Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Prognostic Machine Learning Model for Children with Wilms Tumor: A SEER Database Study

Jingyi Chen
Fujian Maternity and Child Health Hospital
Pingping Yu
The First Hospital of Putian City
Changyuan Wang
1056931841@qq.com

Fujian Maternity and Child Health Hospital

Research Article

Keywords: Wilms tumor, Prognostic factors, Prognostic model, Machine learning, SEER database

Posted Date: March 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8876618/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Objective: This study aimed to establish and validate a machine learning based prognostic model to predict the prognostic risk of childhood Wilms tumor.

Methods: Data of 1958 children with Wilms tumor (data from 2000-2022) were obtained from the Surveillance, Epidemiology, and End Results (SEER) database. The data were divided into a training set (70%) and a validation set (30%). Prognostic factors were analyzed by Cox regression. Five machine learning algorithms were used to construct models, which were evaluated by C-index, receiver operating characteristic curve (ROC), calibration curve, and decision curve analysis (DCA). Meanwhile, a nomogram was drawn to visualize the model.

Results: The results of the multivariate Cox regression analysis showed that age, lymph node density (LND), tumor stage, and place of residence were independent risk factors related to the prognosis. Among the fitted machine learning models, the LASSO-Cox model demonstrated relatively stable and accurate predictive performance. The C-index of the training set was 0.755 (95% CI: 0.704-0.806), and the C-index of the validation set was 0.752 (95% CI: 0.660-0.844).

Conclusion: This study has identified the key prognostic indicators for childhood Wilms tumor, which can assist surgeons in accurately identifying the high-risk population with poor prognosis of Wilms tumor.

1. Introduction

As the most common malignant kidney tumor in children, Wilms tumor seriously threatens the health and life safety of children[1]. According to the latest global cancer statistics, there are approximately 7 to 9 new cases per million children, accounting for a significant proportion of the incidence of pediatric solid tumors[2]. With the advancement of medical technology, comprehensive treatments such as surgery, chemotherapy, and radiotherapy have been increasingly optimized, leading to a significant improvement in the overall survival rate of children with Wilms tumor[3]. Recent studies have shown that the survival rate of some low risk children can reach over 90% after receiving standardized treatment[4]. However, in clinical practice, there are significant differences in the prognosis of different children. Some children who receive the same treatment plan have completely different outcomes. Children with lymph node metastasis, advanced tumor staging, and bilateral disease have a high risk of recurrence and a significantly reduced survival rate[5, 6]. Surgeons mainly rely on traditional factors such as tumor staging, pathological type, and age to assess the prognosis of children with Wilms tumor and formulate treatment plans[7]. For example, tumor staging can describe the scope of tumor growth and the extent of metastasis, there can still be significant differences in prognosis among patients within the same stage[8]. However, these factors are hardly capable of comprehensively and accurately predicting the prognosis of each individual child[9].

Machine learning has been widely applied in tumor diseases recently, providing assistance for disease diagnosis and prognosis prediction[10]. This method is helpful for processing high-dimensional complex clinical data, which has significantly improved the accuracy of prognosis prediction, offering support for clinical decision making[11-13]. Zhong et al found that the XGBoost algorithm can effectively predict the diagnosis and survival of breast cancer bone metastasis[12]. Munai et al proposed a random forest model to predict the risk of brain metastasis in extensive-stage small cell lung cancer[13]. Surveillance, Epidemiology, and End Results (SEER) database contains clinically relevant data on children with malignant tumors, covering populations of different races, regions, etc. There are many clinical prediction models built based on the SEER database, which can help us identify disease-related risk factors[14, 15]. However, the detailed analysis of construct the prognostic model for pediatric Wilms tumor using machine learning is lacked.

In this study, a prognostic prediction model was constructed for pediatric nephroblastoma using machine learning algorithms based on the SEER database. The aim of this study is to investigate a more accurate prognostic assessment tool, which helps to guide individualized treatment decisions for pediatric Wilms tumor patients.

2. Materials and Methods

2.1 Study population

The data are sourced from the SEER database (January 2024 subset). The time span ranges from 2000 to 2022. The inclusion criteria are as follows: (1) The primary site is the kidney (C64.9). (2) The histology is Nephroblastoma (ICCC site recode 3rd edition / IARC 2017). (3) The behavior records indicate malignancy. (4) The age is under 18 years old. (5) The tumor is solitary. The exclusion criteria are as follows: (1) Patients with incomplete information, such as tumor size, ethnic information, survival information, tumor stage, and surgical information; (2) Deaths caused by non-tumor reasons; (3) Cases with unknown or unquantified numbers of examined regional lymph nodes. The detailed selection procedure can be referred to in **Figure 1**.

2.2 Study variables

This study is a prognostic model research, and the outcome indicators are survival time and survival status. In the SEER database, information is extracted through the two fields of “Survival months” and “Vital status recode (study cutoff used)”.

Other indicators included the following demographic and clinically related indicators of the children including gender, age, race, tumor size, residence, laterality, tumor stage, lymph node density (LND), surgery, radiotherapy, and chemotherapy. Among them, the place of residence was reclassified into “Metropolitan” and “Non-metropolitan”. The tumor stage was combined according to the two fields of “Summary stage 2000 (1998-2017)” and “Combined Summary Stage with Expanded Regional Codes

(2004+)", and reclassified into "Localized", "Regional", and "Distant". LND was defined as the proportion of positive regional lymph nodes among the examined regional lymph nodes. The surgical intervention was redefined as "Partial Nephrectomy" and "Total Nephrectomy".

