42.81±15.67	43.13±14.42	0.865
NSE (ng/mL)		42.34±14.67	42.95±15.23	0.738

Note: IIRC staging, International Immune-related Response Criteria Staging; CA199, Carbohydrate Antigen 19-9; CA153, Carbohydrate Antigen 15-3; NSE, Neuron-Specific Enolase.

Cox regression analysis of prognostic factors for 5-year survival

In the univariate Cox regression analysis, several clinical and pathological factors were evaluated for their potential influence on 5-year survival in children with retinoblastoma. These factors included age (P=0.685), sex (P=0.402), tumor diameter (P=0.002), affected side (P=0.273), growth pattern (P=0.376), IIRC stage (P<0.001), histological classification (P=0.828), tumor residual (P=0.843), brain metastasis (P=0.056), choroidal infiltration (P=0.210), CA199 (P=0.218), CA153 (P=0.243), and NSE (P=0.036). Among these factors, tumor diameter (P=0.002), IIRC stage (P<0.001), and NSE (P=0.036) were statistically significant, suggesting they may play an important role in determining survival outcomes (**Table 2**).

Table 2. Univariate Cox regression analysis

Factor	Beta	StdErr	*P* Value	HR	Lower 95 CI	Upper 95 CI
Age	-0.120	0.297	0.685	0.887	0.495	1.587
Sex	-0.342	0.408	0.402	0.710	0.319	1.581
Tumor diameter	1.543	0.500	0.002	4.677	1.755	12.464
Affected side	0.489	0.445	0.273	1.630	0.681	3.903
Growth pattern	-0.361	0.408	0.376	0.697	0.313	1.551
IIRC staging	2.189	0.460	0.000	8.929	3.622	22.013
Histological typing	0.089	0.408	0.828	1.093	0.491	2.433
Tumor residual	-0.081	0.408	0.843	0.922	0.414	2.053
Brain metastasis	0.820	0.429	0.056	2.270	0.980	5.261
Choroidal infiltration	-0.771	0.615	0.210	0.463	0.138	1.546
CA199 (U/mL)	-0.031	0.025	0.218	0.969	0.922	1.019
CA153 (U/mL)	0.016	0.014	0.243	1.016	0.989	1.043
NSE (ng/mL)	0.027	0.013	0.036	1.027	1.002	1.053

Note: IIRC staging, International Immune-related Response Criteria Staging; CA199, Carbohydrate Antigen 19-9; CA153, Carbohydrate Antigen 15-3; NSE, Neuron-Specific Enolase.

Table 3. Multivariate Cox regression analysis

Factor	Beta	StdErr	*P* Value	HR	Lower 95 CI	Upper 95 CI
Tumor diameter	1.103	0.513	0.032	3.014	1.102	8.242
IIRC staging	2.198	0.469	0.000	9.005	3.589	22.596
NSE (ng/mL)	0.030	0.012	0.016	1.030	1.006	1.055

Note: IIRC staging, International Immune-related Response Criteria Staging; NSE, Neuron-Specific Enolase.

In the multivariate Cox regression analysis, variables that were significant in the univariate analysis were further analyzed to identify independent prognostic factors. The results indicated that tumor diameter (P=0.032, HR= 3.014), IIRC stage (P<0.001, HR=9.005), and NSE levels (P=0.016, HR=1.030) were independent predictors of 5-year survival (**Table 3**).

Screening of prognostic factors for 5-year survival by XGBoost

We used the XGBoost machine learning algorithm to identify prognostic factors for children with RB. The input variables for the model included key clinical and pathological data such as age, sex, tumor size, affected side (unilateral or bilateral), growth pattern (intraocular or extraocular), IIRC stage, histological classification, tumor residual, brain metastasis, choroidal infiltration, and the pre-treatment levels of tumor markers including NSE, CA199, and CA153. Through the model's feature selection process, we identified the most important prognostic factors after 66 iterations (**Figure 1A**). The final model highlighted NSE (ng/mL), IIRC staging, CA153 (U/mL), CA199 (U/mL), tumor diameter, and brain metastasis as the most significant variables contributing to the prediction of 5-year survival outcomes (**Figure 1B**). These factors were found to have the highest predictive value in the XGBoost model.

Comparison of risk scores in children

Risk scores for each child in the training and validation groups were calculated based on the formulas derived from both the XGBoost and Cox models. A comparison revealed that in the training group, the risk scores for deceased children were significantly higher than those for surviving children (P<0.001, **Figure 2A**). Similarly, in the validation group, deceased children had higher risk scores in both the XGBoost and Cox models compared to survivors (P<0.001, **Figure 2B**).

The value of XGBoost and Cox models in predicting the 5-year survival of pediatric patients

To further compare the performance of the XGBoost and Cox models in predicting the 5-year survival of children, we used ROC curves

