Research Article

A Nomogram-Based Risk Classification System Predicting the Overall Survival of Childhood with Clear Cell Sarcoma of the Kidney Based on the SEER Database

Jingchao Ding, Huanyin Yao, and Qingchun Chen

Department of Pediatric Internal Medicine, Shaoxing People’s Hospital, Shaoxing 312000, China

Correspondence should be addressed to Jingchao Ding; dijc1986@163.com

Received 1 June 2022; Accepted 16 July 2022; Published 30 August 2022

Copyright © 2022 Jingchao Ding et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Objective. Clear cell sarcoma of the kidney (CCSK) is a lethal pediatric renal malignancy with poor prognosis. A prognostic nomogram needs to be established for overall survival (OS) prediction of patients with CCSK.

Methods. Eligible 2588 CCSK patients (age 0–19) diagnosed between 2000 and 2017 were extracted from the Surveillance, Epidemiology, and End Results (SEER) database. Patients were randomized into training and validation cohorts (7:3). Independent prognostic factors were identified by univariate and multifactorial Cox regression analyses and used to construct a nomogram. Receiver operating characteristics (ROC) analysis, calibration curves, and decision curve analysis (DCA) were used to validate the nomogram. Moreover, a risk classification system was established based on the risk scores of the nomogram.

Results. Cox analyses revealed that age, combined stage, and origin were most significant prognostic factors. Based on these prognostic factors, a nomogram was established for predicting 3- and 5-year OS of patients with CCSK. The area under the ROC curve (AUC) of 3- and 5-year OS was 0.733 and 0.728 in the training cohort, corresponding to 0.69 and 0.674 in the validation cohort. The C-index of calibration curves in the training and validation cohorts was 0.724 and 0.686. DCAs indicated the clinical utility of this nomogram. A risk classification system stratified CCSK patients into three different risk cohorts. OS time of low-, intermediate-, and high-risk patients was 76, 68, and 65 months in the training cohort, corresponding to 69.5, 66, and 72 months in the validation cohort.

Conclusion. A nomogram-based risk classification system has high accuracy for the prognostic prediction of CCSK.

1. Introduction

Clear cell sarcoma of the kidney (CCSK) is one of the most common pediatric renal malignancies, accounting for approximately 5% of all primary childhood kidney tumors [1, 2]. The common therapeutic strategies for CCSK are surgical resection and adjuvant chemotherapy; however, the effects are still unsatisfactory due to late relapses and distant metastasis [2]. Currently, main therapy for CCSK depends on doxorubicin that is accompanied by cardiotoxic side effects [3]. Up to now, there is little information on clinical and histological features of CCSK; therefore, it is difficult for pathologic diagnosis and therapeutic development [4]. Timely detection and prognosis are of paramount importance for CCSK treatment. Thus, it is meaningful and urgent to construct a prognostic model for better treatment of CCSK.

The Surveillance, Epidemiology, and End Results (SEER) database includes basically comprehensive cancer statistics, which is an authoritative source in the United States. SEER database records the demographics, diagnosis, tumor characteristics, survival records, and therapies of patients with malignant tumors [5]. Nomogram is a reliable tool to predict the overall survival (OS) of patients with cancers, which can be established using the SEER database. Nomograms have been widely applied to identify potential prognostic factors that are associated with OS in multiple cancers, such as bladder, gastric, and colorectal cancers [6–8]. Zhang et al. constructed a reliable prognostic nomogram supporting the assessment of OS in bladder cancer patients based on the SEER database [6]. Yu Zhang established the nomograms of colorectal-cancer patients that exhibit favorable clinical values in predicting OS and cancer-
specific survival [8]. Therefore, a nomogram can effectively predict the OS and clinical prognosis of patients with CCSK based on the SEER database.

Based on this background, we explored potential risk parameters associated with CCSK using the SEER database. Furthermore, a prognostic nomogram was constructed to predict 3- and 5-year OS time of CCSK patients. This research provides a useful model to predict survival state of CCSK patients and apply in clinical therapy.

### 2. Materials and Methods

#### 2.1. Data Source

The clinical data of CCSK patients were downloaded from the SEER database (https://seer.cancer.gov/) using the SEER*Stat program (v 8.3.8, National Cancer Institute, MD, USA) [9]. A total of 2588 CCSK patients who were diagnosed from 2000 to 2017 were enrolled. The enrollment criteria of CCSK patients were as follows: [1] patients were 0–19 years old; [2] the diagnosis of CCSK was a positive confirmation; [3] and there was complete survival information. Eligible patients were randomly allocated into the training cohort and the validation cohort (7:3). The prognostic factors, including age, gender, race, origin, laterality, combined stage, radiation, and chemotherapy, were selected for further analyses.

