



OPEN ACCESS

EDITED BY
Darius Mehregan,
Wayne State University, United States

REVIEWED BY
Linjun Shi,
Shanghai Jiao Tong University, China
Peizhen Zhao,
Southern Medical University, China

*CORRESPONDENCE
Jun Lyu
✉ lyujun2020@jnu.edu.cn
Liehua Deng
✉ liehuadeng@126.com

†These authors have contributed equally
to this work and share first authorship

SPECIALTY SECTION
This article was submitted to
Dermatology,
a section of the journal
Frontiers in Medicine

RECEIVED 28 October 2022
ACCEPTED 12 January 2023
PUBLISHED 17 February 2023

CITATION
Zhang J, Yang W, Lian C, Zhao Q, Ming W-k,
Ip CC, Mu H-H, Ching Tom K, Lyu J and Deng L
(2023) A nomogram for predicting survival
in patients with skin non-keratinizing large cell
squamous cell carcinoma: A study based on
the Surveillance, Epidemiology, and End
Results database.
Front. Med. 10:1082402.
doi: 10.3389/fmed.2023.1082402

COPYRIGHT
© 2023 Zhang, Yang, Lian, Zhao, Ming, Ip, Mu,
Ching Tom, Lyu and Deng. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in
other forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

A nomogram for predicting survival in patients with skin non-keratinizing large cell squamous cell carcinoma: A study based on the Surveillance, Epidemiology, and End Results database

Jinrong Zhang^{1†}, Wei Yang^{2†}, Chengxiang Lian¹, Qiqi Zhao^{1,3},
Wai-kit Ming⁴, Cheong Cheong Ip^{1,5}, Hsin-Hua Mu⁶,
Kong Ching Tom⁷, Jun Lyu^{8*} and Liehua Deng^{1,3*}

¹Department of Dermatology, The First Affiliated Hospital of Jinan University and Jinan University Institute of Dermatology, Guangzhou, China, ²Office of Drug Clinical Trial Institution, The First Affiliated Hospital of Jinan University, Guangzhou, China, ³Department of Dermatology, The Fifth Affiliated Hospital of Jinan University, Heyuan, China, ⁴Department of Infectious Diseases and Public Health, Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Hong Kong, Hong Kong SAR, China, ⁵Department of Dermatology, University Hospital Macau, Macau, Macao SAR, China, ⁶General Surgery Breast Medical Center, Taipei Medical University Hospital, Taipei City, China, ⁷Primax Biotech Company, Hong Kong, Hong Kong SAR, China, ⁸Department of Clinical Research, The First Affiliated Hospital of Jinan University, Guangzhou, China

Introduction: This study aimed to develop and validate a nomogram for predicting cancer-specific survival (CSS) in patients with non-keratinized large cell squamous cell carcinoma (NKLCSCC) at 3, 5, and 8 years after diagnosis.

Methods: Data on SCC patients were collected from the Surveillance, Epidemiology, and End Results database. Training (70%) and validation (30%) cohorts were generated using random selection of patients. Independent prognostic factors were selected using the backward stepwise Cox regression model. To predict the CSS rates in patients with NKLCSCC at 3, 5, and 8 years after diagnosis, all of the factors were incorporated into the nomogram. Indicators such as the concordance index (C-index), area under the time-dependent receiver operating characteristic curve (AUC), net reclassification index (NRI), integrated discrimination improvement (IDI), calibration curve, and decision-curve analysis (DCA) were then used to validate the performance of the nomogram.

Results: This study included 9,811 patients with NKLCSCC. Twelve prognostic factors were identified by Cox regression analysis in the training cohort, which were age, number of regional nodes examined, number of positive regional nodes, sex, race, marital status, American Joint Committee on Cancer (AJCC) stage, surgery status, chemotherapy status, radiotherapy status, summary stage, and income. The constructed nomogram was validated both internally and externally. The nomogram had good discrimination ability, as indicated by the comparatively high C-indices and AUC values. The nomogram was properly calibrated, as indicated by the calibration

curves. Our nomogram was superior to the AJCC model, as illustrated by its superior NRI and IDI values. DCA curves indicated the clinical usability of the nomogram.

Conclusion: The first nomogram for prognosis predictions of patients with NKLCSCC has been developed and verified. Its performance and usability demonstrated that the nomogram could be utilized in clinical settings. However, additional external verification is still required.