2.3 Model Establishment, Evaluation and Validation

Randomly divide the total dataset into a training set and a validation set at a ratio of 7:3. Using differential analysis to compare all variables in the two sets to ensure the balance in random splitting. In the training set, conduct univariate COX regression analysis on all independent variables and select variables with a *P*-value less than 0.05 for subsequent machine learning model building. Five machine learning methods, namely Random Survival Forest (RSF), Gradient Boosting Machine (GBM), LASSO-Cox, CoxBoost, and Survival Support Vector Machine (Survival SVM), were used to construct models, and the models were optimized through hyperparameter tuning. Models were evaluated by C-index, ROC curve analysis, decision curve analysis, and calibration curve analysis. Moreover, the stability of the model is verified by the validation set. Use a dynamic nomogram to visualize the results of the LASSO-Cox model.

2.4 Statistical methods

Using the Shapiro-Wilk test to examine the normality of the data. For quantitative data that are not normally distributed, represent them with M (Q1, Q3) and analyze using the Mann-Whitney U test. For categorical data, present as percentages (%) and analyze using the chi-square test or Fisher's exact test. A P-value < 0.05 indicates the difference is statistically significant.

Software including R4.4.2 (https://www.r - project.org/) and SEER*Stat (version 8.4.2, https://seer.cancer.gov/seerstat/) were used. The software packages used included "tidyverse", "survival", "survminer", "caret", "randomForestROC", "gbm", "glmnet", "CoxBoost", "kernlab", "pROC", "rms", "ggplot2".

2.5 Ethical Statement and Reporting Standards

This study used the public SEER database in the United States. The database contained no personally identifiable information and follows ethical standards such as the Declaration of Helsinki. This study adhered to the guidelines stipulated in the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement.

3. Results

3.1 The characteristics of patients

A total of 1958 children with Wilms tumor were included in this study, among which 1370 were in the training set and 588 were in the validation set. As shown in **Table 1**, there were no significant statistical differences in the clinical and demographic characteristics of the children between the two sets

(P>0.05), indicating that the training set and the validation set were balanced and comparable, ensuring the reliability of the model.

Table 1 Clinical Characteristics of Children with Wilms' Tumor in the Training Set and Validation Set

Variables	Total (n=1958)	Training set (n=1370)	Validation set (n=588)	*P*
Time, M (Q_1, Q_3)	103.50 (42.25, 181.00)	104.50 (42.00, 180.75)	102.50 (43.75, 182.00)	0.876
Age, M (Q_1, Q_3)	3.00 (1.00, 5.00)	3.00 (1.00, 5.00)	3.00 (1.00, 4.00)	0.462
LND, M (Q_1, Q_3)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.860
Tumor Size, M (Q_1, Q_3)	110.00 (85.00, 135.00)	110.00 (85.00, 135.00)	110.00 (87.00, 134.00)	0.863
Status, n(%)				0.096
Alive	1849 (94.43)	1286 (93.87)	563 (95.75)	
Dead	109 (5.57)	84 (6.13)	25 (4.25)	
Sex, n(%)				0.503
Male	915 (46.73)	647 (47.23)	268 (45.58)	
Female	1043 (53.27)	723 (52.77)	320 (54.42)	
Race, n(%)				0.537
White	1488 (76.00)	1046 (76.35)	442 (75.17)	
Black	347 (17.72)	235 (17.15)	112 (19.05)	
Other	123 (6.28)	89 (6.50)	34 (5.78)	
Stage, n(%)				0.435
Localized	841 (42.95)	597 (43.58)	244 (41.50)	
Regional	673 (34.37)	473 (34.53)	200 (34.01)	
Distant	444 (22.68)	300 (21.90)	144 (24.49)	
Laterality, n(%)				0.094

Variables	Total (n=1958)	Training set (n=1370)	Validation set (n=588)	*P*
Left	959 (48.98)	692 (50.51)	267 (45.41)	
Right	899 (45.91)	613 (44.74)	286 (48.64)	
Bilateral	100 (5.11)	65 (4.74)	35 (5.95)	
Residence, n(%)				0.939
Metropolitan	1783 (91.06)	1248 (91.09)	535 (90.99)	
Non-metropolitan	175 (8.94)	122 (8.91)	53 (9.01)	
Surgery, n(%)				0.654
Partial Nephrectomy	87 (4.44)	59 (4.31)	28 (4.76)	
Total Nephrectomy	1871 (95.56)	1311 (95.69)	560 (95.24)	
Radiation, n(%)				0.671
No/Unknown	1000 (51.07)	704 (51.39)	296 (50.34)	
Yes	958 (48.93)	666 (48.61)	292 (49.66)	
Chemotherapy, n(%)				0.407
No/Unknown	141 (7.20)	103 (7.52)	38 (6.46)	
Yes	1817 (92.80)	1267 (92.48)	550 (93.54)	

3.2 Univariate and Multivariate Cox Regression Analyses

Taking the prognosis of children with Wilms tumor as the dependent variable and other variables as independent variables, univariate and multivariate COX regression analyses were conducted respectively. As listed in **Table 2**, the multivariate COX regression analysis showed that age (HR=1.065, P=0.049), LND (HR=3.717, P<0.001), tumor stage (Regional: HR=2.099, P=0.032; Distant: HR=4.444, P<0.001), and place of residence (HR=2.158, P=0.012) were independent risk factors affecting the prognosis of Wilms tumor

patients. Specifically, children with older age, higher LND, higher tumor stage, and living in non-metropolitan areas had a higher risk of death in terms of prognosis.