#### 2.2. Nomogram Construction

The associations between clinicopathologic variables and OS were assessed using univariate and multivariate Cox regression analyses, and the results were visualized using forestplot in R package [10]. Hazard ratios (HRs) were presented with 95% confidence interval (CI). P values < 0.05 were identified as final independent risk factors to predict the OS in the multivariate Cox regression analysis. Subsequently, these independent risk factors were applied to construct a nomogram for the prediction of 3- and 5-year OS of CCSK patients. The nomogram was established using regression modeling strategies (BMS) in R package and visualized using ggplot [11, 12].

#### 2.3. Nomogram Validation

The validation of the nomogram was performed by the receiver operating characteristic (ROC) analysis that was achieved using survivalROC in R package [13]. An area-under-the-ROC-curve (AUC) value more than 0.65 suggests the perfect prediction of the nomogram. Calibration curve analysis was achieved using RMS

### Table 1: Baseline characteristics of CCSK patients.

| Characteristic          | Training cohort (n, %) | Validation cohort (n, %) | P value |
|-------------------------|------------------------|--------------------------|---------|
| Age (years)             |                        |                          | 0.1694  |
| Median (IQR)            | 3 (0–19)               |                          |         |
| <3.44                   | 1582 (61.1)            | 1122 (61.9)              |         |
| >3.44                   | 1006 (38.9)            | 690 (38.1)               |         |
| Gender                  |                        |                          | 0.9718  |
| Male                    | 1222 (47.2)            | 856 (47.2)               |         |
| Female                  | 1366 (52.8)            | 956 (52.8)               |         |
| Race                    |                        |                          | 0.8253  |
| White                   | 1955 (75.5)            | 1369 (75.6)              |         |
| Black                   | 448 (17.3)             | 312 (17.2)               |         |
| Asian or Pacific Islander| 121 (4.7)              | 87 (4.8)                 |         |
| American Indian/Alaska native | 34 (1.3)             | 24 (1.3)                 |         |
| Origin                  |                        |                          | 0.9231  |
| Spanish-Hispanic-Latino | 687 (26.5)             | 482 (26.6)               |         |
| Non-Spanish-Hispanic-Latino | 1901 (73.5)         | 1330 (73.4)              |         |
| Laterality              |                        |                          | 0.793   |
| Right                   | 1173 (45.3)            | 823 (45.4)               |         |
| Left                    | 1216 (47.0)            | 852 (47.0)               |         |
| Bilateral               | 183 (7.1)              | 126 (4.9)                |         |
| Combined stage          |                        |                          | 0.08337 |
| Distant                 | 459 (17.7)             | 311 (17.2)               |         |
| Localized               | 816 (31.5)             | 567 (31.3)               |         |
| Regional                | 605 (23.4)             | 439 (17.0)               |         |
| Radiation               |                        |                          | 0.8683  |
| Beam                    | 1201 (46.4)            | 840 (46.4)               |         |
| None                    | 1361 (52.6)            | 956 (52.8)               |         |
| Chemotherapy            |                        |                          | 0.658   |
| Yes                     | 2348 (90.7)            | 1641 (90.6)              |         |
| No                      | 240 (9.3)              | 171 (9.4)                |         |

CCSK, clear cell sarcoma of the kidney; IQR, interquartile range.
in R package [14]. A 45-degree diagonal line represented a perfect prediction. Besides, decision curve analysis (DCA) was performed using ggDCA in R package to assess the clinical value of the nomogram [15].

2.4. Risk Classification System and Survival Analyses. The risk classification system was constructed based on the risk scores of each CCSK patient calculated by the nomogram. The cutoff values were determined using the X-tile program.
Figure 2: A nomogram predicting 3- and 5-year overall survival (OS) of CCSK patients in the training cohort.

Figure 3: Receiver operating characteristic (ROC) curves of the nomogram for predicting the 3- and 5-year OS of CCSK patients. (a, b) ROC curves in the training cohort. (c, d) ROC curves in the validation cohort.
Moreover, survival analysis was performed according to the risk classification system using the survival and survminer in R package [16, 17].

2.5. Statistical Analyses. All data were analyzed using the SPSS version 27.0 (IBM, NY, USA). Two-tailed P value <0.05 was statistical difference.

