

Machine learning-based predictive models for the prognosis of pineoblastoma in pediatric, adolescents, and young adults

Yajun Jing, Hongliang Ge, Yiming Chen, Yiming Chen, Qiu He, JinYE Su, Surmai Shukla, Wenhua Fang, Yuanxiang Lin, Dezhi Kang

Submitted to: JMIR Cancer
on: April 11, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
Supplementary Files.....	27

Preprint
JMIR Publications

Machine learning-based predictive models for the prognosis of pineoblastoma in pediatric, adolescents, and young adults

Yajun Jing¹ MD, PhD; Hongliang Ge¹ MD; Yiming Chen² BD; Yiming Chen² BD; Qiu He³ PhD; Jinye Su¹ MD; Surmai Shukla³ BD; Wenhua Fang⁴ MD; Yuanxiang Lin¹ PhD; Dezhi Kang¹ MD, PhD

¹ First Affiliated Hospital of Fujian Medical University Fuzhou CN

² Fujian Medical University Fuzhou CN

³ Northwestern University Evanston US

⁴ First Affiliated Hospital of Fujian Medical University Fuzhou

Corresponding Author:

Yajun Jing MD, PhD

First Affiliated Hospital of Fujian Medical University
Number twenty Chazhong Road Fuzhou City
Fuzhou
CN

Abstract

Background: Pineoblastoma (PB), a rare entity of the central nervous system, has limited available data on clinical outcomes due to its infrequency. Uncertainty surrounds the optimal treatment approaches and prognostic factors for PB.

Objective: This study aims to identify factors associated with prognosis and to develop a predictive model for PB in adolescents and young adults.

Methods: First, we identified the prognostic factors for PB using data for 304 PB patients (> 30 years) from the Surveillance, Epidemiology, and End Results database. We then developed prediction models utilizing several machine learning approaches and evaluated the predictive power of various models using the concordance index, Brier score, and receiver operating characteristic curve. Finally, we established SHAP graphs for model interpretability and evaluated the clinical usefulness of the model by decision curve analysis (DCA).

Results: The present study revealed that age, year of diagnosis, laterality, tumor-size, stage, surg/rad sequence, and therapeutic approach were significant prognostic factors for PB patients ($p < 0.05$). The examination of model performance revealed that the random forest model has superior predictive power, displaying a potent discriminative ability and accuracy, as well as great consistency. Importantly, the DCA highlighted the clinical usefulness of this model.

Conclusions: In summary, the present study identified the prognostic factors for PB and established the most effective random forest model for predicting the survival of PB patients.

(JMIR Preprints 11/04/2025:75834)

DOI: <https://doi.org/10.2196/preprints.75834>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in [a JMIR journal](#), my article will remain visible to all users.

No. Please do not make my accepted manuscript PDF available to anyone.



Original Manuscript

Title page**Machine learning-based predictive models for the prognosis of pineoblastoma in pediatric and young adult populations**

Yajun Jing^{# a,b}, Hongliang Ge^{# a,b,c}, Yiming Chen^{# d}, Qiu He^{a,d}, Jinye Su^a, Surmai Shukla^{d,e}, Wenhua Fang^{✉ a,g,h}, Yuanxiang Lin^{✉ a,b,c,g,h}, Dezhi Kang^{✉ a,b,c,g,h}

Institution

- a. Department of Neurosurgery, Neurosurgery Research Institute, The First Affiliated Hospital, Fujian Medical University, Fuzhou 350005, Fujian, China.
- b. Department of Neurosurgery, National Regional Medical Center, Binhai Campus of the First Affiliated Hospital, Fujian Medical University, Fuzhou 350212, Fujian, China.
- c. Fujian Provincial Institutes of Brain Disorders and Brain Sciences, First Affiliated Hospital, Fujian Medical University, Fuzhou 350005, Fujian, China.
- d. Department of Medical Laboratory, College of Medical Technology and Engineering, Fujian Medical University, Fuzhou, Taijiang District, Fuzhou, Fujian, China, 350004.
- e. Department of Immunology, Northwestern University, 2122 Sheridan Road, Suite 120, Evanston, IL 60208, USA.
- f. Department of Organ Transplantation, Northwestern University, 2122 Sheridan Road, Suite 120, Evanston, IL 60208, USA.
- g. Department of Neurosurgery, Neurosurgery Research Institute, The First Affiliated Hospital, Fujian Medical University, Fuzhou 350005, Fujian, China.
- h. Institute of Neurology, First Affiliated Hospital, Fujian Medical University, Fuzhou 350005, Fujian, China.

#First author & co-first author

Yajun Jing & Hongliang Ge contributed equally to this work and are considered co-first authors.

Yajun Jing, e-mail: gzsmu15@sina.com

Hongliang Ge, e-mail: yangguang5716@163.com

Authors

Yiming Chen, e-mail: 1067119528@qq.com

Qiu He, e-mail: heqiu0404@163.com

Jinye Su, e-mail: sumile1998@gmail.com

Surmai Shukla, e-mail: surmai.shukla@northwestern.edu

✉Corresponding authors

Wenhua Fang, e-mail: fang.wenhua@fjmu.cn

Yuanxiang Lin, e-mail: lyx99070@163.com

Dezhi Kang, e-mail: kdz99988@vip.sina.com

Abstract

The best treatment strategies and prognosis factors for PB are still unknown because of its rarity. Thus, the aim of this study was to determine prognostic indicators and develop a model for predicting the prognosis in PB patients. First, we identified the prognostic factors for PB using data for 304 PB patients (≤ 30 years) from the Surveillance, Epidemiology, and End Results database. We then developed prediction models utilizing several machine learning approaches and evaluated the predictive power of various models using the concordance index, Brier score, and receiver operating characteristic curve. Finally, we established SHAP graphs for model interpretability and evaluated the clinical usefulness of the model by decision curve analysis (DCA). The present study revealed that age, year of diagnosis, laterality, tumor-size, stage, surg/rad sequence, and therapeutic approach were significant prognostic factors for PB patients ($p < 0.05$). The examination of model performance revealed that the random forest model has superior predictive power, displaying a

potent discriminative ability and accuracy, as well as great consistency. Importantly, the DCA highlighted the clinical usefulness of this model. In summary, the present study identified the prognostic factors for PB and established the most effective random forest model for predicting the survival of PB patients.

Keywords: Pineoblastoma; Machine Learning; Prognosis; Survival analysis



1. Introduction

The pineal gland is an endocrine gland located in the midline of the brain that indirectly modulates circadian rhythms via the secretion of melatonin. A group of primary tumors are termed pineal parenchymal tumors (PPTs), which are considered rare neoplasms and account for less than 1% of all primary central nervous system tumors [1]. In the 2016 classification of tumors of the central nervous system, PPTs were classified into four significant subgroups: pineocytoma (PC) (ICD-O 9361/1, WHO grade I), PPT of intermediate differentiation (PPTID) (ICD-O 9362/3, WHO grade II/III), pineoblastoma (PB) (ICD-O 9362/3, WHO grade IV), and papillary tumor (ICD-O 9395/3, WHO grade II/III) [2]. Notably, the prognosis, clinical features, and biology of these tumors were generally varied. Previous reports have shown that PB generally affects children (mean age of diagnosis: 13 years), and infant and young child cases are typically far more aggressive and challenging to treat; however, therapy is typically tolerated better by older children and teenagers with higher survival rates [3]. Conversely, PPTIDs, PCs, and PTPRs are more common in adults, peaking at ages 2040, 3060, and 30 years, respectively [4,5]. Studies have also shown that PB constitutes a substantial portion of PPTs, accounting for up to 25% to 50% of cases [6], and that PBs occur more frequently in children and adolescents than in adults [7-10].

PB is defined as a high-grade malignant primary supratentorial neuroectodermal tumor in the central nervous system [11]. Similar to other malignancies, patients with PB traditionally have an exceedingly poor prognosis because of the aggressiveness of PB. In addition, a high recurrence rate and strong tendency for metastasis, especially systemic metastasis, are among the main reasons why PB is associated with a poor prognosis. The published literature has reported that PB manifests a predilection for seeding along the craniospinal axis and metastasizing to parts other than the central nervous system [12,13], such as the lung [14], bone [15-18], and peritoneum [19]. However, data on the survival of patients with PB are extremely limited owing to the rarity of PB cases. To date, hundreds of studies have reported on PB, but most have been single case reports or small series [20-22]. Although those studies were a valuable source of information concerning PB at some level, reliable medical evidence for treatment decision making is still lacking due to inherent drawbacks of these studies, such as small sample sizes and inadequate statistical power. Therefore, to date, there are no clearly defined prognostic factors and generally accepted treatment regimens for patients with PB.

To overcome these limitations, we first identified the risk factors affecting PB prognosis based on a large series with 304 PB patients from the Surveillance, Epidemiology, and End Results (SEER) database. Next, we developed prediction models utilizing several machine learning approaches (such as random forest survival, survival tree, and Cox proportional hazard regression models) and evaluated each model's predictive power using the concordance index, Brier score, and area under the ROC curve (AUC). Ultimately, we selected the random forest model that had the best predictive power for predicting prognoses in patients with PB.

2. Methods

2.1. Data Source and Variable Selection

The data of 304 PB cases were acquired from the SEER database using SEER*Sat version 8.4.1 (<https://seer.cancer.gov/>). Comprehensive information, including age, sex, race, year of diagnosis, laterality, stage, therapy, surgical resection area, surg/rad sequence, and type of reporting source, was acquired from source documents to summarize the case. This is not necessarily the original document that identified the case but the source that provided the most complete information. Marital status, income (median household income inflation adj. to 2021), months from diagnosis to therapy, total number of in situ/malignant tumors (total number of in situ/malignant cancers reported by a patient based on the maximum sequence number of all in situ/malignant tumors in SEER up to the last published year of diagnosis; this value is equal to one for all tumors), survival months, vital status (alive or dead) and tumor size were included and analyzed in this study (See Supplementary Files for details). Several of these factors often had different subclasses, such as laterality (including right/left origin of primary and paired sites but no information on laterality) and therapy (consisting of biopsy only; surgical treatment; RT; chemotherapy (CT); RT thereafter). The operations consisted of the following: RT combined with CT, sequence unknown; RT after nonprimary surgery combined with CT, sequence unknown; surg/rad sequence (radiation after surgery, radiation before surgery, radiation combined both before and after surgery); resection (local tumor destruction or excision, partial surgical excision of the primary site, complete surgical removal of the primary site; operation is referred to as “debulking”); and stages (including localized, regional, and distant). Other information included rural-urban areas (countries in metropolitan areas with 1 million or more inhabitants,

countries in metropolitan areas with 250,000 to 1 million inhabitants, countries in metropolitan areas with fewer than 250,000 inhabitants, countries bordering the metropolitan area outside the metropolitan area, nonmetropolitan countries not bordering the metropolitan area), type of reporting source (inpatient/outpatient hospital, other) and tumor size. Survival time was defined as the period from diagnosis to death or last follow-up, and cases lost to follow-up were censored by the SEER database.