Table 2 Univariate and Multivariate COX Regression Analysis of Prognosis in Children with Wilms Tumor

Variables	Univariate analysis		Multivariate analysis	
	HR (95%CI)	*P*	HR (95%CI)	*P*
Age	1.104 (1.045 ~ 1.165)	<0.001*	1.065 (1.001 ~ 1.135)	0.049*
LND	6.820 (3.945 ~ 11.791)	<0.001*	3.717 (2.037 ~ 6.782)	<0.001*
Tumor Size	1.007 (1.001 ~ 1.013)	0.015*	1.002 (0.996 ~ 1.008)	0.562
Stage				
Localized	1.000 (Reference)		1.000 (Reference)	
Regional	2.963 (1.546 ~ 5.681)	0.001*	2.099 (1.068 ~ 4.126)	0.032*
Distant	7.022 (3.762 ~ 13.105)	<0.001*	4.444 (2.269 ~ 8.702)	<0.001*
Race				
White	1.000 (Reference)			
Black	1.352 (0.799 ~ 2.288)	0.261		
Other	1.021 (0.410 ~ 2.541)	0.965		
Laterality				
Left	1.000 (Reference)			
Right	0.652 (0.409 ~ 1.039)	0.072		
Bilateral	1.871 (0.885 ~ 3.955)	0.101		
Sex				
Male	1.000 (Reference)			
Female	0.769 (0.501 ~ 1.181)	0.230		
Residence				
Metropolitan	1.000 (Reference)		1.000 (Reference)	

Variables	Univariate analysis		Multivariate analysis	
	HR (95%CI)	*P*	HR (95%CI)	*P*
Non-metropolitan	1.863 (1.031 ~ 3.364)	0.039*	2.158 (1.188 ~ 3.922)	0.012*
Surgery				
Partial Nephrectomy	1.000 (Reference)			
Total Nephrectomy	3.565 (0.496 ~ 25.613)	0.206		
Radiation				
No/Unknown	1.000 (Reference)			
Yes	2.658 (1.665 ~ 4.243)	<0.001*		
Chemotherapy				
No/Unknown	1.000 (Reference)			
Yes	0.810 (0.391 ~ 1.679)	0.571		

* The difference was statistically significant (P<0.05).

3.3 Establishment, Evaluation and Validation of the Model

Prognostic models were constructed using five machine learning algorithms including RSF, GBM, LASSO-Cox, CoxBoost, and Survival SVM. The LASSO regression method was used to screen variables, and five indicators including age, LND, tumor size, stage and place of residence were selected. These were incorporated into the COX regression model. In the training set, the C-indices of the five models were presented as follows, RSF: 0.968 (95%CI: 0.958-0.978), GBM: 0.906 (95%CI: 0.875-0.937), LASSO-Cox: 0.755 (95%CI: 0.704-0.806), CoxBoost: 0.757 (95%CI: 0.708-0.806), Survival SVM: 0.656 (95%CI: 0.591-0.721). In the validation set, the C-indices were presented as follows, RSF: 0.685 (95%CI: 0.577-0.793), GBM: 0.723 (95%CI: 0.615-0.831), LASSO-Cox: 0.752 (95%CI: 0.660-0.844), CoxBoost: 0.748 (95%CI: 0.656-0.840), Survival SVM: 0.613 (95%CI: 0.501-0.725). Although the RSF and GBM models showed high C-indices in the training set, indicating strong predictive ability, their C-indices in the validation set were poor. The C-indices of the LASSO-Cox model and the CoxBoost model were relatively stable in both the training set and the validation set. Considering the better interpretability and visual effects of the COX model, the LASSO-Cox regression model was adopted for prognostic prediction.

The ROC curves of the five prognostic models for 3-year, 5-year, and 8-year are shown in **Figure 2**. For the LASSO-Cox model, the 3-year, 5-year, and 8-year AUC values in the training set were presented as follows 0.784 (95%CI: 0.730-0.837), 0.761 (95%CI: 0.705-0.816), 0.758 (95%CI: 0.704-0.812). In the validation set, the AUC values at 3-year, 5-year and 8-year were 0.734 (95% CI: 0.623-0.845), 0.754 (95% CI: 0.656-0.853), and 0.754 (95% CI: 0.656-0.853), respectively, indicating the excellent predictive performance of the LASSO-Cox model. In the subsequent stage, we preferentially selected the more stable Cox model to develop the prognostic nomogram.

Further analysis was conducted through the calibration curve for an 8-year prognosis, which confirmed that the Cox model had more excellent predictive performance (**Figure 3A, Figure 3B**). Decision curve analysis (DCA) indicated that this model demonstrated excellent performance in predicting patients' survival probability (**Figure 3C, Figure 3D**). In the subsequent stage, we preferentially selected the more stable LASSO-Cox model to develop the prognostic nomogram (**Figure 4**).

4. Discussion

Based on the SEER database, this study constructed a predictive model for the prognosis of pediatric Wilms tumor using multiple machine learning approaches. Several independent risk factors associated with patient outcomes were identified, including age, LND, tumor stage, and place of residence.