3. Results

3.1. Baseline Characteristics of CCSK Patients. A total of 2588 CCSK patients (0–19 years old) were identified using the SEER database from 2000 to 2017. Patients were divided into the training and validation cohorts (7:3). We identified eight major clinicopathologic characteristics of CCSK patients, including age, gender, race, origin, laterality, combined stage, radiation, and chemotherapy. There was no significant difference in the distribution of these clinical characteristics between the training cohort and validation cohort (P>0.05) (Table 1).

3.2. Independent Prognostic Factors Correlated with OS of CCSK Patients. Hereafter, univariate and multivariate Cox regression analyses were performed to identify significant risk factors of CCSK in the training cohort. Univariate Cox regression analysis revealed that age (P = 7.8e-06), origin (P = 0.0047), combined stage (P < 0.001), and radiation (P = 5.38e-05) were the significant risk factors related to OS of CCSK patients (Figure 1(a)). Multivariate Cox regression analysis further confirmed that age (P = 0.0035), origin (P = 0.025), and combined stage (P < 0.001) were independent prognostic factors of CCSK survival (Figure 1(b)).

3.3. Nomogram Establishment and Validation. A nomogram model for predicting 3- and 5-year OS was constructed, which integrated all three independent prognostic parameters (age, origin, and combined stage). Age (nomogram score range: 0–100) was the most important independent prognostic parameter for evaluating OS, followed by combined stage (0–97.5) and origin (0–30) (Figure 2). Furthermore, ROC analysis was performed for nomogram
validation. The AUC values of 3- and 5-year were 0.733 and
0.728, respectively, in the training cohort, corresponding to
0.690 and 0.674 in the validation cohort (Figures 3(a)–3(d)).
The C-index values of the training and validation cohorts
were 0.724 and 0.686, respectively. Calibration curve
exhibited satisfactory agreement between actual and pre-
diction survival in CCSK patients (Figures 4(a)–4(d)). DCA
showed that this nomogram model has great benefits for
predicting 3- and 5-year OS time of patients with CCSK
within all of threshold probabilities and presented more
positive net benefit than the “all” or “none” strategies be-
tween the training cohort and validation cohort
(Figures 5(a), 5(b)).

3.4. Risk Classification System. Enrolled patients with CCSK
were divided into three groups: low-, intermediate-, and
high-risk groups. The corresponding cutoff points for risk
groups were classified as follows: $<29.28, 29.28 \leq \text{nomogram}
\text{score} \leq 60.71$, and $\geq 60.71$. In the training cohort, the median
OS time of CCSK patients in the low-, intermediate-, and
high-risk groups was 76.0, 68.0, and 65.0 months, respecti-
vely. In the validation cohort, the median OS time of CCSK
patients in the low-, intermediate-, and high-risk groups was
69.5, 66.0 months, and 72 months, respectively.
Significant OS differences were exhibited among the
training, validation, and entire cohorts ($P < 0.05$)
(Figures 6(a)–6(c)).

4. Discussion
CCSK is an aggressive renal malignancy seen in children
with poor prognosis [18]. Nomogram is an effective clinical
decision-making tool for the prognostic prediction of CCSK
patients. In this study, three significant risk factors, in-
cluding age, origin, and combined stage, were identified
based on the SEER database. Nomograms were constructed
and validated for the prediction of 3- and 5-year OS of CCSK
patients, and we found that age was the most important
prognostic factor correlated with OS. Furthermore, a risk
classification system was established, and the OS time among
low-, intermediate-, and high-risk CCSK patients exhibited
significant differences.