2.2. Inclusion and Exclusion Criteria

To meet the research requirements, we initially set up the inclusion and exclusion criteria for sample selection in this study. The inclusion criteria were as follows: (1) patients who were diagnosed with PB with the ICD-O-3 code 9362/3 and (2) patients who were diagnosed between 1975 and 2021. (3) Patients whose PB was younger than 30 years of age at diagnosis; (4) patients for whom the treatment information. Exclusion criteria for patients were as follows: (1) other types of cancer; (2) diagnosis before 1975 and after 2021; (3) older than 30 years of age; (4) unclear treatment information (unknown); (5) died from causes other than PB; and (6) availability of only a single case (e.g., RT before surgery, intraoperative beam radiotherapy (RT), RT before and after surgery, as well as sequences that were unknown, but both were given). Ultimately, 304 patients with PB were selected for this analysis (Figure 1).

2.3 Model Development

All of the cases in this study were randomly divided into training and test datasets at a 7:3 ratio. Prediction models, including the Cox proportional hazards model, survival tree model, and random forest survival model, were developed using the training dataset, and the accuracy of the finished model was externally validated using the test dataset.

2.3.1 Cox Proportional Hazards Model

The Cox proportional hazards model can be represented by the hazard function defined by $h(t)$, which can be expressed as Formula (1):

$$h(t) = h_0(t) \exp(\alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k) \quad (1)$$

where t represents the survival time, and $h_0(t)$ represents the baseline hazard in a situation where all the predictors are zero. The coefficients $(\alpha_1, \alpha_2, \dots, \alpha_k)$ were used to evaluate the effect size of the predictors $(X_1, X_2, \dots, X_k)$.

2.3.2 Survival Tree Model

To construct a survival tree (ST), observations were recursively partitioned into child nodes based on predictors that influence survival outcomes. Partitioning began with the root node containing all observations. At each parent node, every candidate predictor variable was evaluated to find the optimal split that maximizes the survival difference between the resulting child nodes. The predictor and split point leading to the best split was selected to divide the parent node. This recursive splitting process continued, treating child nodes as new parents to be split further. Tree growth stopped when all terminal nodes reached a minimum size, containing only a predefined number of unique events ("minsplit") to prevent overfitting. Tree pruning was also applied to reduce size and improve generalizability.

The log-rank statistic was utilized as the splitting criterion to maximize the dissimilarity between child nodes at each split, using Formula (2):

$$L_{X,C} = \frac{\sum_{i=1}^N \left(d_i, 1 - y_i, 1 \frac{d_i}{y_i} \right)}{\sqrt{i \sum_{i=1}^N \frac{y_i, 1}{y_i} \left(1 - \frac{y_i, 1}{y_i} \right) \left(\frac{y_i - d_i}{y_i - 1} \right) d_i}} \quad (2)$$

2.3.3 Random Forest Model

The RF model employed a nonparametric ensemble approach that introduces two forms of randomness when growing trees: bootstrapping training samples and selecting a subset of predictors for each split (known as "mtry"). During tree growth, a bootstrap sample containing two-thirds of the data (in-bag) was drawn at random B times. Based on these B samples, a node splitting process proceeded as follows: At each node, guided by a predefined splitting criterion ("split rule"), the predictor maximizing survival difference between child nodes was chosen from a random subset ("mtry") to split the node at an optimal cutoff point ("nsplit"). This iterative tree growth was repeated "ntree" times to build a forest

for final prediction. To estimate survival at time t for a new subject ($\hat{s}^{new}$) using the m th tree, $\hat{s}^{new}$ was dropped to a terminal node. The final prediction (Formula 3) averages survival estimates across all trees in the forest:

$$\hat{s}^{new}(t) = \frac{1}{M} \sum_{m=1}^M s_m^{new}(t) \quad (3)$$

2.4 Model Validation, Performance Evaluation, and Comparison

This study utilized the concordance index (C-index) to evaluate and compare the discriminative ability of different prediction models. We also examined the time-dependent C-index, which assesses each model's concordance at specific survival times. More specifically, for the i th patient, T_i denotes the event time, D_i denotes the censoring time, and η_i denotes the predicted risk score from a model. Additionally, $T_i^* = \min(T_i, D_i)$ is defined as the censored time or latest observed time, and $\xi_i = I(T_i < D_i)$ is defined as the event indicator for right censoring. The C-index then provides an estimate of the probability that the sequence of events is correctly predicted by the model for a randomly selected pair of cases (i, j).

In addition to the C-index, we utilized the Brier score (BS) to evaluate the predictive performance of the models. For binary prediction models, the Brier score calculates the mean squared error of predictions. For time-to-event models, the Brier score for a single subject at time t is defined as the squared difference between the observed event status (e.g., death) and the model's predicted probability of surviving beyond time t . The IBS represents a cumulative version of the Brier score over time, calculated as Formula (4). In other words, the IBS measures the accuracy of the predicted survival probabilities against the actual outcomes over the entire follow-up period.

$$IBS = \int_{t_1}^{t_{max}} BS^c(t) d\omega(t) \quad (4)$$

A receiver operating characteristic (ROC) curve depicts the trade-off between the true positive rate (TPR) and false-positive rate (FPR) as the discrimination threshold varies from 0 to 1. The TPR at threshold t , denoted as $TP(t)$, represents the probability that a randomly selected positive instance has a predicted score exceeding t . Assuming that the probability density function of negative instance scores is $fp(t)$, the FPR equals the probability that a randomly chosen negative instance scores above t . Varying t across the full range of $fp(t)$ traces out the ROC curve. The area under the ROC curve (AUC) equals the probability that a randomly selected positive example is assigned a higher score than a randomly selected negative example.

Sensitivity and specificity are relevant for binary classification tasks. A predictor's accuracy can be evaluated by the AUC as follows: Subjects are classified by comparing their polygenic risk scores against a fixed threshold, predicting those above it to manifest the trait. Sensitivity refers to the proportion of affected subjects correctly classified, whereas specificity is the proportion of unaffected subjects correctly classified. Varying the threshold generates different sensitivity and specificity values, tracing the ROC curve by plotting sensitivity against 1-specificity. The AUC represents the probability that a randomly chosen pair of subjects, one affected and one unaffected, are correctly classified. Assuming the normality of polygenic scores, the expected AUC can be derived using central limit theorem principles, which can be calculated using Formula 5.

$$AUC(t) = \int_{-\infty}^{\infty} Se(c, t) d[1 - Sp(c, t)] \quad (5)$$

Validations were conducted internally (in training datasets) and externally (in test datasets) and visualized in the calibration plots. Predictive models for binary outcome Y (1 for event present, 0 for event absent) are considered well calibrated when estimated event risks align with observed event rates. Specifically, for patients with estimated risk $R=p$, the proportion experiencing the event should equal p : $P(Y=1|R=p) = p$. Calibration can be evaluated by regressing Y on the model's linear predictor (LP) $\text{logit}(R)$: $\text{logit}[P(Y)] = \alpha + \beta \times LP$. A perfect calibration implies that $\alpha=0$ and $\beta=1$. Miscalibration frequently occurs when models are applied to new populations, owing to case-mix differences or variable effects. Four primary types of miscalibration exist. If $\beta < 1$, risks are overfit; small risks are underestimated, whereas large risks are overestimated. $\beta > 1$ indicates underfitting. If $\alpha \neq 0$ but $\beta=1$, risks are systematically overestimated ($\alpha < 0$) or underestimated ($\alpha > 0$). More generally, over/underestimation occurs when $\alpha(\beta=1) \neq 0$, as assessed by modeling LP as an

offset. In simulations controlling Y and predictor distributions, miscalibration can be induced by varying $\alpha(\beta=1)$ and β .

A critical role of prediction models is to inform patients on their prognosis, for example, after a cancer diagnosis has been made or when assistant clinicians make a decision. Thus, the usefulness of the prediction model is rooted in the purpose of the prediction model. In this study, decision curve analysis was employed to assess the usefulness of the model. The critical element of decision curve analysis is that a single probability threshold can serve dual purposes—categorizing patients as positive or negative and weighting the harm of false-positive versus false-negative classifications. The selected risk threshold implies relative misclassification costs. The threshold odds equal the harm-to-benefit ratio—the harm of a false-positive divided by the benefit of a true positive. For a 20% threshold, the odds are 1:4, meaning that the harm of a false positive is one quarter the benefit of a true positive. This suggests accepting 4 false positives for 1 true positive, which a clinician may articulate as "I would do no more than 5 biopsies to find one cancer." Thus, when a model is applied, we can correct true positives (TPs) for false positives (FPs) using the threshold odds (w): $TP - FP \times w$. Dividing by the total sample size N yields the net benefit—the net proportion of TPs.

Net benefit compares a model to the default strategies of treating none (no one has the event) or all (all have the event). Unlike cost-effectiveness analysis, decision-analytic net benefit does not require monetary value. Varying the risk threshold plot model and default net benefit by threshold as a decision curve. A lower model net benefit indicates clinical harm at that harm-benefit ratio. Treat-none net benefit is always 0; treat-all is positive below and negative above the event rate. The net benefit difference between two models at a given threshold reflects their difference in true positive proportions at identical false-positive levels.

2.5 Model Hyperparameter Optimization

The performance quality of a machine learning model depends heavily on its hyperparameter settings. All machine learning systems have hyperparameters, the values for which must be adjusted to optimize performance. This process, known as hyperparameter optimization (HPO), is particularly important for machine learning models, which rely heavily on architectural, regularization, and optimization decisions. Notably, in HPO, the optimal combination of hyperparameters is selected to maximize model performance. Unlike model parameters that are learned from data, hyperparameters are specified a priori and configure the learning algorithm itself. Proper tuning of hyperparameters can significantly improve the accuracy, robustness, and generalizability of the model. By finding the right balance between underfitting and overfitting, HPO enables models to better capture the underlying relationships in data and achieve stronger performance on new unseen data. In this study, HPO was used to optimize the random forest model.

2.6 Explanation of Models

Explaining machine learning (ML) models is crucial to understanding how they make predictions. However, there are many explainability techniques, each of which uses different approaches. Broadly speaking, model explanation involves the analysis of ML models, particularly "black boxes" with opaque internal logic. A key technique is feature importance (FI), which ranks input features according to their influence on the output. This provides a comprehensive explanation of the key drivers of the model predictions. In contrast, local methods such as SHapley Additive exPlanations (SHAPs) explain individual predictions. SHAP uses game theory to map the output of a prediction to the contributions of each input feature. It can quantify the impact of features locally on a specific prediction or globally across the entire model. Here, we applied both FI for global explanation and SHAP for local understanding in explaining a random forest model.