There is an association between age and both tumor biological characteristics and treatment outcomes. As reported by Mansfield et al, older pediatric patients are more likely to present with invasive tumor phenotypes[16]. Moreover, their tolerance to chemotherapy varies considerably, and their overall recovery capacity tends to be reduced[16]. David et al pointed out that for older children with Wilms' tumor, the proliferation rate of tumor cells increases, and their sensitivity to chemotherapy drugs decreases, this results in poor treatment outcomes and a poorer prognosis[5]. Chen et al found that age is a significant factor influencing the immune response characteristics in pediatric patients with solid tumors[17]. Older children typically possess a more mature immune system, which may increase their susceptibility to immune-related complications during the treatment of malignant tumors, potentially adversely affecting their prognosis[17].

LND is an important indicator for evaluating the degree of lymph node metastasis in malignant tumors. Elevated LND levels indicate a greater degree of lymph node involvement, which is strongly associated with an increased risk of long-term tumor metastasis and a higher likelihood of adverse clinical outcomes[18]. Furthermore, increased LND reflects extensive infiltration of peritumoral tissues, suggesting that tumor cells may enter the bloodstream via the lymphatic system, thereby facilitating distant organ metastasis[19]. According to the study by Nseir et al, patients with higher LND values exhibit significantly lower 5-year survival rates and elevated recurrence risks[20]. Therefore, detecting the level of LND is of great value in predicting the prognosis of children with Wilms tumor.

The findings of Jiang et al provide evidence that tumor stage is a significant prognostic factor for children with Wilms' tumor. Patients with regional or distant metastasis demonstrate a markedly elevated mortality risk in comparison to those with localized disease[21]. The progression of tumors can be divided into growth and metastasis. As the tumor advances and infiltrates surrounding tissues, the clinical stage increases progressively, resulting in heightened treatment complexity and a significantly reduced survival duration among pediatric patients[22, 23].

Taye et al reported that patients residing in remote areas frequently present with advanced-stage clinical manifestations and demonstrate a reduced survival rate[24]. Geographic remoteness may delay access to timely diagnostic evaluations and therapeutic interventions for pediatric patients. However, studies by Arash Delavar et al have shown that the cancer survival rates of children and adolescents in the United States do not vary based on their metropolitan or non-metropolitan residence at the time of diagnosis[25]. This study shows that the survival rate of children with Wilms tumor living in metropolitan areas is higher than that of children living in non-metropolitan areas. Among various machine learning models, the LASSO-Cox model demonstrates good stability and predictive performance. This model combines the variable-selection ability of the LASSO algorithm with the survival-analysis advantage of Cox regression. It effectively screens out key prognostic factors and avoids the problem of over fitting[26].

However, our study is not without limitations. First, although the SEER database encompasses a large number of cases, it lacks certain potentially important clinical and molecular data, such as genetic test results, detailed treatment protocols, and adverse reaction records. The absence of these variables may compromise the accuracy and comprehensiveness of the model[27]. Second, this study only performed internal validation using data from the SEER database. Future research should include external validation using independent cohorts to confirm the model’s stability and generalizability[28]. Patients from different regions and ethnic groups may have differences in genetic background, living environment, and medical conditions. These differences may affect the predictive performance of the model. Therefore, validation in a broader population is required[29].

5. Conclusion

In summary, this study constructed a machine learning model for the prognosis of pediatric Wilms tumor based on the SEER database. It was found that age, LND, tumor stage, and place of residence are independent risk factors related to prognosis. This model plays a supportive role in the treatment decisions and follow-up management of patients. Future research can further incorporate external data for external validation, and integrate more predictive factors, thereby significantly enhancing the practical application efficiency of the model.

Declarations

Ethical approval

This study uses the public SEER database in the United States. The database contains no personally identifiable information and follows ethical standards such as the Declaration of Helsinki.

Consent to participate

Not applicable, as no direct patient interaction occurred and data were anonymized.

Consent to publish

Not applicable due to the use of de-identified registry data.

Author Contributions Statement

Jingyi Chen: made contributions to the conception and design of the study, acquisition of data (laboratory or clinical), data analysis and/or interpretation, as well as drafting of the manuscript and/or critical revision.

Pingping Yu: contributed to the conception and design of the study, acquisition of data (laboratory or clinical), and drafting of the manuscript and/or critical revision.

Changyuan Wang: participated in the conception and design of the study, acquisition of data (laboratory or clinical), data analysis and/or interpretation, and approval of the final version of the manuscript.

Declaration of Competing Interest

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

Sources of Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgments

The authors gratefully acknowledge all support from Fujian Maternity and Child Health Hospital.

Data Availiability declaration

The data that support the findings of this study are openly available in software package SEER*Stat 8.4.5 (https://seer.cancer.gov/seerstat/).