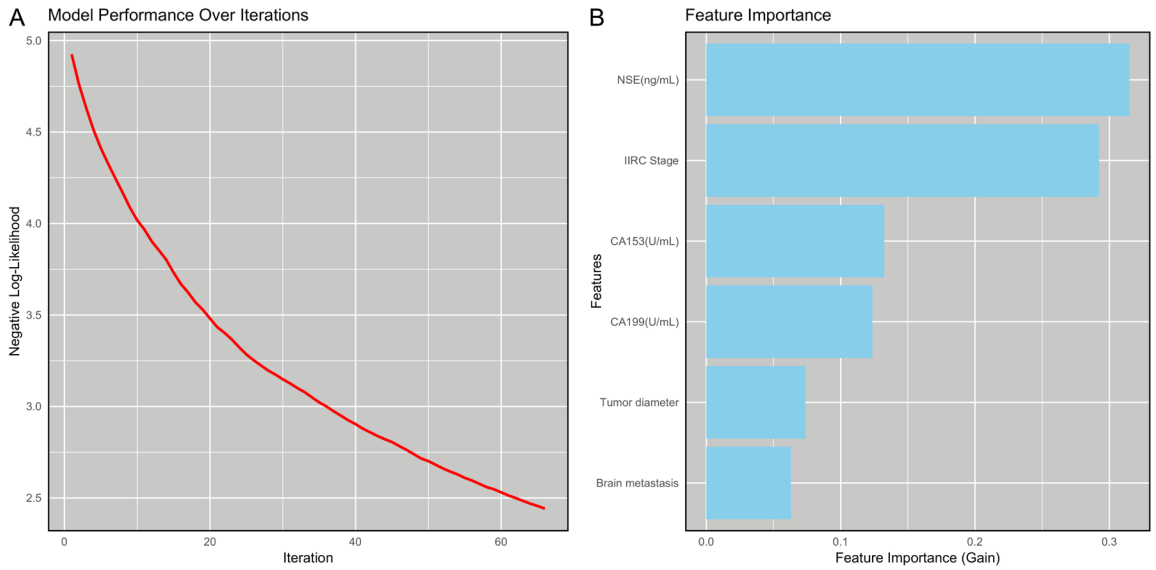

Figure 1. XGBoost model training and feature factor screening. Note: XGBoost, Extreme Gradient Boosting.

to assess the AUC of the two models, and their predictive performance was compared using Delong tests. The results showed that the XGBoost model in the training group achieved an AUC of 0.951 and an F1 score of 65.67%; while the Cox model had an AUC of 0.837 and an F1 value of 57.63% (**Figure 3A**). In the validation group, the AUC and F1 of the XGBoost model were 0.902 and 75.68%, respectively, compared with 0.953 and 85.11% of the Cox model (**Figure 3B**). The Delong test revealed an obviously higher AUC for the XGBoost model compared to the Cox model in the training group (P=0.001, **Tables 4**, **5**), but their difference in the validation group did not reach statistical significance (P=0.117, **Tables 4**, **5**). In addition, NRI analysis revealed that the NRI of the XGBoost model was 0.265, while that for Cox model was 0.065, indicating an improvement in the reclassification of negative-class children (**Table 6**). Overall, the XGBoost model demonstrated higher predictive value in predicting the 5-year survival rate of pediatric RB patients.

Analysis of the stability and clinical utility of the models via DCA, calibration curve, and survival curve

In this study, we analyzed the stability and clinical utility of the models through DCA and calibration curves. The results indicated that in the training group, the net benefit curve of the XGBoost model consistently outperformed the Cox model. Moreover, the calibration curve exhibited a high degree of alignment with the actual observed results, indicating that the XGBoost model provides higher clinical net benefits and superior calibration across various threshold probabilities. In contrast, the calibration curve of the Cox model displayed some deviations, with noticeable discrepancies between the predicted and observed outcomes (**Figures 4A**, **5A**, **5B**). In the validation group, the net benefit curve of the XGBoost model remained significantly higher than that of the Cox model, emphasizing its superiority in clinical decision-making. The calibration curve for the XGBoost model maintained a strong fit with observed results, further substantiating its stability and predictive accuracy. In contrast, the calibration effect of the Cox model in the validation group remained suboptimal, with a considerable deviation between the prediction results and the observed results (**Figures 4B**, **5C**, **5D**). Survival curve analysis indicated that, in both the XGBoost and Cox models, the 5-year survival rate of children in the low-score group was higher than those in the high-score group (**Figure 6**). Therefore, the XGBoost model demonstrated high stability and significant clinical application value in both the training and validation groups.