2.7 Software

The continuous variables in our study (including age, years of diagnosis, and tumor size) were categorized based on optimal cutoff values identified using X-tile software. The optimal cutoff value was defined as the most significant splitting point among the survival distributions of the categorized factors. The newly categorized variables were considered predictors and analyzed in the survival analysis. Statistical analysis in this study was performed using R (version 3.6.3) and X-tile software (Yale University, version 3.61). The criterion for excluding a variable from the multivariate Cox proportional hazards model was a p value > 0.1 . For all other tests, a p value < 0.05 indicated statistical significance.

3. Results

3.1. Clinical and Pathological Characteristics of Patients

Ultimately, 304 individuals with PBs younger than 30 years of age were selected for the present study. A total of 212 people were in the training cohort, and 92 patients were in the test cohort. Table 1 shows the baseline characteristics of the patients. In this study, the overall ratio of men to women was approximately 1:1. With a mean age of 13 years, most patients were non-Hispanic ($n = 165$, 77.8% in the training dataset and $n = 68$, 73.9% in the test dataset). In addition, a significant proportion of patients had tumor extension features confined and localized to the gland of origin not otherwise specified ($n = 134$, 63.2% in the training dataset and $n = 64$, 69.6% in the test dataset) and unpaired laterality features ($n = 206$, 97.2% in the training dataset and $n = 91$, 98.9% in the test dataset). In this work, approximately 80% of patients ($n = 166$, 78.3% in the training dataset and $n = 76$, 82.6% in the test dataset) were diagnosed between 2001 and 2021. Importantly, up to 97% of patients had positive histology and a confirmed diagnosis. Additionally, the vast majority of cases ($n = 209$, 98.6% in the training dataset; $n = 91$, 98.8% in the test dataset) were recorded as either hospital inpatients or outpatients. The vast majority of patients were unmarried ($n = 194$, 91.5% in the training dataset; $n = 90$, 97.83% in the test dataset). Regarding treatment options, patients either underwent isolated surgery or additional radiation (RT) or chemotherapy (CT) after surgical resection. Notably, over two-thirds of patients were from countries with cities with a population of one million or less.

Table 1. Clinical and pathological characteristics of patients

Variables	Test Dataset	Training Dataset	Standardize diff.	P value	P value*
N	92	212			
Age	13.0 ± 9.1	13.4 ± 9.7	0.1 (-0.2, 0.3)	0.681	0.872
Tumor size	32.1 ± 11.5	31.7 ± 11.2	0.0 (-0.2, 0.3)	0.748	0.950
Income[#]	6.1 ± 2.1	6.1 ± 2.1	0.0 (-0.2, 0.2)	0.977	0.897
TNM[#]	1.1 ± 0.2	1.1 ± 0.3	0.0 (-0.2, 0.3)	0.758	0.813
Months[#]	0.4 ± 0.7	0.5 ± 0.8	0.1 (-0.1, 0.4)	0.256	0.219
Race			0.1 (-0.2, 0.3)	0.458	-
Hispanic	24 (26.1%)	47 (22.2%)			
Non-Hispanic	68 (73.9%)	165 (77.8%)			
Sex			0.0 (-0.2, 0.3)	0.745	-
Male	47 (51.1%)	104 (49.1%)			
Female	45 (48.9%)	108 (50.9%)			
Year			0.1 (-0.1, 0.4)	0.392	-
1978-2000	16 (17.4%)	46 (21.7%)			
2001-2021	76 (82.6%)	166 (78.3%)			
Laterality			0.1 (-0.1, 0.4)	0.352	0.679
Not paired site	91 (98.9%)	206 (97.2%)			
Paired site	1 (1.1%)	6 (2.8%)			
Stages			0.1 (-0.1, 0.4)	0.560	-
Localized	64 (69.6%)	134 (63.2%)			
Regional	15 (16.3%)	43 (20.3%)			
Distant	13 (14.1%)	35 (16.5%)			
Resection[#]			0.3 (0.0, 0.5)	0.521	-
Autopsy	16 (17.4%)	46 (21.7%)			
LTD	26 (28.3%)	56 (26.4%)			
PSR	27 (29.3%)	49 (23.1%)			
TSR	6 (6.5%)	25 (11.8%)			
Debulking	8 (8.7%)	11 (5.2%)			
Radical surgery	3 (3.3%)	6 (2.8%)			
Surgery, NOS	6 (6.5%)	19 (9.0%)			
Therapy[#]			0.2 (-0.0, 0.5)	0.877	-
Autopsy only	2 (2.2%)	9 (4.2%)			
Surgery	8 (8.7%)	20 (9.4%)			
Radiation	5 (5.4%)	10 (4.7%)			
Chemotherapy	1 (1.1%)	4 (1.9%)			
SR	20 (21.7%)	37 (17.5%)			
SC	14 (15.2%)	24 (11.3%)			

<i>RC</i>	9 (9.8%)	27 (12.7%)			
<i>SRC</i>	33 (35.9%)	81 (38.2%)			
<i>Sequence[#]</i>			1.6 (1.3, 1.9)	<0.001	-
<i>Chemo after Sur</i>	39 (42.4%)	2 (0.9%)			
<i>Rad after Sur</i>	50 (54.3%)	111 (52.4%)			
<i>RC</i>	3 (3.3%)	99 (46.7%)			
<i>Source[#]</i>			0.0 (-0.2, 0.3)	0.818	1.000
<i>Hospital</i>	91 (98.9%)	209 (98.6%)			
<i>Others</i>	1 (1.1%)	3 (1.4%)			
<i>Material status</i>			0.3 (0.2, 0.7)	0.304	-
<i>Married</i>	2 (2.17%)	18 (8.5%)			
<i>Unmarried</i>	90 (97.83%)	194 (91.5%)			
<i>Rural-Urban[#]</i>			0.3 (0.0, 0.5)	0.301	-
<i>1</i>	63 (68.5%)	137 (64.6%)			
<i>2</i>	12 (13.0%)	42 (19.8%)			
<i>3</i>	5 (5.4%)	16 (7.5%)			
<i>4</i>	7 (7.6%)	7 (3.3%)			
<i>5</i>	5 (5.4%)	10 (4.7%)			

Note:

Results in the table: Mean+SD/N (%); P value*: Kruskal-Wallis's rank test for continuous variables, Fisher's exact test for categorical variables with expectations<10.

Income#: Median household income inflation adj. to 2021

TNM#: Total number of in situ/malignant

Sequence#: Surg/Rad Sequence

Chemo after Sur, chemotherapy after surgery; Rad after Sur, radiotherapy after surgery; RC, radiotherapy combined with chemotherapy

Source#: Type of Reporting Source

Therapy#:

SR, radiotherapy after surgery; SC, chemotherapy after surgery; RC, radiotherapy combined with chemotherapy; SRC, radiotherapy after surgery combined with chemotherapy

Resection#: Site-specific surgery

LTD, local tumor destruction or excision; PSR, partial surgical removal of the primary site; TSR, total surgical removal of the primary site; "debulking", surgery stated to be "debulking".

Rural-Urban#: Rural-Urban Continuum Code

1, Countries in metropolitan areas of 1 million pops; 2, Countries in metropolitan areas of 250,000 to 1 million pops; 3, Countries in metropolitan areas of lt 250 thousand pops; 4, Nonmetropolitan countries adjacent to metropolitan areas; 5, Nonmetropolitan countries not adjacent to metropolitan areas.

3.2. Survival Analysis and Identification of Potential Prognostic Factors

To identify variables that potentially affect the survival of PB patients, we first performed survival analysis on the variables included in this study. In the present study, age, year of diagnosis, laterality, tumor size, stage, surg/rad sequence, and therapeutic approach were found to have a significant effect on the survival of PB patients ($p<0.05$) (Figure S1 and S2). In addition, the survival curves indicated that the total number of in situ/malignant and rural-urban cases affected patient survival, even if the p value was not significant ($p<0.2$) (Figure S2). Notably, the results of the survival analysis revealed that other characteristics, including race, sex, resection, months between diagnosis and therapy, type of reporting source, income, and marital status, had no significant effect on PB survival ($p>0.2$) (Figure S1 and S2). Given that univariate analysis findings often cannot completely rule out the impact of confounding variables, we chose variables with p values less than 0.2 [23] to be included in the prediction model.

3.3. Establishment and Comparison of Models

Various prediction models for survival in this study were developed using the Cox proportional hazards model, tree-based machine learning algorithms (survival tree) and random forest. Nevertheless, the predictive power of various models varies widely. Notably, machine learning (ML) algorithms identify the best model by use of automated procedures that analyze the data to find patterns, including interactions between variables and within variable subsets

[24]. Additionally, machine learning (ML) is more accurate than the Cox proportional hazards model is [25-27]. As a result, we assessed the metrics of several prediction models using AUC values, the C-index, and Brier scores (BS) in this work. Interestingly, our findings, as illustrated in Figure 2A–C, revealed that the random forest survival model presented a greater AUC value (AUC values of 0.98, 0.99, and 0.96 for 3-, 5-, and 10-year survival in the training dataset, respectively) than both the Cox regression model (AUC values of 0.81, 0.79, and 0.75 for 3-, 5-, and 10-year survival in the training dataset, respectively) and the survival tree model (AUC values of 0.81, 0.79, and 0.75 for 3-, 5-, and 10-year survival in the training dataset, respectively), indicating that the random forest survival model had a higher predictive power than the other models. More importantly, we achieved consistent results in the test dataset, as shown in Figure 2D–F. Additionally, we discovered that the random forest survival model had the highest C index in both the training and test datasets, which further demonstrated the superior predictive power of the random forest model (Figure 2G–H). Finally, through a comparative analysis of the three models' Brier scores, we discovered that, in both the training and test datasets, the random forest survival model displayed the lowest Brier scores (Figure 2I). This further supports the idea that the random forest survival model performs better in predictions than other models. As mentioned above, since it has greater predictive potential than the other models, we will eventually select the random forest technique to create a predictive model for PB.

3.4. Random Forest Survival Model Optimization and Specification

The hyperparameters of a machine learning model significantly impact its performance quality [28]. The machine learning (ML) model and its hyperparameters are usually manually assigned a goal and a dataset. This work thus applies hyperparameter optimization to optimized machine learning models. As shown in Figure S3, the model returned the minimum out-of-bag (OOB) error values when the random forest survival model met the following parameters: number of trees (ntree) = 200, number of variables tested in any split (mtry) = 8, and minimum size of terminal node (nodesize) = 3. Therefore, the above parameters are used to define the optimal random forest model.