References

1. Spreafico F, Fernandez CV, Brok J et al. Wilms tumour. Nat Rev Dis Primers. 2021 7(1):75.
2. Libes J, Hol J, Neto JCA et al. Pediatric renal tumor epidemiology: Global perspectives, progress, and challenges. Pediatr Blood Cancer. 2023 70(1):e30006.
3. Saltzman AF, Cost NG, Romao RLP. Wilms Tumor. Urol Clin North Am. 2023 50(3):455-464.
4. Bhutani N, Kajal P, Sharma U. Many faces of Wilms Tumor: Recent advances and future directions. Ann Med Surg (Lond). 2021 64:102202.
5. Qian DC, Sykes-Martin KD, Tobillo R et al. Impact of Age on Overall Survival Among Children With Wilms Tumor: A Population-based Registry Analysis. Am J Clin Oncol. 2023 46(5):213-218.
6. Chen SY, Lee WGH, Laifman E et al. A Single Center Experience With Bilateral Wilms Tumor. Am Surg. 2023 89(10):4101-4104.
7. Neagu MC, David VL, Iacob ER et al. Wilms' Tumor: A Review of Clinical Characteristics, Treatment Advances, and Research Opportunities. Medicina (Kaunas). 2025 61(3).
8. Parsons LN. Wilms Tumor: Challenges and Newcomers in Prognosis. Surg Pathol Clin. 2020 13(4):683-693.
9. Balis F, Green DM, Anderson C et al. Wilms Tumor (Nephroblastoma), Version 2.2021, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2021 19(8):945-977.
10. Margue G, Ferrer L, Etchepare G et al. UroPredict: Machine learning model on real-world data for prediction of kidney cancer recurrence (UroCCR-120). NPJ Precis Oncol. 2024 8(1):45.
11. Kourou K, Exarchos TP, Exarchos KP et al. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J. 2015 13:8-17.
12. Zhong X, Lin Y, Zhang W et al. Predicting diagnosis and survival of bone metastasis in breast cancer using machine learning. Sci Rep. 2023 13(1):18301.
13. Munai E, Zeng S, Yuan Z et al. Machine learning-based prediction model for brain metastasis in patients with extensive-stage small cell lung cancer. Sci Rep. 2024 14(1):28790.
14. Xiong Y, Cao H, Zhang Y et al. Nomogram-Predicted Survival of Breast Cancer Brain Metastasis: a SEER-Based Population Study. World Neurosurg. 2019 128:e823-e834.
15. Tan X, Wang J, Tang J et al. A Nomogram for Predicting Cancer-Specific Survival in Children With Wilms Tumor: A Study Based on SEER Database and External Validation in China. Front Public Health. 2022 10:829840.
16. Mansfield SA, Lamb MG, Stanek JR et al. Renal Tumors in Children and Young Adults Older Than 5 Years of Age. J Pediatr Hematol Oncol. 2020 42(4):287-291.
17. Chen Q, Zhao B, Tan Z et al. Systems-level immunomonitoring in children with solid tumors to enable precision medicine. Cell. 2025 188(5):1425-1440.e1411.
18. Walker JP, Johnson JS, Eguchi MM et al. Factors affecting lymph node sampling patterns and the impact on survival of lymph node density in patients with Wilms tumor: a Surveillance, Epidemiology,

and End Result (SEER) database review. J Pediatr Urol. 2020 16(1):81-88.
19. Reticker-Flynn NE, Zhang W, Belk JA et al. Lymph node colonization induces tumor-immune tolerance to promote distant metastasis. Cell. 2022 185(11):1924-1942.e1923.
20. Nseir S, Zeineh N, Capucha T et al. The impact of lymph node density as a predictive factor for survival and recurrence of tongue squamous cell carcinoma. Int J Oral Maxillofac Surg. 2022 51(4):441-449.
21. Jiang L, Tong Y, Wang J et al. A dynamic visualization clinical tool constructed and validated based on the SEER database for screening the optimal surgical candidates for bone metastasis in primary kidney cancer. Sci Rep. 2024 14(1):3561.
22. Zhang L, Wang WQ, Chen JH et al. Tumor-infiltrating immune cells and survival in head and neck squamous cell carcinoma: a retrospective computational study. Sci Rep. 2024 14(1):6390.
23. Miotke L, Nevala-Plagemann C, Ying J et al. Treatment outcomes in recurrent versus de novo metastatic pancreatic adenocarcinoma: a real world study. BMC Cancer. 2022 22(1):1054.
24. Taye BW, Clark PJ, Hartel G et al. Remoteness of residence predicts tumor stage, receipt of treatment, and mortality in patients with hepatocellular carcinoma. JGH Open. 2021 5(7):754-762.
25. Delavar A, Feng Q, Johnson KJ. Rural/urban residence and childhood and adolescent cancer survival in the United States. Cancer. 2019 125(2):261-268.
26. Wang Q, Qiao W, Zhang H et al. Nomogram established on account of Lasso-Cox regression for predicting recurrence in patients with early-stage hepatocellular carcinoma. Front Immunol. 2022 13:1019638.
27. Nelson MV, van den Heuvel-Eibrink MM, Graf N et al. New approaches to risk stratification for Wilms tumor. Curr Opin Pediatr. 2021 33(1):40-48.
28. Wang Q, Sun Z, Xu X et al. The Evaluation of a SEER-Based Nomogram in Predicting the Survival of Patients Treated with Neoadjuvant Therapy Followed by Esophagectomy. Front Surg. 2022 9:853093.
29. Nie Y, Ying B, Lu Z et al. Predicting survival and prognosis of postoperative breast cancer brain metastasis: a population-based retrospective analysis. Chin Med J (Engl). 2023 136(14):1699-1707.