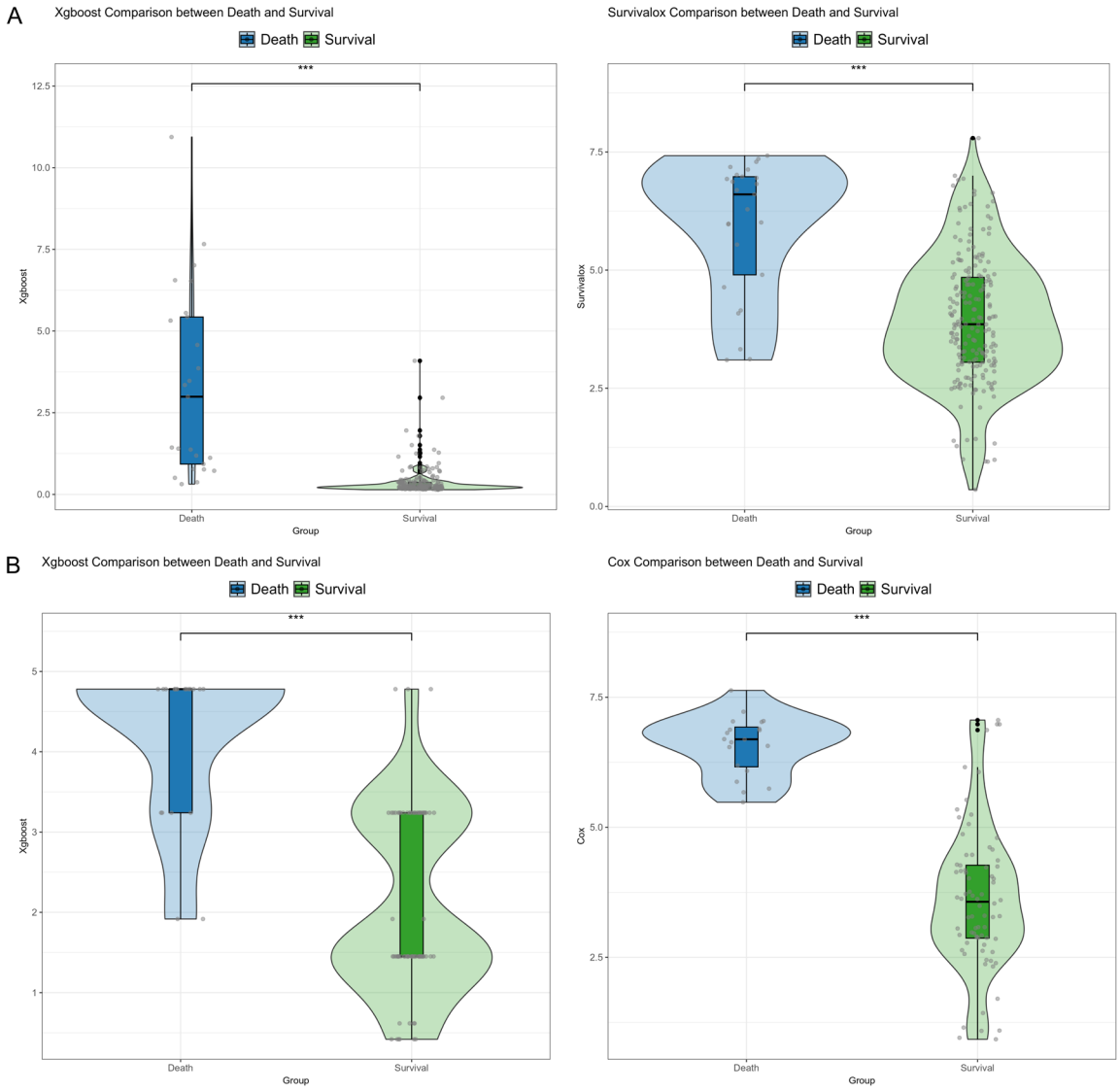

Figure 2. Comparison of risk scores between survived and dead patients based on XGBoost and Cox regression models in both training and validation groups. A. Comparison of risk scores between survived and dead patients based on XGBoost and Cox regression models in the training group; B. Comparison of risk scores between survived and dead patients based on XGBoost and Cox regression models in the validation. Note: XGBoost, Extreme Gradient Boosting; Cox, Cox Proportional Hazards Model.

Discussion

With the advancement of machine learning technology, novel algorithms like XGBoost have witnessed an expanding application scope in medical predictions [18]. Compared with the traditional Cox regression model, XGBoost shows better performance when dealing with complex and high-dimensional data [19]. The objective of this study is to establish and validate a prognostic risk early-warning model for retinoblastoma (RB) based on the XGBoost algorithm, aiming to enhance the prediction accuracy of the 5-year survival rate of children and provide more effective support for clinical decision-making.

In this study, we found that the XGBoost model demonstrated high accuracy in predicting the 5-year survival rate of children with RB. The AUCs in the training and validation groups were 0.951 and 0.902, respectively, significantly higher than the AUC of the Cox regression model in the training group (0.837). While the