3.5. Random Forest Survival Model Performance

3.5.1 Accuracy of Random Forest Survival Model

Accurate predictions effectively discriminate between individuals with and without the outcome of interest. Currently, several quantitative measures exist to evaluate the performance of binary classification models. The receiver operating characteristic (ROC) curve has emerged as the predominant metric used to quantify the discriminative ability of predictive models. Notably, several accuracy indices, including the AUC, partial area, and TPF/FPF, have been proposed in the literature to concisely summarize the salient features of ROC curves [29-31], in which the area under the curve (AUC), as a one-dimensional summary metric, is frequently used for assessing the predictive efficacy of a model [32]. In the present study, the accuracy of the random forest survival model was assessed using the Brier scores and time-dependent receiver operating characteristic (ROC) curves. As shown in Figure 3A and 3B, the model consistently demonstrated robust performance in predicting PB-specific survival, with time-dependent AUC values ranging from 0.70 to 0.95 and BS values ranging from 0.01 to 0.08. Additionally, to measure the global importance of variables, we use permutational variable importance. We interpret the result as follows: the greater the decrease in performance of the model after permuting a variable, the more important the variable is for making the prediction; therefore, the variables higher on the plot are the most important ones. As shown in Figure 3C, age was the most important factor affecting the survival of PB patients. More importantly, we found that the importance of age for PB prognosis was much greater than that of other variables. Finally, we use a time-dependent metric (cumulative/dynamic AUC, C/D AUC) [33,34] to determine whether some variables are more important at certain times of the study. Our results confirmed again that age was the most important predictor (Figure 3D). Notably, our findings were in line with those of previous studies [35-37]. These results suggest that the random forest survival model performs exceptionally well in terms of accuracy.

3.5.2 Discrimination Power of Random Forest Survival Model

In addition to the AUC, calibration is a vital metric for validating the accuracy of predictive models. Calibration gauges how closely a model's predicted probabilities align, on average, with actual observed outcomes. Calibration plots depicting observed versus expected event rates across risk strata can localize potential deviations. The calibration curve visualizes model performance across subsets, complementing overall accuracy. Well-calibrated models exhibit minimal deviation from the 45° perfect fit line, whereas miscalibration manifests as a growing divergence of expected probabilities from observed probabilities. In this study, the random forest survival model developed for training was

evaluated using calibration plots, as shown in Figure 4, where the calibration plots for the training and test datasets exhibited high consistency.

3.5.3 Clinical Usefulness of Random Forest Survival Model

Developed by Vickers and Elkin in a 2006 paper in Medical Decision Making [38], decision curve analysis is an approach for assessing prediction models and diagnostic tests. This strategy aims to overcome the limitations of both traditional statistical measures and comprehensive decision analytic tools, such as calibration and discrimination, which are not always indicators of therapeutic usefulness. Therefore, the usefulness of the random forest survival model in this study was assessed using decision curve analysis. As shown in Figure 5, the black line represents the net benefit if all patients died, and the yellow line represents the net benefit if no patients died. If the blue line, which represents the model, falls between the black and yellow lines, it is a sign of the model's clinical usefulness. A greater net benefit, or the model curve's distance from the black and yellow lines, indicates better clinical usefulness of the random forest survival model. Notably, almost all of the threshold probabilities at 36 months (Figure 5A), 60 months (Figure 5B), and 120 months (Figure 5C) showed a significant net gain in the random forest survival model for the DCA curve analysis in the training cohort, indicating the potential clinical usefulness of this model. Importantly, the DCA results for the test cohort demonstrated a substantial net improvement in the random forest survival model for almost all of the threshold probabilities at 36 months (Figure 5D), 60 months (Figure 5E), and 120 months (Figure 5F), further suggesting the model's potential for clinical application. In summary, these results emphasize the potential clinical utility of this model.

3.6 Explanation of Random Forest Model Survival

Recent studies have reported that more advanced machine learning (ML) models, including random forest survival, gradient boosting machines (GBMs), and artificial neural networks (ANNs), typically demonstrate greater accuracy than simpler models when predicting nonlinear, faint, or rare phenomena [39,40]. However, these complex algorithms frequently achieve superior performance at the expense of interpretability. Model interpretability remains crucial for practical application. In this study, we utilize established explanatory techniques to elucidate the model we have developed. Specifically, we provide both global and local explanations to facilitate understanding of the overall model and its individual components. Among the nine 9 variables found in the random forest survival model, the top 7 variable importance values (VIMPs) are displayed in Figure 6A and B. Remarkably, age seemed to have the highest predictive significance (VIMP=0.251). In addition, tumor stage (VIMP=0.054) for all 9 variables had the second largest importance in terms of prediction. These findings suggest that different therapy regimens, especially preoperative or postoperative chemoradiation, are important predictors of prognosis in children, adolescents, and young adults. The third most important predictor in this model was surg/rad sequence (VIMP=0.052). Notably, the individual predictions in this investigation were subjected to local explanations. The important ranks of the predictors for a random case are displayed, and the patient's overall survival is estimated, as shown in Figure 6C and D.

4. Discussion

It is of great importance for rare disorders such as PB to identify the prognostic factors affecting the survival of patients and evaluate the effectiveness of current treatment regimens. Thus, the acquisition of abundant clinical materials from PB patients is critical for achieving this goal. However, conducting prospective clinical trials seems to be an exceedingly challenging issue owing to the limited number of PB cases. Therefore, to date, the optimal management strategies for PB confined by extremely limited data have not been defined. Furthermore, the clinical risk factors associated with prognosis were still less clear in groups with PB, especially in children, adolescents, and young adults. Given the greater precision of the advanced machine learning models developed previously than the classical staging system for prognosis in some cancers [41,42], a random forest survival model for PB, consequently, was developed on the basis of enormous amounts of clinical PB data collected from the SEER database from 1975 to 2021 in this study. The aim of this study was to clarify the ideal treatment regimen and identify crucial prognostic factors for PB. Notably, to the best of our knowledge, this study represents the largest series investigation of PB in children, adolescents, and young adults.

Our findings verified that the primary prognostic indicators for PB patients in children, adolescents, and young adults were age at diagnosis, year of diagnosis, surg/rad sequence, therapy choices, laterality, tumor stage, and tumor size. Interestingly, the results of this analysis indicated that the overall survival rate of patients with PB has little to do with the application of RT or CT alone. However, the combination of RT and CT was found to be a prognostic factor for survival.

Multiple lines of evidence illustrate that PB, compared with infratentorial primitive neuroectodermal tumors, has poorer outcomes irrespective of treatment regimens [43-45]. However, aggressive surgical resection is still recommended for the management of PB, and surgical treatment is regarded as the first-line therapeutic strategy [11,22]. In this analysis, the effects of various treatment regimens on the prognosis of patients with PB was accessed by the random forest model, which revealed that those treatment options had different influences on clinical outcomes. Importantly, the risk of death was considerably lower in patients who received RT after surgery combined with CT than in patients who received only RT or surgical resection. Although surgical resection is known as the first-line therapeutic option and is strongly recommended [46], the results of this study could not support that surgical treatment alone could significantly improve the prognosis of patients with PB in young populations. Additionally, the risk of death significantly increased after partial surgical removal of the primary site in terms of the impact of surgical resection on survival. Interestingly, the results from a survey conducted by Mallick and colleagues revealed that the clinical outcome was markedly improved in patients who received surgical resection and adjuvant radiotherapy [47]. In line with these findings, the findings of this study revealed that RT after surgery combined with CT was a principal prognostic factor that could remarkably improve clinical survival outcomes.

The role of CT in PB is still a contentious issue. PB appears to be responsive to CT, as reported by Schild et al. [48] and Ghim et al. [49]. Interestingly, the findings of this study revealed that CT could improve the overall survival of patients. Interestingly, a study conducted by Mynarek et al. [7] drew conclusions that were inconsistent with our findings and could not support the effectiveness of CT in PB. Similarly, there is also no consensus regarding the role of RT against PB. A study by Clark and colleagues [50] demonstrated that PB was resistant to RT. However, several studies reported that PB was sensitive to radiotherapy [44,51,52]. Notably, our results revealed that radiotherapy alone did not effectively improve the survival of PB patients. Additionally, the dose of RT and selection of local radiotherapy (locRT) or craniospinal irradiation (CSI) are controversial topics in the management of PB. In view of the severe side effects of CSI in young patients, whether to choose RT typically depends on patient age [7]. Nonetheless, "reduced" CSI with a dose of 23.4/24.0 Gy was also used during first-line therapy in patients <4 years of age [6]. In addition, the application of CSI at a dose of 31.2 Gy and a dose range of 45.0–59.4 Gy has been reported in the literature [6]. Notably, those studies failed to conclude that RT is effective for PB. Interestingly, the findings of this analysis could not determine that patients who only received RT benefited from RT treatment. The limited number of PB patients may be the primary reason for the discrepancy in the effectiveness of RT and CT. Hence, high-quality clinical studies are crucial to clarify these disagreements between RT and CT.

Notably, age is typically regarded as a risk factor for various disorders, such as Parkinson's disease [53], venous thromboembolism [54], diabetes [55], cardiovascular disease [56], stroke [57], and all types of tumors [58-60]. Interestingly, the present study revealed that age was the primary risk factor for PB. Consistent with findings of previous research [61-63], our results showed that overall survival increased markedly with increasing age. Specifically, compared with patients aged <5 years at diagnosis, the risk of death was significantly lower for patients aged 5-30 years at diagnosis. The subgroup analysis revealed that patients older than five years of age had a significantly greater overall survival probability than patients younger than five years of age. However, patients older than twelve years of age and within the age range of 5 years presented a much smaller overall survival probability improvement than those younger than five years of age and within the age range of 5 years. The biological behavior of PB may vary with increasing age, but many studies have been conducted to confirm this conclusion. In addition to age, the results of survival analysis indicated that the year of diagnosis was a prognostic factor that influenced the survival of PB patients. Specifically, we found that patients diagnosed before 1998 had significantly lower overall survival than those diagnosed after 1998. Our data clearly reflect lead-time bias and length bias, particularly for PB. Additionally, the differential effect of screening over time due to variation in screening rates is a confounding variable when analyzing temporal changes in survival rates in diverse racial or ethnic groups. We recognized that changes in survival over time might be confounded by the differential effects of various early interventions, and our analyses included the year of diagnosis in the multivariate model and adjusted for its effect on survival. This differential effect may be partially due to a change in the classification of deaths due to cancer. Mortality reflects a mixture of cancer diagnoses made at various points in the past, with differences in available diagnostic practices, treatment options, and levels of supportive care. Much of the overall mortality decline has been due to those cancers for which interventions have been introduced into the general population.

Aggressiveness, the major characteristic of malignancy, is the chief cause of poor prognosis in most cancer patients. The results of this study revealed, almost without exception, that the overriding risk factor for PB was the continuous extension of disease, which was in line with results of other studies [64,65]. In particular, the random forest survival model revealed that overall survival for groups whose tumors had spread to distant organs was significantly lower than that for groups whose tumors were localized to the gland of origin; however, there was no significant difference in overall survival between groups with tumors that had spread to regional organs and those that had tumors localized to the gland of origin. Notably, tumor size is another crucial prognostic marker for various tumors. Currently, there is much debate concerning the relationship between tumor size and prognosis. The results of a study conducted by Tirada and colleagues revealed that increased tumor size often portended a poor prognosis for patients with cancers [66]. Consistent with this, a study from Dittmar et al. showed that one of the high-risk factors for mortality was a larger tumor size [67]. In addition, it has been reported that the risk of death and recurrence increases with increasing tumor size [68]. However, there is evidence that patients with tumors tend to have poorer survival when the tumor size is smaller [69,70]. In contrast, the results of this analysis revealed that a larger tumor size was the main prognostic factor for poor survival.