Figures

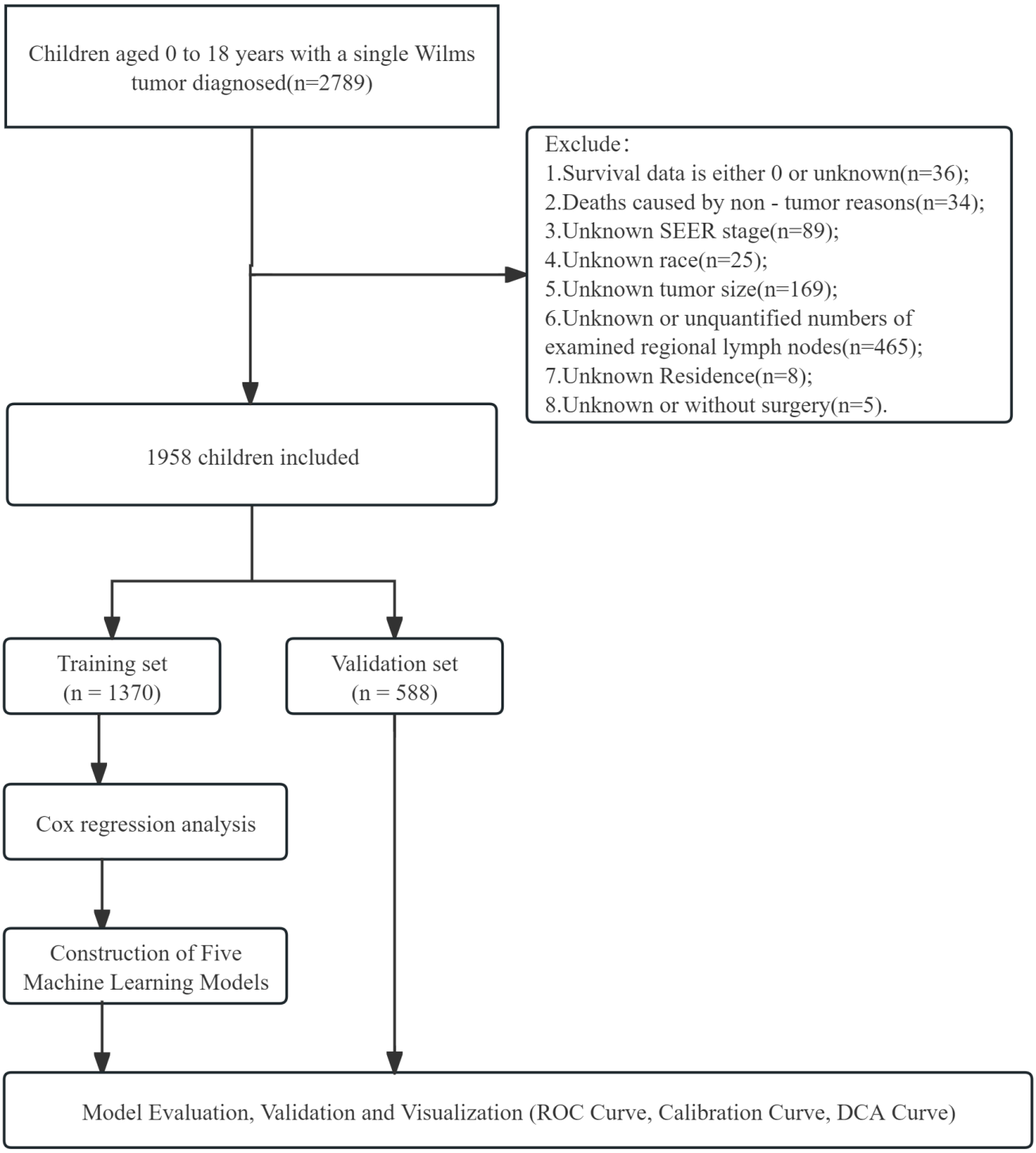


Figure 1

Flowchart of Inclusion and Exclusion Criteria for Patients with Wilms Tumor

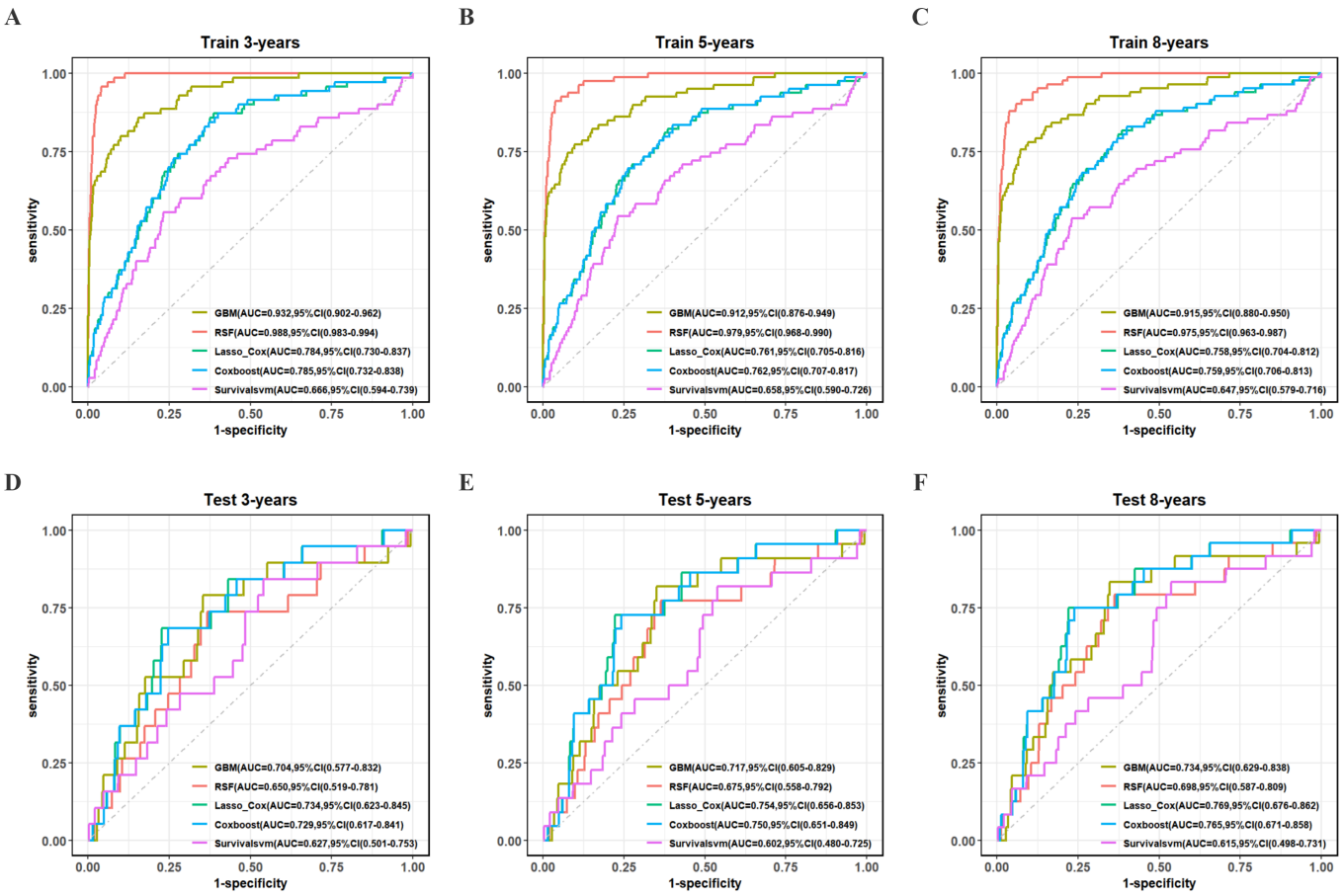


Figure 2

ROC curve analysis of five machine-learning models in the training set and validation set for 3-year, 5-year, and 8-year time-points. The LASSO-COX model demonstrates greater stability.