KEYWORDS

Surveillance, Epidemiology, and End Results, cancer-specific survival, nomogram, non-keratinizing large cell squamous cell carcinoma, SEER

Introduction

Cutaneous squamous cell carcinoma (cSCC) is the second most common type of non-melanoma skin cancer. It accounts for 20% of skin cancers, with 1 million cases and an estimated 9,000 deaths each year in the United States (1). The reported incidence of cSCC ranges from 5 to 499 per 100,000 patients (2–5). Among the non-Hispanic white population in the United States, the lifetime risk of developing SCC is 14–20% (6, 7). This risk has continued to increase each year, with an estimated increase of between 50 and 200%, and is likely to continue to increase due to the aging population (8). However, the current understanding of non-keratinized large cell squamous cell carcinoma (NKLCSCC) of the skin is inadequate, and its increasing incidence and distinct characteristics from other types of squamous cell carcinoma require it to be independently analyzed.

Some researchers have proposed that the most significant risk factors for cSCC include age, sex, race, and surrounding environment (9), but there is currently no definite prognostic implication for the subtypes of cSCC. The present study addressed the postoperative recurrence of NKLCSCC, which is one of the most common subtypes of squamous cell carcinoma.

A fundamental standard of care for cancer treatment is the standard American Joint Committee on Cancer (AJCC) staging system. But when used to predict the prognosis of NKLCSCC, the AJCC staging system is restricted by the lack of precise demographic and clinical characteristics. Providing clinicians with convenient and thorough guidance requires more detailed and extensive prediction models.

Nomograms are highly accurate and easy-to-use tools that are based on many types of tumor prediction models (10) and make it possible to estimate the survival probability of a specific patient. Many researchers have established nomograms for various cancers, including of the tonsils, parotid gland, and breast (11–13), but no nomogram specifically designed for NKLCSCC has been developed. Therefore we developed and evaluated a nomogram for NKLCSCC using pertinent data from the Surveillance, Epidemiology, and End Results (SEER) database to further investigate the prognosis factors for NKLCSCC and its specific treatment.

We aimed to develop a comprehensive nomogram for patients with NKLCSCC in the SEER database that accounted for key demographics, clinicopathological features, and therapeutic approaches in addition to some fundamental traits. We analyzed the treatments applicable to these patients. Our novel nomogram can provide clinicians with more-thorough and personalized patient survival predictions, which makes it clinically superior to conventional methods.

Patients and methods

Data sources and research factors

Data were filtered and extracted from the SEER database using SEER*Stat software. Part of the SEER database is available to the public, and we requested additional access to the SEER Plus database (14). We collected NKLCSCC cases from the SEER database by adopting the ICD-O-3 (third revision of the International Classification of Diseases for Oncology) histology/behavior code for NKLCSCC (“8072/3: Squamous cell carcinoma, large cell, non-keratinizing, NOS”) and the cases where the site was cutaneous were selected.