The prognostic model exhibited in this study was based
on 2588 CCSK patients (aged 0–19) included in the SEER
database from 2000–2017. We evaluated eight potential
prognostic factors related to OS of CCSK patients, including
age, gender, race, origin, laterality, combined stage, radia-
tion, and chemotherapy. Through the univariate and mul-
tivariate Cox regression analyses, we demonstrated that age,
origin, and combined stage were the significant independent
prognostic factors closely related with the OS of CCSK
patients. It has been reported that young age is a significant
unfavorable prognostic factor for CCSK patient survival
[19]. Seibel et al. indicated that stage is highly predictive for
survival outcome of CCSK patients [20]. Combined with
previous studies, we further confirmed that age, origin, and
combined stage are three major prognostic factors for predicting OS of CCSK patients. Nomogram is a graphic representation widely applied to depict a statistical prognostic model for clinical events and OS in patients with cancer. Zheng et al. developed and validated a nomogram that can postoperatively evaluate OS and cancer-specific survival of patients with pediatric adren al cancer [21]. Liu et al. established a novel nomogram with favorable discrimination ability to predict prognosis for newly diagnosed pediatric patients with atypical teratoid/ rhabdoid tumors [22]. However, to our knowledge, there is no reliable nomogram for CCSK prognosis. In the present study, we constructed a promising nomogram for predicting the 3- and 5-year OS of CCSK patients based on three independent prognostic factors (age, origin, and combined stage). The nomogram illuminated that age (nomogram score 0–100) is the most significant parameter for CCSK prognosis, followed by combined stage (0–97.5), whereas origin (0–30) presents limited impact on OS outcomes. Subsequently, the prognostic nomogram was validated via
ROC analysis. AUC of ROC curve indicates the discrimination ability of a prognostic nomogram [23]. The AUC values of 3- and 5-year ROC curves were, respectively, 0.733 and 0.728 in the training cohort (vs. 0.69 and 0.674 in the validation cohort). This result suggests that the nomogram we established is a stable and reliable prognostic model for CCSK patients. In addition, calibration curves are utilized to evaluate whether the nomogram-predicted survival is consistent with the actual survival of CCSK patients [24]. The C-indexes of calibration curves of the training and validation cohorts were 0.724 and 0.686, respectively, indicating that there is excellent agreement between the nomogram-predicted and the actual survival of CCSK patients. DCA further evaluated the clinical utility of the prognostic nomogram and confirmed its viability and accuracy.

Furthermore, a risk classification system was established according to the nomogram risk scores from each CCSK patient. The median OS time in patients with low-risk was 76 months, whereas the OS time of high-risk patients was one year less than that of low-risk patients. This finding indicates that older age and distant stage present a poor survival outcome for high-risk CCSK patients. Indeed, CCSK has the propensity of distant metastasis to other sites, including the bone, brain, lymph nodes, lungs, and liver [25]. Therefore, age and distant stage can be as the independent prognostic factors for predicting the OS of CCSK patients, which provides new guidance for CCSK treatment.

5. Conclusion
In summary, a novel and reliable nomogram was established and validated for the prediction of the 3- and 5-year OS time of CCSK patients based on eight potential prognostic factors from the SEER database between 2000 and 2017. Moreover, a risk classification system was established to stratify patients with CCSK into three different risk cohorts, and older age and distant stage in the high-risk patients mean a poor survival outcome. This prognostic model provides bright prospects for survival prediction of CCSK patients and can be applied in clinical practice. The nomogram established in this study can be applied for early diagnostic and prognostic prediction for CCSK, thereby assisting clinicians in making individualized decisions. However, the present study needs to further identify other potential risk factors, such as chemotherapy, radiotherapy, and genetic characteristics. The nomogram we constructed should include more detailed clinical factors for survival prediction of CCSK patients.

Data Availability
The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest
The authors declare that there are no conflicts of interest regarding the publication of this paper.