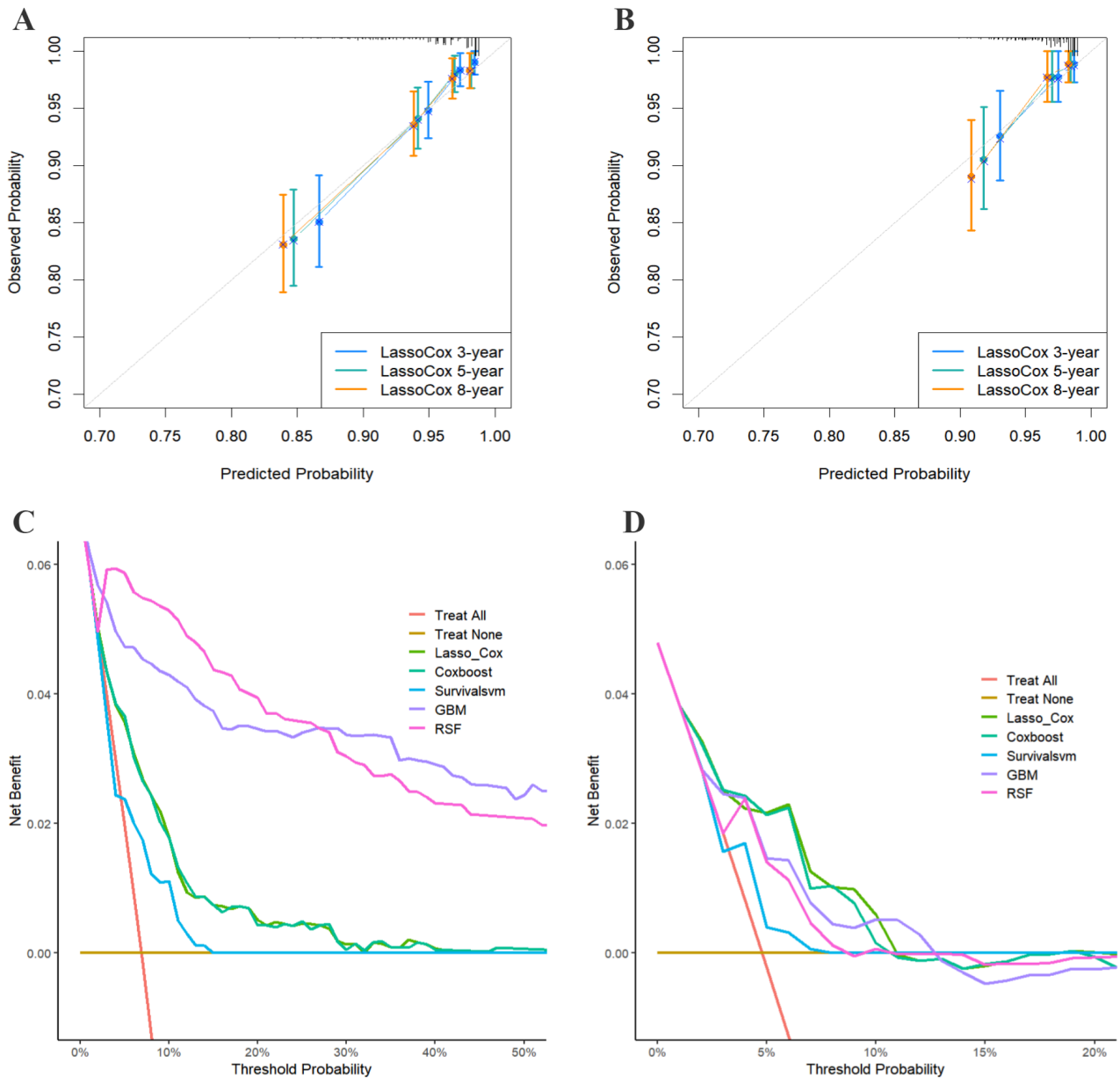


Figure 3

Calibration curves of the LASSO-COX model for the training set and validation set over 3-year, 5-year, and 8-year timeframes. A: Training set; B: Validation set. DCA curves of the five machine-learning