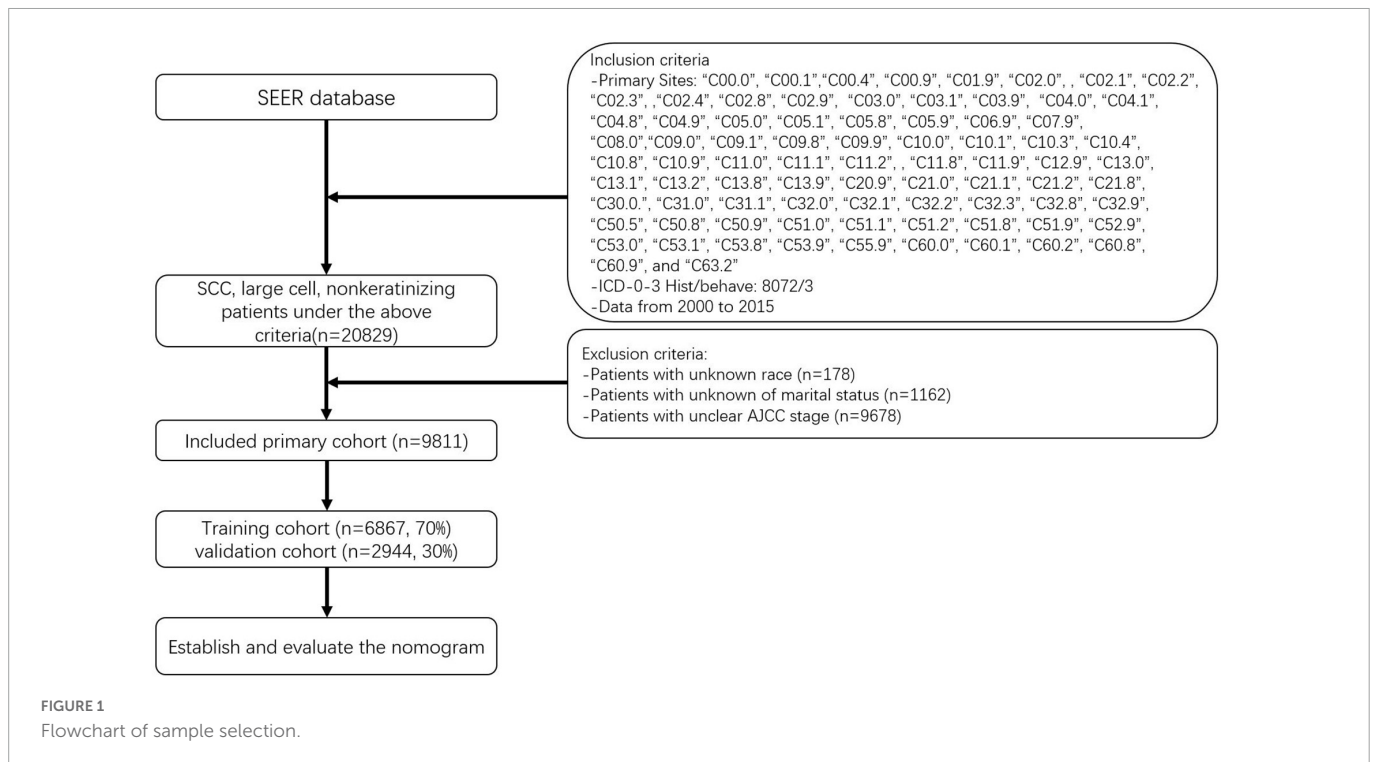
We selected several factors that may be relevant to disease prognosis, including age, race, sex, marital status, tumor grade, AJCC stage, income, number of regional nodes examined (RNE), number of positive regional nodes (RNP), summary stage, surgery status, radiotherapy status, and chemotherapy status. Tumor extension, lymph node metastasis, and distant metastasis as assessed using the TNM (tumor, node, and metastasis) staging system are all included in the AJCC staging system. Due to substantial multicollinearity caused by including all of these factors in the analysis, this study only used the AJCC staging system. Cancer-specific survival (CSS) was the outcome variable. Since the SEER database used in this study did not contain any personally identifying information, it was not necessary to obtain patient-informed permission.

The data of patients whose baseline and survival data were fully available were selected. The seventh edition of the AJCC staging system was adopted. Using the methods described above, we initially identified 20,839 patients with NKLCSCC between 2000 and 2015. After excluding 178 patients with unknown race, 1,162 with unknown marital status, and 9,678 with unknown AJCC stage, 9,811 NKLCSCC patients were finally included (15). These patients were randomly divided into training (70%) and validation (30%) cohorts to test the model using R software (version 4.2.0¹). **Figure 1** illustrates the data screening process.

Nomogram and statistical analysis

A log-rank test performed after assigning all subjects to the training and validation cohorts indicated no statistically significant differences between them. The initial baseline characteristics of

¹ <http://www.r-project.org>



each variable in the study cohort were then summarized using SPSS Statistics software (version 27.0, IBM SPSS, Chicago, IL, United States). Other variable data were presented as frequencies and percentages, while age at diagnosis was expressed as median and interquartile range (IQR) values. Nomograms were utilized to estimate the 3-, 5-, and 8-year CSS probabilities for NKLCSCC, and Cox regression was conducted to identify CSS factors related to NKLCSCC, with a significance cut-off of $p = 0.05$. After constructing the nomogram, we evaluated the model using a set of metrics. Two metrics were used to evaluate the discrimination capability of the nomogram: the concordance index (C-index) and the area under the time-dependent receiver operating characteristic (ROC) curve (AUC). Although AUC and C-index are often utilized, their improvements were not significant relative to the existing model. To ascertain if the new model was superior, we additionally used two relatively new metrics: net reclassification index (NRI) and integrated discriminant improvement (IDI). IDI accounts for numerous tangents that can be utilized to assess the overall performance of the model, and NRI is primarily used to evaluate the prediction capacity of old and new models at a certain tangent level (16, 17). These two markers are better understood in actual clinical applications.