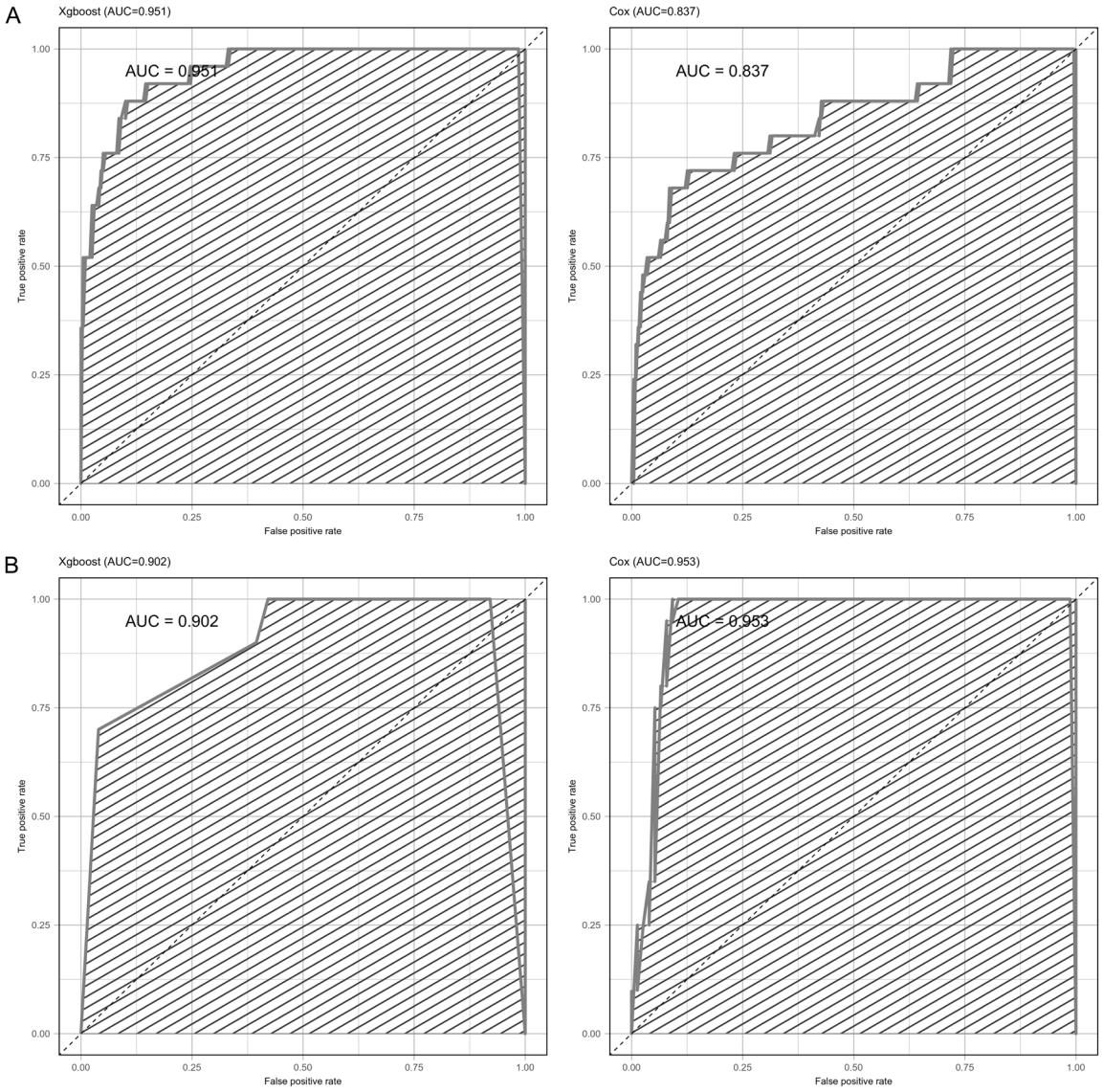

Figure 3. ROC curves of XGBoost and Cox regression models for predicting children's survival in the training and validation groups. A. ROC curves for XGBoost and Cox regression models in the training group; B. ROC curves for XGBoost and Cox regression models in the validation group. Note: ROC curves, Receiver Operating Characteristic curves; XGBoost, Extreme Gradient Boosting; Cox, Cox Proportional Hazards Model.

AUCs of the two models did not differ significantly in the validation group, the XGBoost model showed superior clinical net benefits and calibration ability across various thresholds, as evidenced by DCA and calibration curves. Additionally, the NRI analysis confirmed that the XGBoost model offered better reclassification of survival risks, further proving its efficacy in prognostic prediction.

These findings align with previous studies that have demonstrated the superiority of XGBoost over traditional statistical methods. For example, Mao et al. [20] applied the XGBoost algorithm to a cohort of follicular thyroid cancer patients, achieving improved prediction accuracy compared to the Kaplan-Meier and Cox models. Similarly, Zhang et al. [21] employed XGBoost, along with other machine learning algorithms, to predict bone metastasis in lung adenocarcinoma patients, yielding AUC values of 0.913 and 0.853 in the training and test sets, respectively, further validating the algorithm's robustness. Moreover, Liu et al. [22]

Table 4. ROC curve parameters of two prediction models

Marker	Dataset	AUC	CI lower upper	Specificity	Sensitivity	Youden index	Cut off	Accuracy	Precision	F1 score
XGBoost	Training group	0.951	0.915-0.987	89.95%	88.00%	77.95%	0.714	89.73%	88.00%	65.67%
Cox		0.837	0.742-0.932	91.46%	68.00%	59.46%	5.928	88.84%	68.00%	57.63%
XGBoost	Validation group	0.902	0.835-0.969	96.05%	70.00%	66.05%	4.01	90.62%	70.00%	75.68%
Cox		0.953	0.912-0.994	90.79%	100.00%	90.79%	5.413	92.71%	100.00%	85.11%

Note: AUC, Area Under the Curve; XGBoost, Extreme Gradient Boosting; Cox, Cox Proportional Hazards Model.

Table 5. Comparison of AUCs of the two prediction models

Marker 1	Marker 2	Dataset	Z value	*P* value	AUC difference	CI lower upper
XGBoost	Cox	Training set	3.29	0.001	0.114	0.046-0.182
XGBoost	Cox	Validation set	-1.567	0.117	-0.051	-0.115-0.013

Note: AUC, Area Under the Curve; XGBoost, Extreme Gradient Boosting; Cox, Cox Proportional Hazards Model.

Table 6. NRI results

Variable	Estimate	Lower	Upper
NRI	0.265	-0.187	0.559
NRI+	0.200	-0.259	0.500
NRI-	0.065	0.005	0.092
Pr (Up|Case)	0.280	<0.001	0.550
Pr (Down|Case)	0.080	<0.001	0.296
Pr (Down|Ctrl)	0.065	0.010	0.092
Pr (Up|Ctrl)	0.000	<0.001	0.005

Note: NRI, Net Reclassification Improvement.

demonstrated the high prediction accuracy and stability of the XGBoost algorithm in colon cancer recurrence prediction, with an AUC of 0.962 in the training set and 0.952 in the validation set. Collectively, these studies underscore the advantages of XGBoost in managing complex, high-dimensional data and providing accurate predictions, reinforcing its value in clinical decision-making. By citing these studies, we emphasize that XGBoost is particularly suited for handling complex datasets, such as those involving multiple clinical and biomarker variables, which traditional models like Cox regression may struggle to process. The enhanced predictive accuracy of XGBoost observed in our study and others demonstrates its potential to significantly enhance clinical outcomes in retinoblastoma and other cancers.