Importantly, exploring the molecular mechanism of PB is highly important for its management. However, due to the limited number of PB cases, reports regarding PB at the molecular level are fewer than case reports and clinical studies. A genome-wide DNA methylation study by Capper et al. revealed that PB exhibited distinctive molecular features despite its histologic resemblance to other central nervous system embryonal tumors [71]. In particular, several parallel studies have shown that PB typically presents various distinct molecular subgroups [72-74]. A series reported by Anthony et al. indicated that patients in the PB-miRNA1 or PB-miRNA2 groups were more commonly older children with a median age at diagnosis of 8.5 years or young adolescents with a median age at diagnosis of 11.8 years, whereas PB-RB1 was found among younger children with a median age at diagnosis of 1.4 years [75]. Interestingly, tumor behavior differed significantly among the molecular groups; specifically, patients in the PB-miRNA1 and PB-miRNA2 groups often presented with localized disease. However, PB-RB1 was more common in patients with metastatic disease at diagnosis. Similarly, Elke et al. reported that PB presented distinct subtypes with different subgroups, such as PB-Grp1A, PB-Grp1B, and PB-Grp2, as well as Pin-RB1 and PB-MYC, two novel molecular groups [76]. In line with the results of Anthony et al., distinct tumor behaviors and survival differed significantly among various subtypes. In summary, it is critical for the management of rare tumors to explore molecular subtypes or novel clinical classifications combined with molecular subtypes.

There was no doubt that this study has some limitations. First, the long span of a few decades of patient data collection led to heterogeneity in examination, diagnosis, and treatment modalities. Additionally, several crucial data points related to patient prognosis, such as the resection technique, chemotherapy regimen, and radiotherapy dose, were unavailable. Third, the surg/rad sequence was unknown, which might lead to an unrealized evaluation of the effect of the sequence on patient prognosis. Despite the shortcomings of this study, its advantages are also considerable and conspicuous. Compared with previous studies, more information about the prognostic factors of PB was revealed in this study. Additionally, inherent bias and heterogeneity were limited through regression analysis, and the statistical power was improved significantly because of the increased number of PB patients. More importantly, the risk factors in the model were more accessible to achieve than those realized by a high-priced technique and a great deal of time.

5. Conclusions

In summary, age, surg/rad sequence, stage, treatment approach, laterality, and tumor size were identified as major prognostic factors of PB in this study. Additionally, we developed and validated a prognostic model that presented an accurate and nonsubjective prediction for the prognosis of patients with PB. Most importantly, the model may help clinicians assess patient prognosis and therapy decisions.

Declaration

Ethics declarations

The SEER Research Data Agreement to protect the privacy of the patients was signed by the authors, which is consistent with ethical principles.

Availability of data and materials

These data can be found in the SEER database (<https://seer.cancer.gov/data/>) and we cited publicly released codebase for reproducibility of the reported results (see: https://github.com/dumizai/SEER_OPCs_survival_prediction_code_data).

Author contributions

All authors contributed to the study conception and design. Y.J.J., D.Z.K., and Y.X.L. contributed to Conceptualization. Y.J.J., J.Y.S. and H.L.G., to the Methodology. Y.J.J., H.L.G. to the Software. J.Y.S. and Q.H. to the Validation. Y.J.J. and Q.H. to the Formal Analysis. J.Y.S. and Y.M.C. to the Investigation. D.Z.K. and W.H.F. to the Resources. Y.M.C. and S.S. to the Data Curation. Y.J.J. and J.Y.S. to the Writing – Original Draft. J.Y.S., W.H.F., and S.S. to the Writing – Review & Editing. J.Y.S. and Q.H. to the Visualization. D.Z.K. and W.H.F. to the Supervision. D.Z.K. and Y.X.L. to the Project Administration. D.Z.K., Hongliang Ge, Wenhua Fang, and Y.X.L. to the Funding Acquisition. All authors reviewed the manuscript.

Acknowledgments

Not Applicable

Funding

This study was generously supported by the National Natural Science Foundation of China (No. 82171327), the Technology Platform Construction Project of Fujian Province (No.2021Y2001) and Leading Project Foundation of Science and Technology, Fujian Province (No.2021Y0013).

Competing Interest

All authors report no conflicts of interest in this study.

Consent for publication

Not applicable.

Abbreviations

PB: pineoblastoma

C-index: concordance index

ROC: receiver operating characteristic

DCA: decision curve analysis

AUC: area under the curve

PPT: pineal parenchymal tumor

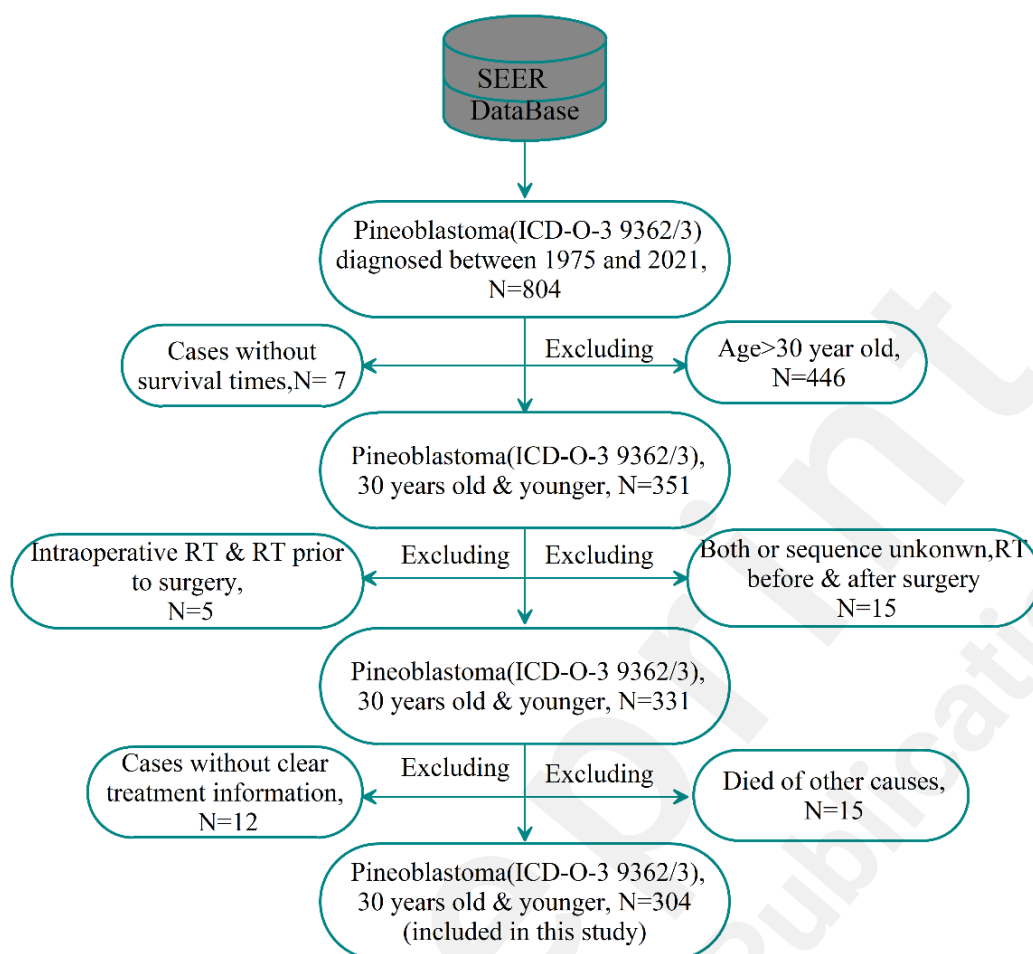
SEER database: Surveillance, Epidemiology, and End Results database

ICD-O-3: International Classification of Diseases for Oncology, third Edition; WHO, World Health Organization

CT: chemotherapy

locRT: local radiotherapy

CSI: craniospinal irradiation

Figure legends**Figure1. Flow chart for enrollment.**

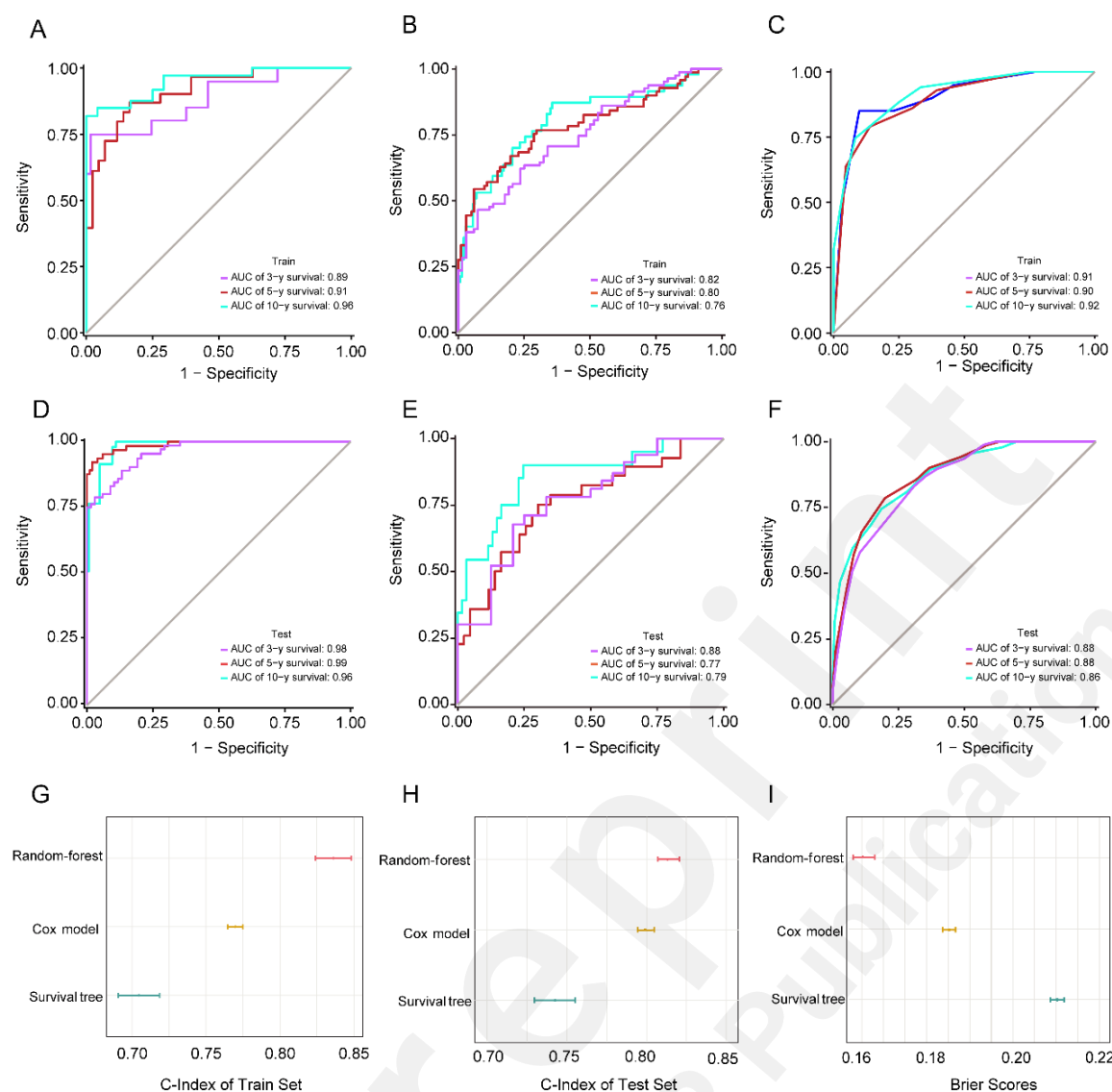


Figure 2. Comparison of the predictive power of the Cox regression, survival tree, and random forest survival models.