References
[1] S. J. Aw and K. T. E. Chang, “Clear cell sarcoma of the kidney,” Archives of Pathology and Laboratory Medicine, vol. 143, pp. 1022–1026, 2019.
[2] A. P. Aldera and K. Pillay, “Clear cell sarcoma of the kidney,” Archives of Pathology and Laboratory Medicine, vol. 144, pp. 119–123, 2020.
[3] C. Kenny, N. McDonagh, A. Lazaro et al., “Dysregulated mitogen-activated protein kinase signalling as an oncogenic basis for clear cell sarcoma of the kidney,” The Journal of Pathology, vol. 244, no. 3, pp. 334–345, 2018.
[4] M. Cao, J. Zhang, H. Ma, and Y. Liang, “Clear cell sarcoma of the kidney in an adult: a case report and literature review,” Translational Cancer Research, vol. 11, no. 1, pp. 288–294, 2022.
[5] T. M. Kuo and L. R. Mobley, “How generalizable are the seer registries to the cancer populations of the USA?” Cancer Causes and Control, vol. 27, no. 9, pp. 1117–1126, 2016.
[6] Y. Zhang, Y. K. Hong, D. W. Zhuang, X. J. He, and M. E. Lin, “Bladder cancer survival nomogram: development and validation of a prediction tool, using the seer and tca databases,” Medicine (Baltimore), vol. 98, no. 44, Article ID e17725, 2019.
[7] Y. Zhang and C. Yu, “Development and validation of a surveillance, epidemiology, and end results (seer)-based prognostic nomogram for predicting survival in elderly patients with gastric cancer after surgery,” Journal of Gastrointestinal Oncology, vol. 12, no. 2, pp. 278–296, 2021.
[8] C. Yu and Y. Zhang, “Establishment of prognostic nomogram for elderly colorectal cancer patients: a seer database analysis,” BMC Gastroenterology, vol. 20, no. 1, p. 347, 2020.
[9] S. Davis and P. S. Meltzer, "GEOquery: a bridge between the gene expression omnibus (GEO) and BioConductor," *Bioinformatics*, vol. 23, no. 14, pp. 1846-1847, 2007.
[10] H. Jiang, Y. Li, and T. Wang. “Comparison of billroth i, billroth ii, and roux-en-y reconstructions following distal gastrectomy: a systematic review and network meta-analysis,” Cirugia Espanola, vol. 99, no. 6, pp. 412–420, 2021.
[11] Z. Huang, Y. Tong, H. Tian, and C. Zhao, “Establishment of a prognostic nomogram for lung adenocarcinoma with brain metastases,” World Neurosurgery, vol. 141, pp. e700–e709, 2020.
[12] Y. Wang, C. Mi, and Y. Guo, “Satellite tracking reveals a new migration route of black-necked cranes (grus nigricollis) in qinghai-tibet plateau,” PeerJ, vol. 8, Article ID e9715, 2020.
[13] W. Weng, X. Chen, S. Gong, L. Guo, and X. Zhang, “Pre-operative neutrophil-lymphocyte ratio correlated with glioma grading and glioblastoma survival,” Neurological Research, vol. 40, no. 11, pp. 917–922, 2018.
[14] B. Van Calster, D. Nieboer, Y. Vergouwe, B. De Cock, M. J. Pencina, and E. W. Steyerberg, “A calibration hierarchy for risk models was defined: from utopia to empirical data,” Journal of Clinical Epidemiology, vol. 74, pp. 167–176, 2016.
[15] T. Fu, S. Coulter, E. Yoshihara et al., “Fxr regulates intestinal cancer stem cell proliferation,” Cell, vol. 176, no. 5, pp. 1098–1112.e18, 2019.
[16] N. Tong and C. C. L. Wyatt, “Five-year survival rate of bonded dental restorations in frail older adults,” JDR Clinical and Translational Research, vol. 6, no. 1, pp. 77–86, 2021.
[17] L. Zeng, X. Fan, X. Wang et al., “Bioinformatics analysis based on multiple databases identifies hub genes associated with hepatocellular carcinoma,” Current Genomics, vol. 20, no. 5, pp. 349–361, 2019.
W. Zekri, A. S. Alfaar, D. Yehia et al., “Clear cell sarcoma of the kidney: patients’ characteristics and improved outcome in developing countries,” *Pediatric Blood and Cancer*, vol. 61, no. 12, pp. 2185–2190, 2014.

J. J. Dong, X. Y. He, X. Liu et al., “Retrospective analysis of outcomes in patients with clear cell sarcoma of the kidney: a tertiary single-institution experience,” *Journal of Pediatric Surgery*, vol. 56, no. 3, pp. 580–586, 2021.

N. L. Seibel, Y. Y. Chi, E. J. Perlman et al., “Impact of cyclophosphamide and etoposide on outcome of clear cell sarcoma of the kidney treated on the national wilms tumor study-5 (nwts-5),” *Pediatric Blood and Cancer*, vol. 66, no. 1, Article ID e27450, 2019.

J. Zheng, J. Cai, X. Diao et al., “Nomograms for the prediction of survival for patients with pediatric adrenal cancer after surgery,” *Journal of Cancer*, vol. 11, no. 8, pp. 2080–2090, 2020.

Y. Liu, X. Peng, and T. Zeng, “Development and validation of a nomogram for predicting overall survival in pediatric patients with atypical teratoid/rhabdoid tumors,” *Turkish Neurosurgery*, 2021.

Q. Q. Wei, Y. Chen, X. Chen et al., “Prognostic nomogram associated with longer survival in amyotrophic lateral sclerosis patients,” *Aging and Disease*, vol. 9, no. 6, pp. 965–975, 2018.

W. Yu, L. Huang, Z. Zhong et al., “A nomogram-based risk classification system predicting the overall survival of patients with newly diagnosed stage ivb cervix uteri carcinoma,” *Frontiers of Medicine*, vol. 8, Article ID 693567, 2021.

J. Weaver, T. Ho, A. Lang et al., “Bladder recurrence of clear cell sarcoma of the kidney seven years after initial presentation,” *Urology*, vol. 106, pp. 193–195, 2017.