To further explore the prognostic factors in our study, we analyzes the relationship between several key clinical variables and the prognosis of RB. These variables include IIRC stage, tumor diameter, and brain metastasis, all of which have been consistently identified in existing literature as significant predictors of prognosis. By conducting a detailed analysis of these factors, we seek not only to validate the effectiveness of the XGBoost model in predicting survival outcomes but also to gain a deeper understanding of which factors most strongly influence long-term survival in children. This insight is crucial for guiding individualized treatment plans. The IIRC staging system is widely used to assess the severity and treatment approach for RB. Higher IIRC stages (such as D and E) indicate larger, more invasive tumors, which are associated with a poorer prognosis [23]. Previous research has demonstrated that IIRC stage is inversely correlated with patient survival rates, meaning that higher stages are linked to a lower 5-year survival rate [24]. Similarly, tumor diameter is another critical factor, as larger tumors typically reflect a more advanced disease stage, complicating treatment and leading to worse outcomes. Studies have shown that patients with a tumor diameter greater than 16 mm have significantly lower 5-year survival rates compared to those with smaller tumors [25]. Furthermore, the presence of brain metastasis is often indicative of advanced, aggressive disease and is associated with significantly reduced survival rates [26]. Understanding the influence of these key factors reinforces the predictive accuracy of our model and provides valuable support for developing tailored treatment strategies based on individual risk profiles.

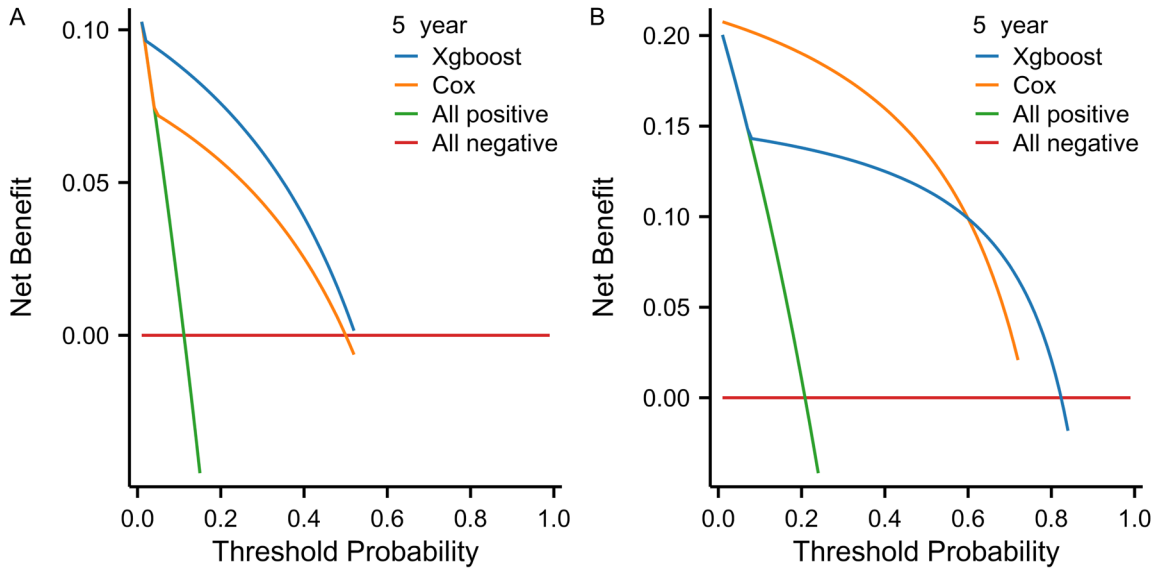

Figure 4. DCA curves of XGBoost and Cox regression models in the training and validation groups. A. DCA curves of the XGBoost and Cox regression models in the training group; B. DCA curves of the XGBoost and Cox regression models in the validation group. Note: DCA, Decision Curve Analysis; XGBoost, Extreme Gradient Boosting; Cox, Cox Proportional Hazards Model.

Although no existing literature directly links CA199 and CA153 with the prognosis of RB, some studies have highlighted the high specificity of these tumor markers in diagnosing RB [27]. Nevertheless, our study found that these tumor markers were associated with prognosis, suggesting a potential role that warrants further exploration in future research. NSE is a well-established prognostic marker in RB. Research has indicated [28] that NSE level in children with RB decrease significantly after chemotherapy compared to pre-chemotherapy levels. Elevated NSE levels in RB patients are typically associated with more aggressive tumors and worse prognosis. For example, Lu et al. [29] found that the average NSE level in patients with a postoperative survival time ≥5 years was notably higher than those with a survival time <5 years. Similarly, Comoy et al. [30] observed that the postoperative NSE levels of two RB patients who survived for more than 5 years were 2.8 and 8.7 mg/L, respectively, considerably lower than those of the other five patients, further suggesting that NSE is an effective prognostic indicator in RB. In addition, a recent study by Guo et al. [31] developed a prognostic nomogram for RB, incorporating factors such as tumor size, laterality, and residence. This nomogram, based on data from the SEER database, found that tumor size was a key predictor of overall survival (OS), aligning with our findings that tumor diameter and other clinical markers such as NSE are strongly associated with prognosis in RB. A nomogram developed by Guo et al. demonstrated a C-index of 0.71 for predicting survival, whereas our XGBoost model achieved a higher predictive accuracy, with an AUC of 0.951 in the training group. This comparison highlights the potential of integrating additional biomarkers like CA199, CA153, and NSE into prognostic models to further enhance the accuracy of survival predictions for retinoblastoma patients.