(A-C) Time-dependent ROC curves showing the sensitivity and specificity of the cancer-specific survival prediction by the random forest survival, Cox regression, and survival tree models in the training dataset. (D-F) Time-dependent ROC curves showing the sensitivity and specificity of the cancer-specific survival prediction by the random forest survival, Cox regression, and survival tree models in the test dataset. (G & H) Dot-line plot of the C-index values of the Cox regression, survival tree, and random forest survival models. (I) Dot-line plot of the Brier scores from the Cox regression, survival tree, and random forest survival models.

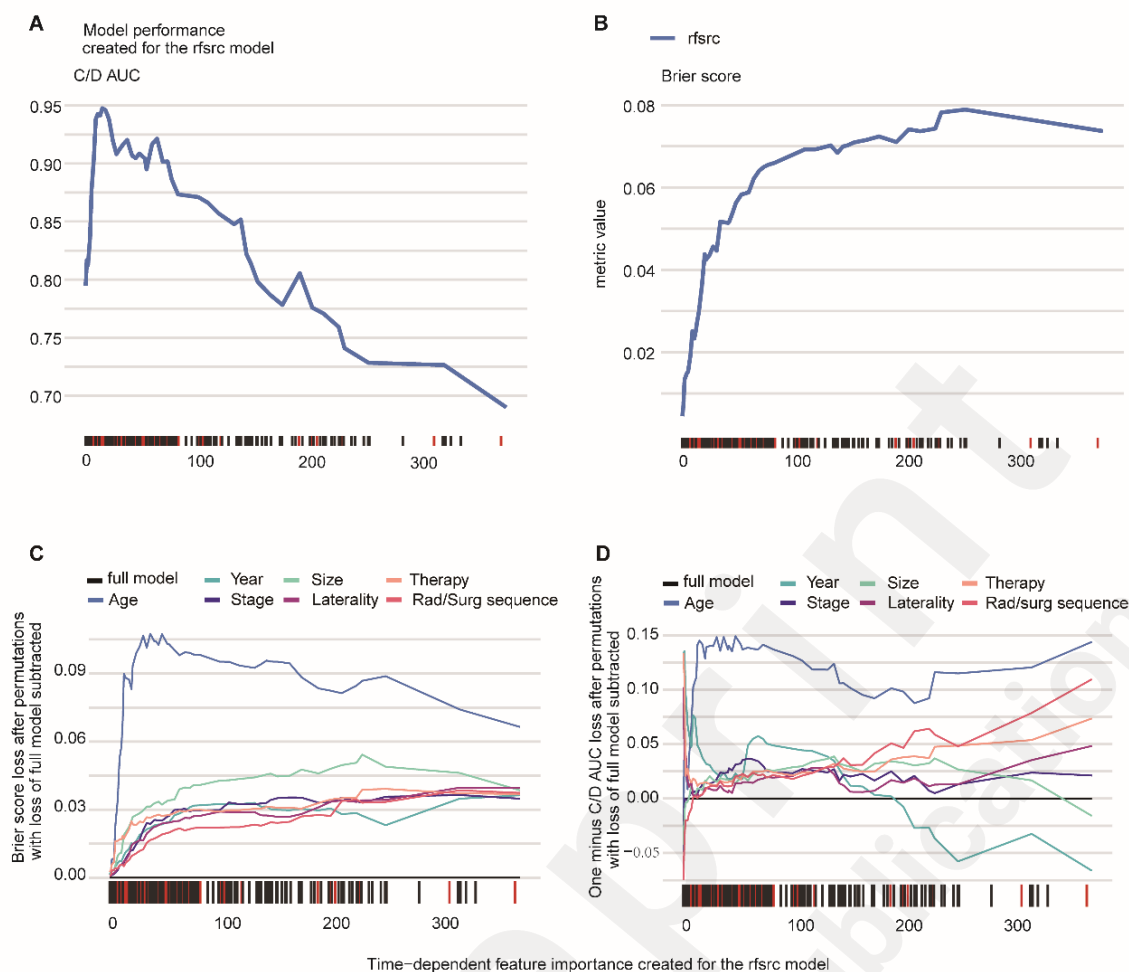


Figure 3. Evaluation of the prediction accuracy of the optimized random forest survival model.

(A) Time-dependent receiver operating characteristic (ROC) curve for the optimized random forest survival model. (B) Time-dependent Brier score curves. (C) Time-dependent Brier score curves for Brier score loss after permutations with subtraction of the full model loss in the optimized random forest survival model. (D) Time-dependent receiver operating characteristic (ROC) curve for one minus C/D AUC loss after permutations with subtraction of the full model loss in the optimized random forest survival model.

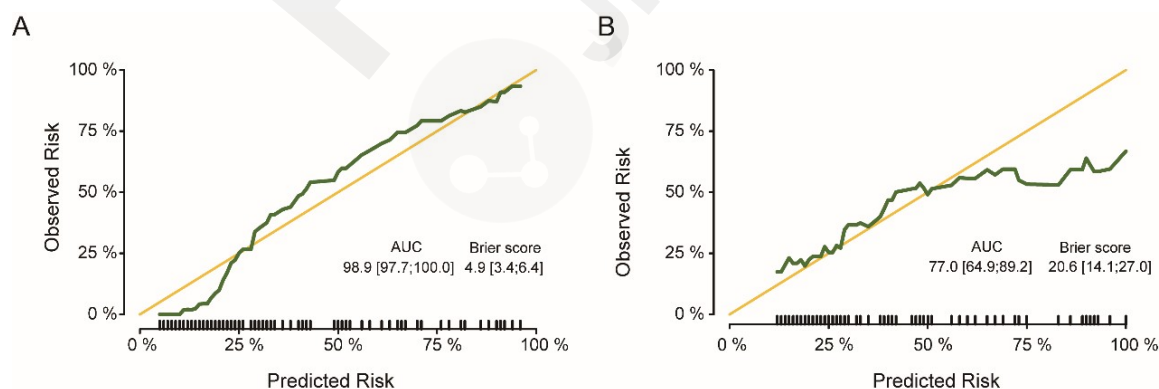


Figure 4. Discrimination power of the random forest survival model.

Calibration plots for the training (A) and test (B) datasets, depicting the concordance between the predicted and observed outcomes and validating the reliability of the random forest survival model for both datasets.

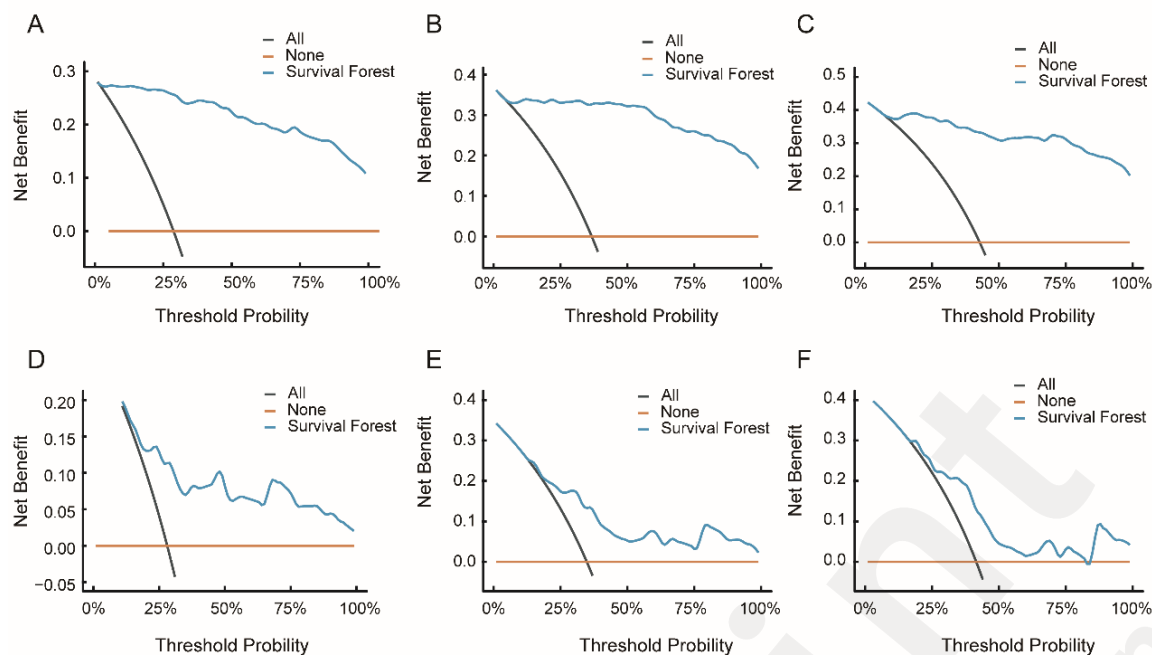


Figure 5. Clinical utility of the optimized random forest survival model.

Decision curves of the optimized random forest survival model for predicting 3- (A), 5- (B), and 10-year (C) cancer-specific survival in the training dataset; decision curves of the random forest survival model for predicting 3- (D), 5- (E), and 10-year (F) cancer-specific survival in the test dataset. The black line assumes that all patients experienced cancer-specific death, whereas the yellow line on the plot assumes that none of the patients experienced cancer-specific death. The blue line represents the net benefit of using the random forest survival model.