A calibration plot was also constructed to graphically display the variation between the two values. The level of model calibration indicates the consistency of predicted and actual values. When the calibration curve is aligned with the 45° standard line, model consistency has advanced. Decision-curve analysis (DCA) curves were employed to evaluate the clinical validity of the model. The abscissa and ordinate of the DCA curve represent the threshold probability and net profit of the model, respectively. The net profit of a model will be higher if the DCA curve is higher (18).

The R software and SPSS Statistics were used to conduct all statistical analyses. SPSS Statistics were used to characterize the fundamental features of the cohort. R software was then used to

randomly divide the data into training and validation cohorts, and a log-rank test was conducted. The following R software packages were used: survival, rms, foreign, survival, survivalROC, nricens, and DCA. These analyses included Cox regression analysis, proportional-hazards regression, nomogram establishment, and assessment. Significance was defined as two-sided probability values of $p < 0.05$.

Results

General characteristics

After randomizing 9,811 patients into two cohorts, the log-rank test obtained a probability value ($p = 1$) that indicated no significant differences between them. The fundamental demographic and clinical characteristics of the two cohorts were then described using SPSS (Table 1). The median age at diagnosis was 57 years (IQR = 48–66 years) in the training cohort and 57 years (IQR = 48–65 years) in the validation cohort. The distributions of sex and surgery status were fairly even. Most patients in both the training and validation cohorts were white (78.8 and 79.8%, respectively) and married (56.8 and 55.1%). AJCC stage IV was observed in most cases. Most patients were in the regional cancer summary stage. Most patients were treated with radiotherapy, chemotherapy, and earned US\$ 60,000–74,999 per year.

Constructing a nomogram using the training cohort

Age at diagnosis, RNE, RNP, sex, race, marital status, AJCC stage, surgery status, chemotherapy status, radiotherapy status, summary

TABLE 1 Demographic and clinical characteristics of the two cohorts of patients.

Variable	Training cohort (%)	Validation cohort (%)
N	6,867	2,944
Age of diagnosis	57 (48–66)	57 (48–65)
Sex		
Male	3,389 (49.4)	1,483 (50.4)
Female	3,478 (50.6)	1,461 (49.6)
Race		
White	5,413 (78.8)	2,349 (79.8)
Black	647 (9.4)	247 (8.4)
Other	807 (11.8)	348 (11.8)
Marital status		
Single	1,502 (21.9)	655 (22.2)
Married	3,901 (56.8)	1,623 (55.1)
DSW	1,464 (21.3)	666 (22.6)
AJCC stage		
I	1,484 (21.6)	610 (20.7)
II	904 (13.2)	396 (13.5)
III	1,583 (23.1)	697 (23.7)
IV	2,896 (42.2)	1,241 (42.2)
Summary of stage		
Localized	1,727 (25.1)	723 (24.6)
Regional	3,763 (54.8)	1,632 (55.4)
Distant	1,377 (20.1)	589 (20)
Radiation		
Yes	5,318 (77.4)	2,257 (76.7)
No/unknown	1,549 (22.6)	687 (23.3)
Chemotherapy		
Yes	4,381 (63.8)	1,868 (63.5)
No/unknown	2,486 (36.2)	1,076 (26.5)
Income		
<\$35,000, \$35,000–\$44,999	542 (7.9)	237 (8.1)
\$45,000–\$59,999	1,456 (21.2)	615 (20.9)
\$60,000–\$74,999	2,888 (42.1)	1,283 (43.6)
\$75,000+	1,981 (28.8)	809 (27.5)
Surgery		
Yes	3,166 (46.1)	1,367 (46.4)
No/unknown	3,701 (53.1)	1,577 (53.6)

stage, and income were the 12 variables that were included after performing multivariate Cox stepwise regression ($p < 0.05$). **Table 2** lists the factors that were found to be significant following the multivariate Cox regression analysis, which were age at diagnosis (hazard ratio [HR] = 1.031, $p < 0.0001$), RNE (HR = 0.997, $p < 0.0001$), RNP (HR = 1.003, $p < 0.0001$), sex (HR = 1.237, $p < 0.0001$), black race (versus white: HR = 1.243, $p < 0.0001$), married (versus single: HR = 0.697, $p < 0.0001$), AJCC stage II (versus stage I: HR = 1.429, $p < 0.0001$), AJCC stage III (versus stage I:

HR = 1.937, $p < 0.0001$), AJCC stage IV (versus stage I: HR = 2.095, $p < 0.0001$), distant summary stage (versus localized: HR = 2.343, $p < 0.0001$), no/unknown radiotherapy status (versus radiotherapy: HR = 1.702, $p < 0.0001$), no/unknown chemotherapy status (versus chemotherapy: HR = 1.268, $p < 0.0001$), no/unknown surgery status (versus surgery: HR = 1.428, $p < 0.0001$), and income of \$75,000+

TABLE 2 Selected variables by multivariate Cox stepwise regression analysis.