Although we have successfully constructed an early-warning prognostic model for RB based on XGBoost, there are still some limitations. First, the retrospective design may introduce selection bias and limit the determination of causality. Second, the five-year follow-up period was insufficient to fully capture late complications or long-term survival outcomes. Third, the complexity and parameter adjustments required for the XGBoost model could limit its practical clinical application. Additionally, we did not account for treatment variables, such as surgery and chemotherapy, which can sig-

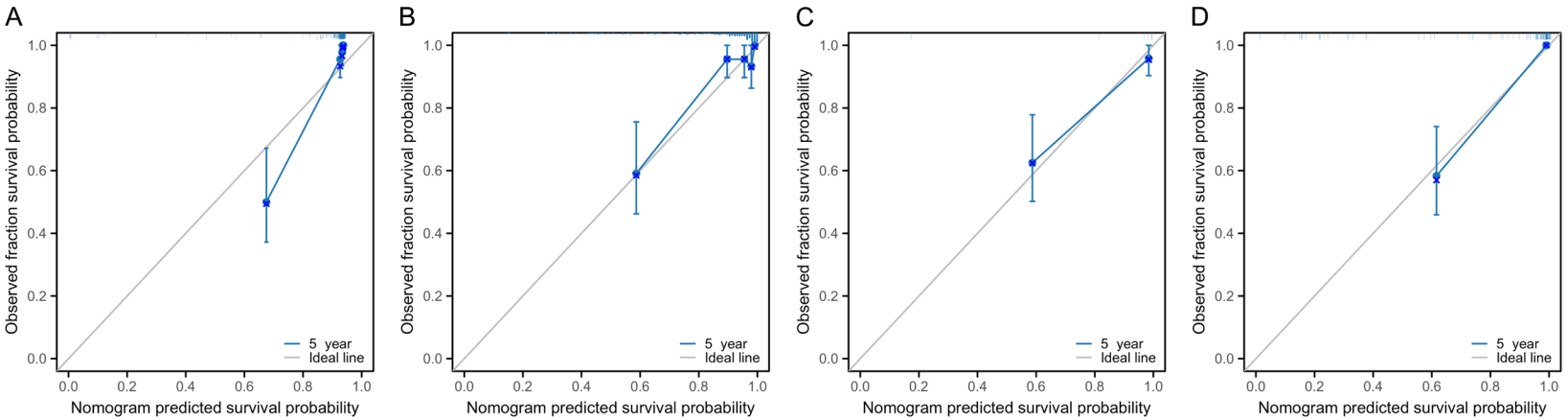

Figure 5. Calibration curves of XGBoost and Cox regression models in the training and validation groups. A, B. Calibration curves of XGBoost and Cox regression models in the training group; C, D. Calibration curves of XGBoost and Cox regression models in the validation group. Note: XGBoost, Extreme Gradient Boosting; Cox, Cox Proportional Hazards Model.