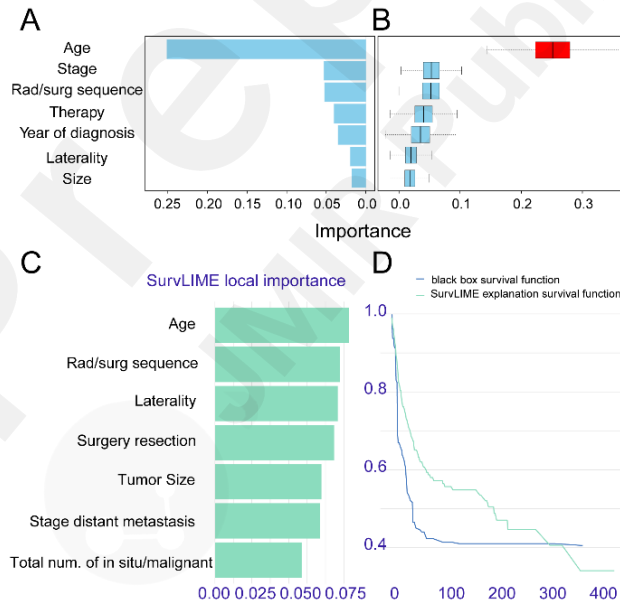


Figure 6. Model interpretability of the random forest survival model.

The bar plot (A) and box plot (B) in the random forest survival analysis display the VIMPs of the eight variables that are most crucial for global interpretability. The VIMPs of the eight variables with the highest significance for local interpretability in the random forest survival analysis and survival rate estimation are displayed in the bar plot (C) and survival curves (D).

References

- [1] Lee JY, Wakabayashi T, Yoshida J. Management and survival of pineoblastoma: an analysis of 34 adults from the brain tumor registry of Japan. *Neurol Med Chir (Tokyo)*. 2005 Mar;45(3):132-41; discussion 141-2.
- [2] Hoffman HJ, Yoshida M, Becker LE, Hendrick EB, Humphreys RP. Pineal region tumors in childhood. Experience at the Hospital for Sick Children. 1983. *Pediatr Neurosurg*. 1994;21(1):91-103; discussion 104.
- [3] Liu APY, Li BK, Vasiljevic A, Dewan MC, Tamrazi B, Ertl-Wagner B, Hansford JR, Pfaff E, Mynarek M, Ng HK, Tsang DS, Gottardo NG, Gajjar A, Bouffet E, Dufour C, Pizer B, Schiff D, Jenkinson MD, Lombardi G, Wen PY, van den Bent MJ, Huang A. SNO-EANO-EURACAN consensus on management of pineal parenchymal tumors. *Neuro Oncol*. 2024 Dec 5;26(12):2159-2173.
- [4] Wilson ER, Takei Y, Bikoff WT, O'Brien MS, Tindall GT, Boehm WM. Abdominal metastases of primary intracranial yolk sac tumors through ventriculoperitoneal shunts: report of three cases. *Neurosurgery*. 1979 Sep;5(3):356-64.
- [5] Lombardi G, Poliani PL, Manara R, Berhouma M, Minniti G, Tabouret E, Razis E, Cerretti G, Zagonel V, Weller M, Idbaih A. Diagnosis and Treatment of Pineal Region Tumors in Adults: A EURACAN Overview. *Cancers (Basel)*. 2022 Jul 27;14(15):3646.
- [6] Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol*. 2016 Jun;131(6):803-20.
- [7] Farnia B, Allen PK, Brown PD, Khatua S, Levine NB, Li J, et al. Clinical outcomes and patterns of failure in pineoblastoma: a 30-year, single-institution retrospective review. *World Neurosurg*. 2014 Dec;82(6):1232-41.
- [8] Mynarek M, Pizer B, Dufour C, van Vuurden D, Garami M, Massimino M, et al. Evaluation of age-dependent treatment strategies for children and young adults with pineoblastoma: analysis of pooled European Society for Paediatric Oncology (SIOP-E) and US Head Start data. *Neuro Oncol*. 2017 Apr 1;19(4):576-585.
- [9] Tate MC, Rutkowski MJ, Parsa AT. Contemporary management of pineoblastoma. *Neurosurg Clin N Am*. 2011 Jul;22(3):409-12, ix.
- [10] Mena H, Rushing EJ, Ribas JL, Delahunt B, McCarthy WF. Tumors of pineal parenchymal cells: a correlation of histological features, including nucleolar organizer regions, with survival in 35 cases. *Hum Pathol*. 1995 Jan;26(1):20-30.
- [11] Tian Y, Liu R, Qin J, Wang J, Ma Z, Gong J, et al. Retrospective Analysis of the Clinical Characteristics, Therapeutic Aspects, and Prognostic Factors of 18 Cases of Childhood Pineoblastoma. *World Neurosurg*. 2018 Aug;116:e162-e168.
- [12] Gaillard F, Jones J. Masses of the pineal region: clinical presentation and radiographic features. *Postgrad Med J*. 2010 Oct;86(1020):597-607.
- [13] Villà S, Miller RC, Krengli M, Abusaris H, Baumert BG, Servagi-Vernat S, et al. Primary pineal tumors: outcome and prognostic factors--a study from the Rare Cancer Network (RCN). *Clin Transl Oncol*. 2012 Nov;14(11):827-34.
- [14] Banerjee AK, Kak VK. Pineoblastoma with spontaneous intra and extracranial metastasis. *J Pathol*. 1974 Sep;114(1):9-12.
- [15] Golbin D, Nikitin KV, Konovalov AN, Pitskhelauri DI, Shishkina LV, Golanov AV, et al. Intraosseous Metastasizing of Pineoblastoma into the Anterior Skull Base, Calvarial Bones, and Vertebrae. *Cureus*. 2015 Dec 28;7(12):e437.
- [16] Nguyen AL, Chong EG, Lee J, Mirshahidi S, Mirshahidi H. A rare case of extremely delayed osseous metastasis of pineoblastoma. *Rare Tumors*. 2021 Jun 20;13:2036361320975752.
- [17] Charafe-Jauffret E, Lehmann G, Fauchon F, Michiels JF, Paquis P, Maraninchi D, et al. Vertebral metastases from pineoblastoma. *Arch Pathol Lab Med*. 2001 Jul;125(7):939-43.
- [18] Zhao P, Strohl A, Gonzalez C, Fishbein T, Rosen-Bronson S, Kallakury B, et al. Donor transmission of pineoblastoma in a two-yr-old male recipient of a multivisceral transplant: a case report. *Pediatr Transplant*. 2012 Jun;16(4):E110-4.
- [19] Fraser G, Rampling R, Smith C, Nicoll J, Stephen M. Long-term survival following extranextra-neural metastasis from a pineoblastoma. *J Neurooncol*. 2000 Jun;48(2):141-4.
- [20] Cranston PE, Hatten MT, Smith EE. Metastatic pineoblastoma via a ventriculoperitoneal shunt: CT demonstration. *Comput Med Imaging Graph*. 1992 Sep-Oct;16(5):349-51.
- [21] Ghosal N, Furtado SV, Hegde AS. Pinealoblastoma with prominent retinoblastic differentiation: an unusual case in an adult. *Neuropathology*. 2010 Aug;30(4):439-42.

- [22] Gadish T, Tulchinsky H, Deutsch AA, Rabau M. Pinealoblastoma in a patient with familial adenomatous polyposis: variant of Turcot syndrome type 2? Report of a case and review of the literature. *Dis Colon Rectum*. 2005 Dec;48(12):2343-6.
- [23] Zhang Z. Model building strategy for logistic regression: purposeful selection. *Ann Transl Med*. 2016 Mar;4(6):111.
- [24] Tate M, Sughrue ME, Rutkowski MJ, Kane AJ, Aranda D, McClintock L, et al. The long-term postsurgical prognosis of patients with pineoblastoma. *Cancer*. 2012 Jan 1;118(1):173-9.
- [25] Ashrafuzzaman M. Artificial Intelligence, Machine Learning and Deep Learning in Ion Channel Bioinformatics. *Membranes (Basel)*. 2021 Aug 31;11(9):672.
- [26] Mirza B, Wang W, Wang J, Choi H, Chung NC, Ping P. Machine Learning and Integrative Analysis of Biomedical Big Data. *Genes (Basel)*. 2019 Jan 28;10(2):87.
- [27] Moncada-Torres A, van Maaren MC, Hendriks MP, Siesling S, Geleijnse G. Explainable machine learning can outperform Cox regression predictions and provide insights in breast cancer survival. *Sci Rep*. 2021 Mar 26;11(1):6968.
- [28] Stüber AT, Coors S, Schachtner B, Weber T, Rügamer D, Bender A, Mittermeier A, Öcal O, Seidensticker M, Ricke J, Bischl B, Ingrisch M. A Comprehensive Machine Learning Benchmark Study for Radiomics-Based Survival Analysis of CT Imaging Data in Patients With Hepatic Metastases of CRC. *Invest Radiol*. 2023 Dec 1;58(12):874-881.
- [29] Shin S, Austin PC, Ross HJ, Abdel-Qadir H, Freitas C, Tomlinson G, et al. Machine learning vs. conventional statistical models for predicting heart failure readmission and mortality. *ESC Heart Fail*. 2021 Feb;8(1):106-115.
- [30] Muschelli J. ROC and AUC with a Binary Predictor: a Potentially Misleading Metric. *J Classif*. 2020 Oct;37(3):696-708.
- [31] Ma H, Bandos AI, Rockette HE, Gur D. On use of partial area under the ROC curve for evaluation of diagnostic performance. *Stat Med*. 2013 Sep 10;32(20):3449-58.
- [32] Indrayan A, Malhotra RK, Pawar M. Use of ROC curve analysis for prediction gives fallacious results: Use predictivity-based indices. *J Postgrad Med*. 2024 Apr 1;70(2):91-96.
- [33] Spytek M, Krzyżiński M, Langbein SH, Baniecki H, Wright MN, Biecek P. survex: an R package for explaining machine learning survival models. *Bioinformatics*. 2023 Dec 1;39(12):btad723.
- [34] Martinez-Cambor, Pablo, Gustavo F. Bayón, and Sonia Pérez-Fernández. "Cumulative/dynamic ROC curve estimation." *Journal of Statistical Computation and Simulation* 86.17 (2016): 3582-3594.
- [35] Mynarek M, Pizer B, Dufour C, van Vuurden D, Garami M, Massimino M, Fangusaro J, Davidson T, Gil-da-Costa MJ, Sterba J, Benesch M, Gerber N, Juhnke BO, Kwiecien R, Pietsch T, Kool M, Clifford S, Ellison DW, Giangaspero F, Wesseling P, Gilles F, Gottardo N, Finlay JL, Rutkowski S, von Hoff K. Evaluation of age-dependent treatment strategies for children and young adults with pineoblastoma: analysis of pooled European Society for Paediatric Oncology (SIOP-E) and US Head Start data. *Neuro Oncol*. 2017 Apr 1;19(4):576-585.
- [36] Nandoliya KR, Sadagopan NS, Thirunavu V, Houskamp EJ, Karras CL, Chaliparambil RK, Sriram N, Jamshidi P, Raleigh DR, Lukas RV, Magill ST. Post-Surgical Prognosis of Patients with Pineoblastoma: A Systematic Review and Individual Patient Data Analysis with Trends over Time. *Cancers (Basel)*. 2023 Jun 27;15(13):3374.
- [37] Greppin K, Cioffi G, Waite KA, Ostrom QT, Landi D, Takaoka K, Kruchko C, Barnholtz-Sloan JS. Epidemiology of pineoblastoma in the United States, 2000-2017. *Neurooncol Pract*. 2022 Jan 27;9(2):149-157.
- [38] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006 Nov-Dec;26(6):565-74.
- [39] Abbey CK, Gallas BD, Boone JM, Niklason LT, Hadjiiski LM, Sahiner B, et al. Comparative statistical properties of expected utility and area under the ROC curve for laboratory studies of observer performance in screening mammography. *Acad Radiol*. 2014 Apr;21(4):481-90.
- [40] Speiser JL, Miller ME, Tooze J, Ip E. A Comparison of Random Forest Variable Selection Methods for Classification Prediction Modeling. *Expert Syst Appl*. 2019 Nov 15;134:93-101.
- [41] Bostwick DG, Burke HB. Prediction of individual patient outcome in cancer: comparison of artificial neural networks and Kaplan--Meier methods. *Cancer*. 2001 Apr 15;91(8 Suppl):1643-6.
- [42] Chen S, Liu Y, Yang J, Liu Q, You H, Dong Y, et al. Development and Validation of a Nomogram for Predicting Survival in Male Patients with Breast Cancer. *Front Oncol*. 2019 May 14;9:361.
- [43] Martini A, Kumarasamy S, Beksac AT, Abaza R, Eun DD, Bhandari A, et al. A Nomogram to Predict Significant Estimated Glomerular Filtration Rate Reduction After Robotic Partial Nephrectomy. *Eur Urol*. 2018 Dec;74(6):833-