Variable	Multivariate analysis		
	HR	95% CI	p-Value
Age of diagnosis	1.031	1.028–1.033	<0.0001
RNE	0.997	0.995–0.998	<0.0001
RNP	1.003	1.002–1.004	<0.0001
Sex			
Male	Reference		
Female	1.237	1.154–1.326	<0.0001
Race			
White	Reference		
Black	1.243	1.127–1.371	<0.0001
Other	0.977	0.884–1.08	0.6478
Marital status			
Single	Reference		
Married	0.697	0.644–0.755	<0.0001
DSW	0.998	0.912–1.092	0.9641
AJCC stage			
I	Reference		
II	1.429	1.22–1.674	<0.0001
III	1.937	1.616–2.321	<0.0001
IV	2.095	1.734–2.532	<0.0001
Summary of stage			
Localized	Reference		
Regional	1.163	0.995–1.359	0.0572
Distant	2.343	1.968–2.79	<0.0001
Radiation			
Yes	Reference		
No/unknown	1.702	1.554–1.865	<0.0001
Chemotherapy			
Yes	Reference		
No/unknown	1.268	1.167–1.379	<0.0001
Income			
<\$35,000, \$35,000–\$44,999	Reference		
\$45,000–\$59,999	0.891	0.791–1.004	0.0573
\$60,000–\$74,999	0.823	0.736–0.919	0.0006
\$75,000+	0.738	0.655–0.831	<0.0001
Surgery			
Yes	Reference		
No/unknown	1.428	1.317–1.547	<0.0001

per year (versus <US\$ 35,000 and US\$ 35,000–44,999: HR = 0.738, $p < 0.0001$).

The finally constructed nomogram is shown in **Figure 2**. Based on the relevant factors stated above, the multiple regression model of the nomogram may be utilized to predict CSS probabilities. According to **Figure 2**, the summary stage had the greatest effect on survival rate, followed by AJCC stage, radiotherapy status, surgery status, chemotherapy status, race, sex, age at diagnosis, RNP, marital status, and income. Each component is represented as a line segment on the nomogram, and the numerical scale of the line defines the risk level presented by that factor. The sum of the scores for all of the criteria for each patient produces a total score that corresponds to their 3-, 5-, and 8-year CSS probabilities.

Evaluating the nomogram using the validation cohort

The C-index in the nomogram model was 0.710 for the training cohort and 0.725 for the validation cohort. The C-index for AJCC training cohort is 0.595, and validation cohort is 0.599. The new model's C-index outranks the AJCC for training cohort by 0.115, and for validation cohort by 0.126. The AUCs at years 3, 5, and 8 were 0.739, 0.732, and 0.745, respectively, for the training cohort, and 0.766, 0.751, and 0.759 for the validation cohort (**Figure 3**).

The discrimination ability of the nomogram was evaluated using the NRI and IDI. The NRI values for the 3-, 5-, and 8-year CSS