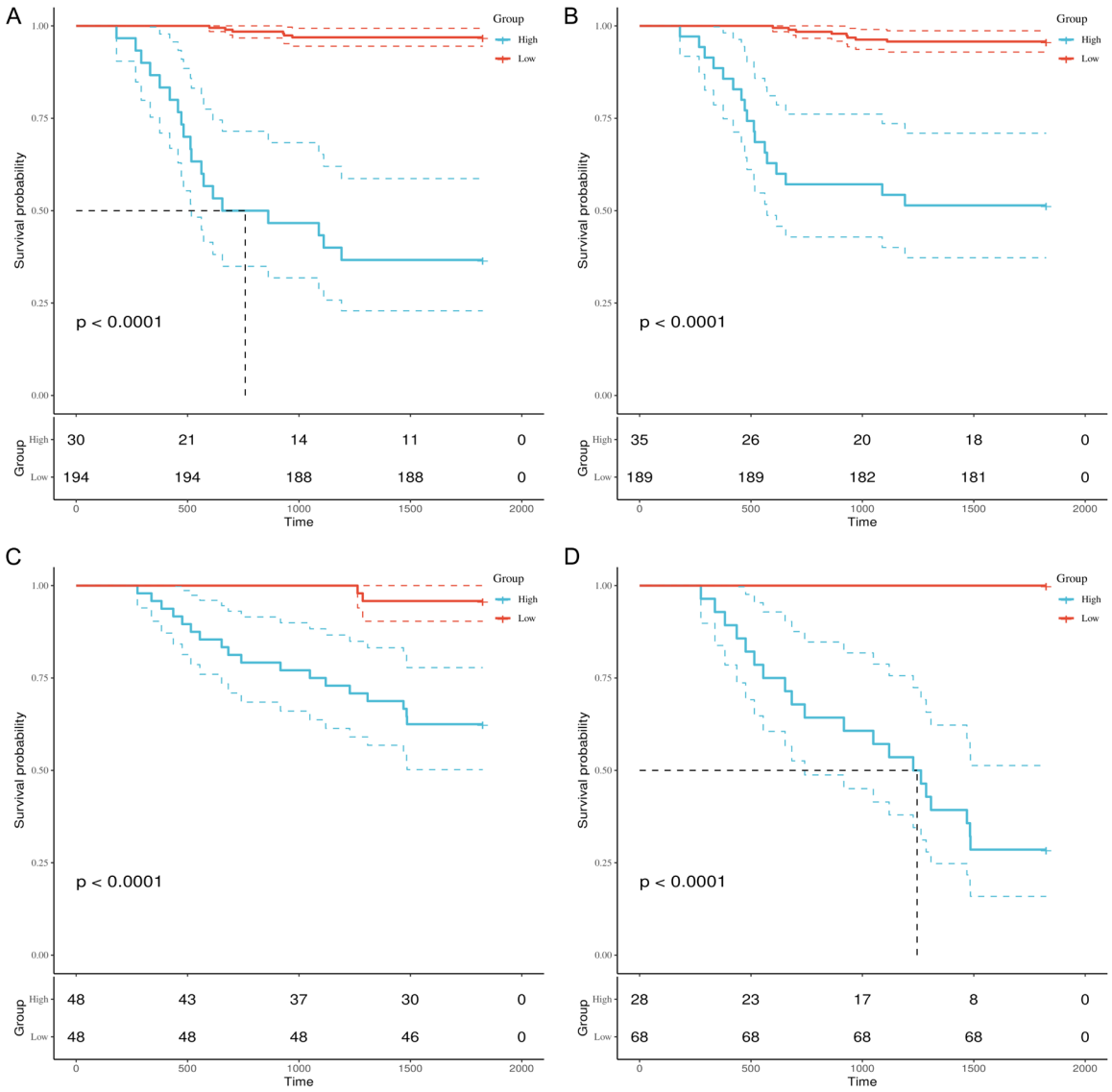

Figure 6. Survival curves for patients with high/low risk scores based on XGBoost and Cox regression models in both training and validation groups. A, B. Survival curves of patients with high/low risk scores based on XGBoost and Cox regression models in the training group; C, D. Survival curves of patients with high/low risk scores based on XGBoost and Cox regression models in the validation group. Note: XGBoost, Extreme Gradient Boosting; Cox, Cox Proportional Hazards Model.

nificantly influence patient outcomes and provide a more comprehensive assessment of prognosis. Finally, the lack of external dataset validation may affect the stability and prediction accuracy of the model. Future studies should address these limitations by adopting a prospective design, extending the follow-up period to 10 years or longer, simplifying and optimizing the model, incorporating treatment data, and utilizing multi-center datasets for external validation. These improvements will enhance the clinical applicability and reliability of the model.

Conclusion

The XGBoost model shows high accuracy and clinical utility in predicting the 5-year survival of children with RB. It provides superior clinical net benefits and better calibration ability under different threshold probabilities, as confirmed by decision curve analysis (DCA) and calibration curve analysis.

Disclosure of conflict of interest

None.

Address correspondence to: Jian Wang, Department of Radiology, Shanxi Provincial People's Hospital, Taiyuan 030012, Shanxi, China. Tel: +86-0351-4960023; E-mail: wangjtysx@163.com

References

[1] Peeler CE and Gonzalez E. Retinoblastoma. N Engl J Med 2022; 386: 2412.
[2] Cobrinik D. Retinoblastoma origins and destinations. N Engl J Med 2024; 390: 1408-1419.
[3] Mehyar M, Mosallam M, Tbakhi A, Saab A, Sultan I, Deebajah R, Jaradat I, AlJabari R, Mohammad M, AlNawaiseh I, Al-Hussaini M and Yousef YA. Impact of RB1 gene mutation type in retinoblastoma patients on clinical presentation and management outcome. Hematol Oncol Stem Cell Ther 2020; 13: 152-159.
[4] Abramson DH, Dunkel IJ and Francis JH. Magnetic resonance imaging of metastatic retinoblastoma. J Pediatr Ophthalmol Strabismus 2023; 60: 152.
[5] Madan R, Radhakrishnan V, Meel R, Chinnaswamy G, Singh L, Kulkarni S, Sasi A, Kaur T, Sharma J, Dhaliwal RS, Haldorai M, Rath GK and Bakhshi S. Management of extraocular retinoblastoma: ICMR consensus guidelines. Indian J Pediatr 2024; 91: 1157-1165.
[6] Rating P, Bornfeld N, Schlüter S, Westekemper H, Kiefer T, Stuschke M, Göricke S, Ketteler P, Ting S, Metz KA, Bechrakis NE and Biewald E. Long-term results after intraocular surgery in treated retinoblastoma eyes. Ocul Oncol Pathol 2022; 8: 161-167.
[7] Yousef YA, Mohammad M, Jaradat I, Shatnawi R, Mehyar M and Al-Nawaiseh I. Prognostic factors for eye globe salvage by external beam radiation therapy for resistant intraocular retinoblastoma. Oman J Ophthalmol 2020; 13: 123-128.
[8] Nag A and Khetan V. Retinoblastoma - a comprehensive review, update and recent advances. Indian J Ophthalmol 2024; 72: 778-788.
[9] Carnevale JA, Goldberg J, Kocharian G, Rivera M, Giantini Larsen A, Garton A, Ramos A, Francis JH, Abramson DH and Pierre Gobin Y. Intra-arterial chemotherapy for retinoblastoma. J Neurointerv Surg 2023; 15: 303-304.
[10] Global Retinoblastoma Study Group. The global retinoblastoma outcome study: a prospective, cluster-based analysis of 4064 patients from 149 countries. Lancet Glob Health 2022; 10: e1128-e1140.
[11] Haug CJ and Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. N Engl J Med 2023; 388: 1201-1208.
[12] Chen L, Wang Y, Cai C, Ding Y, Kim RS, Lipchik C, Gavin PG, Yothers G, Allegra CJ, Petrelli NJ, Suga JM, Hopkins JO, Saito NG, Evans T, Jujjavarapu S, Wolmark N, Lucas PC, Paik S, Sun M, Pogue-Geile KL and Lu X. Machine learning predicts oxaliplatin benefit in early colon cancer. J Clin Oncol 2024; 42: 1520-1530.
[13] Moon I, LoPiccolo J, Baca SC, Sholl LM, Kehl KL, Hassett MJ, Liu D, Schrag D and Gusev A. Machine learning for genetics-based classification and treatment response prediction in cancer of unknown primary. Nat Med 2023; 29: 2057-2067.
[14] Wong D and Yip S. Machine learning classifies cancer. Nature 2018; 555: 446-447.
[15] Zhang Y, Shen S, Li X, Wang S, Xiao Z, Cheng J and Li R. A multiclass extreme gradient boosting model for evaluation of transcriptomic biomarkers in Alzheimer's disease prediction. Neurosci Lett 2024; 821: 137609.
[16] Singh L, Chinnaswamy G, Meel R, Radhakrishnan V, Madan R, Kulkarni S, Sasi A, Kaur T, Dhaliwal RS and Bakhshi S. Epidemiology, diagnosis and genetics of retinoblastoma: ICMR consensus guidelines. Indian J Pediatr 2024; 91: 1147-1156.
[17] Randhawa JK, Kim ME, Polski A, Reid MW, Mascarenhas K, Brown B, Fabian ID, Kaliki S, Stacey AW, Burner E, Sayegh CS, Poblete RA, Ji X, Zou Y, Sultana S, Rashid R, Sherief ST, Cassoux N, Garcia J, Coronado RD, López AMZ, Ushakova T, Polyakov VG, Roy SR, Ahmad A, Reddy MA, Sagoo MS, Al Harby L, Astbury NJ, Bascaran C, Blum S, Bowman R, Burton MJ, Gomel N, Keren-Froim N, Madgar S, Zondervan M and Berry JL. The effects of breastfeeding on retinoblastoma development: results from an international multicenter retinoblastoma survey. Cancers (Basel) 2021; 13: 4773.
[18] Li B, Eisenberg N, Beaton D, Lee DS, Aljabri B, Verma R, Wijeysundera DN, Rotstein OD, de Mestral C, Mamdani M, Roche-Nagle G and Al-Omran M. Using machine learning (XGBoost) to predict outcomes after infrainguinal bypass for peripheral artery disease. Ann Surg 2024; 279: 705-713.
[19] Cai C, Li H, Zhang L, Li J, Duan S, Fang Z, Li C, Chen H, Alharbi M, Ye L, Liu Y and Zeng Z. Machine learning identification of nutrient intake variations across age groups in metabolic syndrome and healthy populations. Nutrients 2024; 16: 1659.
[20] Mao Y, Huang Y, Xu L, Liang J, Lin W, Huang H, Li L, Wen J and Chen G. Surgical methods and social factors are associated with long-term survival in follicular thyroid carcinoma: construction and validation of a prognostic model based on machine learning algorithms. Front Oncol 2022; 12: 816427.
[21] Zhang Y, Xiao L, LYu L and Zhang L. Construction of a predictive model for bone metastasis from first primary lung adenocarcinoma within