models in the training set and validation set. C: Training set; D: Validation set.

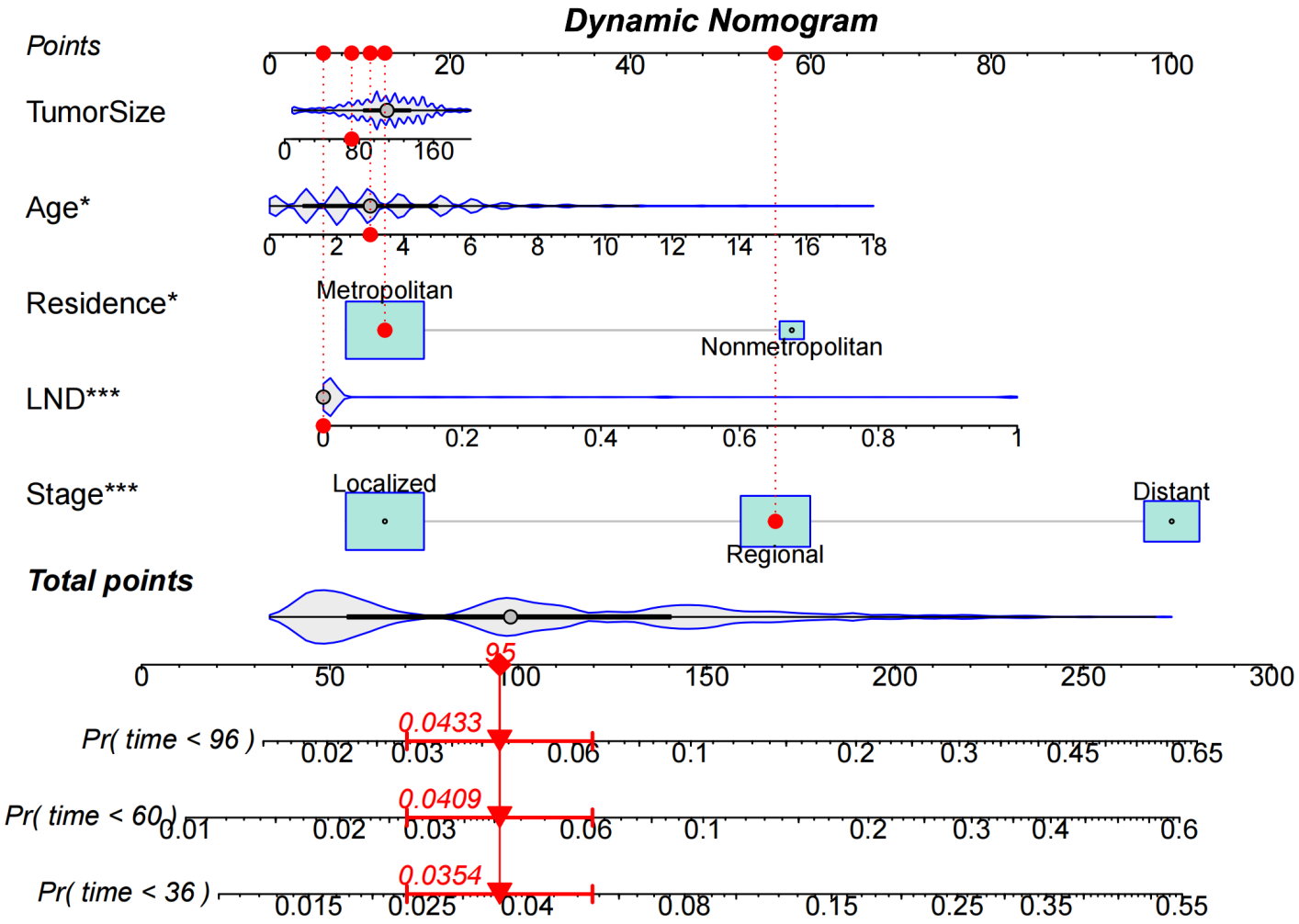


Figure 4

Dynamic Nomogram: predict the 3-year, 5-year, and 8-year Survival Probabilities of Children with Wilms Tumor