839.

- [44] Reddy AT, Janss AJ, Phillips PC, Weiss HL, Packer RJ. Outcome for children with supratentorial primitive neuroectodermal tumors treated with surgery, radiation, and chemotherapy. *Cancer*. 2000 May 1;88(9):2189-93.
- [45] Gururangan S, McLaughlin C, Quinn J, Rich J, Reardon D, Halperin EC, et al. High-dose chemotherapy with autologous stem-cell rescue in children and adults with newly diagnosed pineoblastomas. *J Clin Oncol*. 2003 Jun 1;21(11):2187-91.
- [46] Gilheeney SW, Saad A, Chi S, Turner C, Ullrich NJ, Goumnerova L, et al. Outcome of pediatric pineoblastoma after surgery, radiation and chemotherapy. *J Neurooncol*. 2008 Aug;89(1):89-95.
- [47] Jing Y, Deng W, Zhang H, Jiang Y, Dong Z, Fan F, et al. Development and Validation of a Prognostic Nomogram to Predict Cancer-Specific Survival in Adult Patients With Pineoblastoma. *Front Oncol*. 2020 Jul 24;10:1021.
- [48] Mallick S, Benson R, Rath GK. Patterns of care and survival outcomes in patients with pineal parenchymal tumor of intermediate differentiation: An individual patient data analysis. *Radiother Oncol*. 2016 Nov;121(2):204-208.
- [49] Schild SE, Scheithauer BW, Schomberg PJ, Hook CC, Kelly PJ, Frick L, et al. Pineal parenchymal tumors. Clinical, pathologic, and therapeutic aspects. *Cancer*. 1993 Aug 1;72(3):870-80.
- [50] Ghim TT, Davis P, Seo JJ, Crocker I, O'Brien M, Krawiecki N. Response to neoadjuvant chemotherapy in children with pineoblastoma. *Cancer*. 1993 Sep 1;72(5):1795-800.
- [51] Clark AJ, Sughrue ME, Ivan ME, Aranda D, Rutkowski MJ, Kane AJ, et al. Factors influencing overall survival rates for patients with pineocytoma. *J Neurooncol*. 2010 Nov;100(2):255-60.
- [52] Massimino M, Gandola L, Spreafico F, Luksch R, Collini P, Giangaspero F, et al. Supratentorial primitive neuroectodermal tumors (S-PNET) in children: A prospective experience with adjuvant intensive chemotherapy and hyperfractionated accelerated radiotherapy. *Int J Radiat Oncol Biol Phys*. 2006 Mar 15;64(4):1031-7.
- [53] Timmermann B, Kortmann RD, Kühl J, Meisner C, Dieckmann K, Pietsch T, et al. Role of radiotherapy in the treatment of supratentorial primitive neuroectodermal tumors in childhood: results of the prospective German brain tumor trials HIT 88/89 and 91. *J Clin Oncol*. 2002 Feb 1;20(3):842-9.
- [54] Collier TJ, Kanaan NM, Kordower JH. Ageing as a primary risk factor for Parkinson's disease: evidence from studies of non-human primates. *Nat Rev Neurosci*. 2011 Jun;12(6):359-66.
- [55] Keenan CR, White RH. Age as a risk factor for venous thromboembolism after major surgery. *Curr Opin Pulm Med*. 2005 Sep;11(5):398-402.
- [56] Basina M, Maahs DM. Age at type 1 diabetes onset: a new risk factor and call for focused treatment. *Lancet*. 2018 Aug 11;392(10146):453-454.
- [57] Najjar SS, Scuteri A, Lakatta EG. Arterial aging: is it an immutable cardiovascular risk factor? *Hypertension*. 2005 Sep;46(3):454-62.
- [58] Marinigh R, Lip GY, Fiotti N, Giansante C, Lane DA. Age as a risk factor for stroke in atrial fibrillation patients: implications for thromboprophylaxis. *J Am Coll Cardiol*. 2010 Sep 7;56(11):827-37.
- [59] Yuan C, Kim J, Wang QL, Lee AA, Babic A; PanScan/PanC4 I-III Consortium; Amundadottir LT, et al. The age-dependent association of risk factors with pancreatic cancer. *Ann Oncol*. 2022 Jul;33(7):693-701.
- [60] Guliyeva G, Huayllani MT, Boczar D, Avila FR, Lu X, Forte AJ. Age as a risk factor for breast cancer-related lymphedema: a systematic review. *J Cancer Surviv*. 2023 Feb;17(1):246-253.
- [61] White MC, Holman DM, Boehm JE, Peipins LA, Grossman M, Henley SJ. Age and cancer risk: a potentially modifiable relationship. *Am J Prev Med*. 2014 Mar;46(3 Suppl 1):S7-15.
- [62] Parikh KA, Venable GT, Orr BA, Choudhri AF, Boop FA, Gajjar AJ, et al. Pineoblastoma-The Experience at St. Jude Children's Research Hospital. *Neurosurgery*. 2017 Jul 1;81(1):120-128.
- [63] Tate M, Sughrue ME, Rutkowski MJ, Kane AJ, Aranda D, McClinton L, et al. The long-term postsurgical prognosis of patients with pineoblastoma. *Cancer*. 2012 Jan 1;118(1):173-9.
- [64] Jin MC, Prolo LM, Wu A, Azad TD, Shi S, Rodrigues AJ, et al. Patterns of Care and Age-Specific Impact of Extent of Resection and Adjuvant Radiotherapy in Pediatric Pineoblastoma. *Neurosurgery*. 2020 May 1;86(5):E426-E435.
- [65] Lutterbach J, Fauchon F, Schild SE, Chang SM, Pagenstecher A, Volk B, et al. Malignant pineal parenchymal tumors in adult patients: patterns of care and prognostic factors. *Neurosurgery*. 2002 Jul;51(1):44-55; discussion 55-6.
- [66] Jakacki RI, Zeltzer PM, Boyett JM, Albright AL, Allen JC, Geyer JR, et al. Survival and prognostic factors following radiation and/or chemotherapy for primitive neuroectodermal tumors of the pineal region in infants and children: a report of the Childrens Cancer Group. *J Clin Oncol*. 1995 Jun;13(6):1377-83.
- [67] Dittmar Y, Ardelt M, Scheuerlein H, Rauchfuss F, Dondorf F, Settmacher U. The Impact of Tumor Size on the

- Prognosis of Gastric Cancer: Experiences from a European Study Group. *J Surg.* 2016; 3(1): 1082.
- [68] Park JY, Kim DY, Kim JH, Kim YM, Kim YT, Nam JH. Outcomes after radical hysterectomy according to tumor size divided by 2-cm interval in patients with early cervical cancer. *Ann Oncol.* 2011 Jan;22(1):59-67.
- [69] Huang B, Feng Y, Mo SB, Cai SJ, Huang LY. Smaller tumor size is associated with poor survival in T4b colon cancer. *World J Gastroenterol.* 2016 Aug 7;22(29):6726-35.
- [70] Han, H. S., Paik, H. J., Ryu, J. M., Park, S., Yi, H. W., Bae, S., et al. Comparison of prognosis and specific features according to tumor size in small-sized breast cancer with extensive lymph node involvement. *Journal of Clinical Oncology*, 2015 Oct 1; 33: 81-81.
- [71] Tirada N, Aujeiro M, Khorjekar G, Richards S, Chopra J, Dromi S, et al. Breast Cancer Tissue Markers, Genomic Profiling, and Other Prognostic Factors: A Primer for Radiologists. *Radiographics.* 2018 Nov-Dec;38(7):1902-1920.
- [72] Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, et al. DNA methylation-based classification of central nervous system tumours. *Nature.* 2018 Mar 22;555(7697):469-474.
- [73] Li BK, Vasiljevic A, Dufour C, Yao F, Ho BLB, Lu M, et al. Pineoblastoma segregates into molecular sub-groups with distinct clinico-pathologic features: a Rare Brain Tumor Consortium registry study. *Acta Neuropathol.* 2020 Feb;139(2):223-241.
- [74] Liu APY, Guden B, Lin T, Orr BA, Klimo P Jr, Kumar R, et al. Risk-adapted therapy and biological heterogeneity in pineoblastoma: integrated clinico-pathological analysis from the prospective, multi-center SJMB03 and SJYC07 trials. *Acta Neuropathol.* 2020 Feb;139(2):259-271.
- [75] Pfaff E, Aichmüller C, Sill M, Stichel D, Snuderl M, Karajannis MA, Schuhmann MU, et al. Molecular subgrouping of primary pineal parenchymal tumors reveals distinct subtypes correlated with clinical parameters and genetic alterations. *Acta Neuropathol.* 2020 Feb;139(2):243-257.
- [76] Liu APY, Li BK, Pfaff E, Guden B, Vasiljevic A, Orr BA, et al. Clinical and molecular heterogeneity of pineal parenchymal tumors: a consensus study. *Acta Neuropathol.* 2021 May;141(5):771-785.

Supplementary Files