3 cm based on machine learning algorithm: a retrospective study. PeerJ 2024; 12: e17098.
[22] Liu Y, Du W, Guo Y, Tian Z and Shen W. Identification of high-risk factors for recurrence of colon cancer following complete mesocolic excision: an 8-year retrospective study. PLoS One 2023; 18: e0289621.
[23] Xing L, Zhang L, Feng Y, Cui Z and Ding L. Downregulation of circular RNA hsa_circ_0001649 indicates poor prognosis for retinoblastoma and regulates cell proliferation and apoptosis via AKT/mTOR signaling pathway. Biomed Pharmacother 2018; 105: 326-333.
[24] Mohammad M, Shehada R, Al-Nawaiseh I, Mehyar M, AlHussaini M, Jaradat I, Sultan I, Halalsheh H, Khzouz J and Yousef YA. A comparison of high risk pathological features between primary and secondary enucleation for retinoblastoma. Eur J Ophthalmol 2023; 33: 2014-2023.
[25] Guo X, Wang L, Beeraka NM, Liu C, Zhao X, Zhou R, Yu H, Fan R and Liu J. Incidence trends, clinicopathologic characteristics, and overall survival prediction in retinoblastoma children: SEER prognostic nomogram analysis. Oncologist 2024; 29: e275-e281.
[26] Hu H, Zhang W, Huang D, Wang Y, Zhang Y, Yi Y, Liu A and Li J. Clinical characteristics, treatment and prognosis of paediatric patients with metastatic neuroblastoma to the brain. Clin Neurol Neurosurg 2019; 184: 105372.
[27] Liu ZP, Zhou KY, Chen LL, Xiao ZH and Chen YZ. A preliminary study of retinoblastoma-related serum tumor markers. Zhongguo Dang Dai Er Ke Za Zhi 2017; 19: 318-321.
[28] Hu H, Zhang W, Wang Y, Huang D, Shi J, Li B, Zhang Y and Zhou Y. Characterization, treatment and prognosis of retinoblastoma with central nervous system metastasis. BMC Ophthalmol 2018; 18: 107.
[29] Lu Y and Wang JF. Analysis of prognostic factors of retinoblastoma. Zhonghua Yan Ke Za Zhi 2009; 45: 935-939.
[30] Comoy E, Roussat B, Henry I and Bloch-Michel E. Neuron-specific enolase in the aqueous humor. Its significance in the differential diagnosis of retinoblastoma. Ophtalmologie 1990; 4: 233-235.
[31] Guo X, Wang L, Beeraka NM, Liu C, Zhao X, Zhou R, Yu H, Fan R and Liu J. Incidence trends, clinicopathologic characteristics, and overall survival prediction in retinoblastoma children: SEER prognostic nomogram analysis. Oncologist 2024; 29: e275-e281.