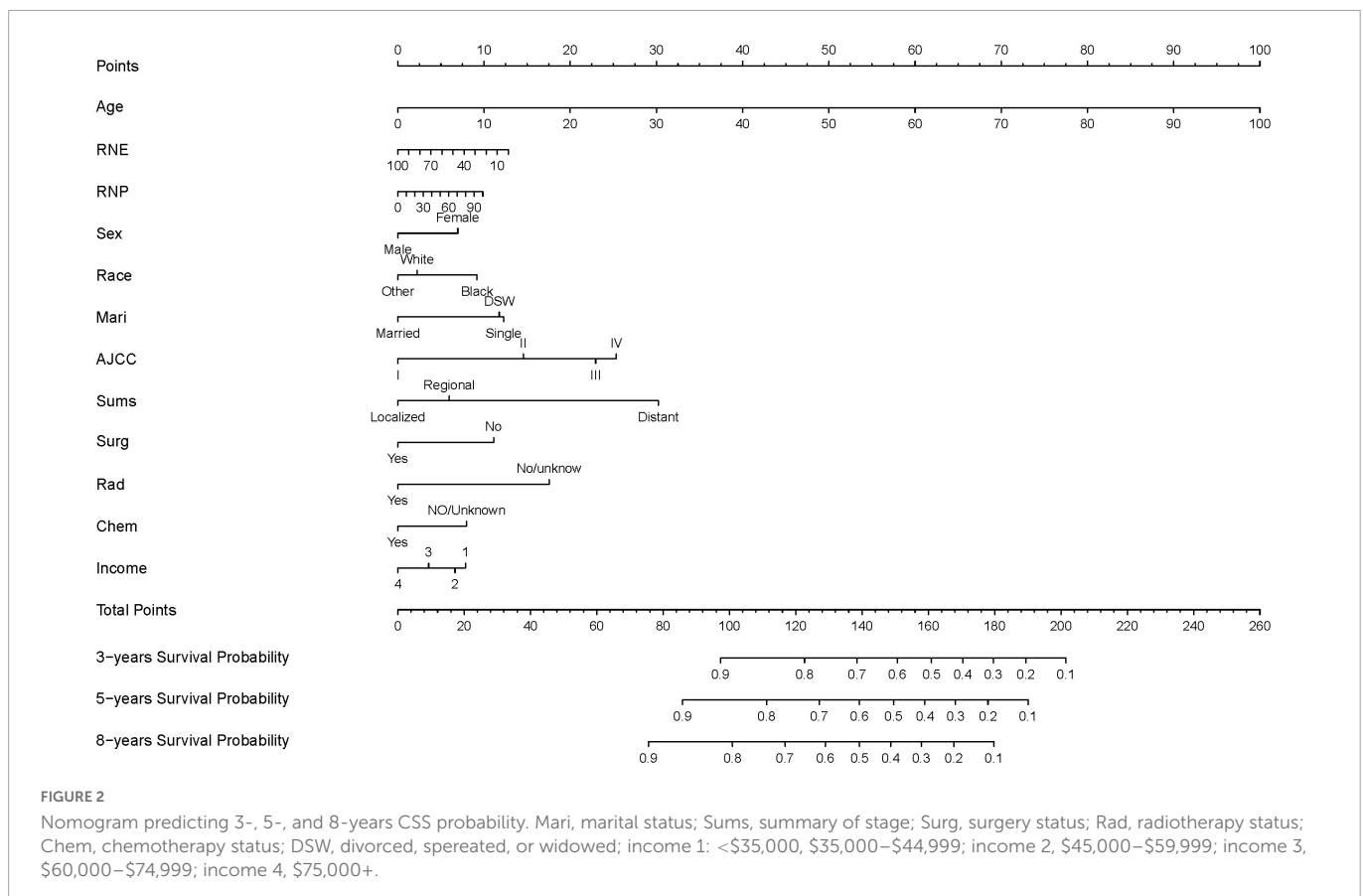
probabilities were 0.568 (95% confidence interval [CI] = 0.514–0.625), 0.581 (95% CI = 0.542–0.630), and 0.625 (95% CI = 0.587–0.682), respectively, for the training cohort, and 0.708 (95% CI = 0.623–0.796), 0.689 (95% CI = 0.612–0.762), and 0.748 (95% CI = 0.666–0.817) for the validation cohort. The IDI values for the 3-, 5-, and 8-year CSS probabilities were 0.125, 0.142, and 0.151, respectively, for the training cohort ($p = 0.001$), and 0.152, 0.158, and 0.168 for the validation cohort ($p = 0.001$).

In order to test between the real and ideal values of the model, the calibration plot was first utilized to confirm the discrimination ability of the model. The calibration plots of 3-, 5-, and 8-year CSS probabilities in the model were very close to the standard line (**Figure 4**), demonstrating that it had a high level of calibration.

Finally, a DCA curve was drawn to demonstrate the clinical validity of the nomogram. The survival probability curves for the new model were all higher than those for the AJCC model (**Figure 5**), indicating that using the new method to predict 3-, 5-, and 8-year CSS probabilities is superior overall.

Discussion

The newest guideline by the American Academy of Dermatology in 2018 paid more attention to individual factors in assessing NKLCSCC (19). Other classifications focus on molecular and immunohistochemical information, and biopsy techniques (20, 21). However, these current guidelines do not provide definite prognostic predictions for NKLCSCC. This study was therefore designed to analyze the combined prognostic factors for NKLCSCC in detail for



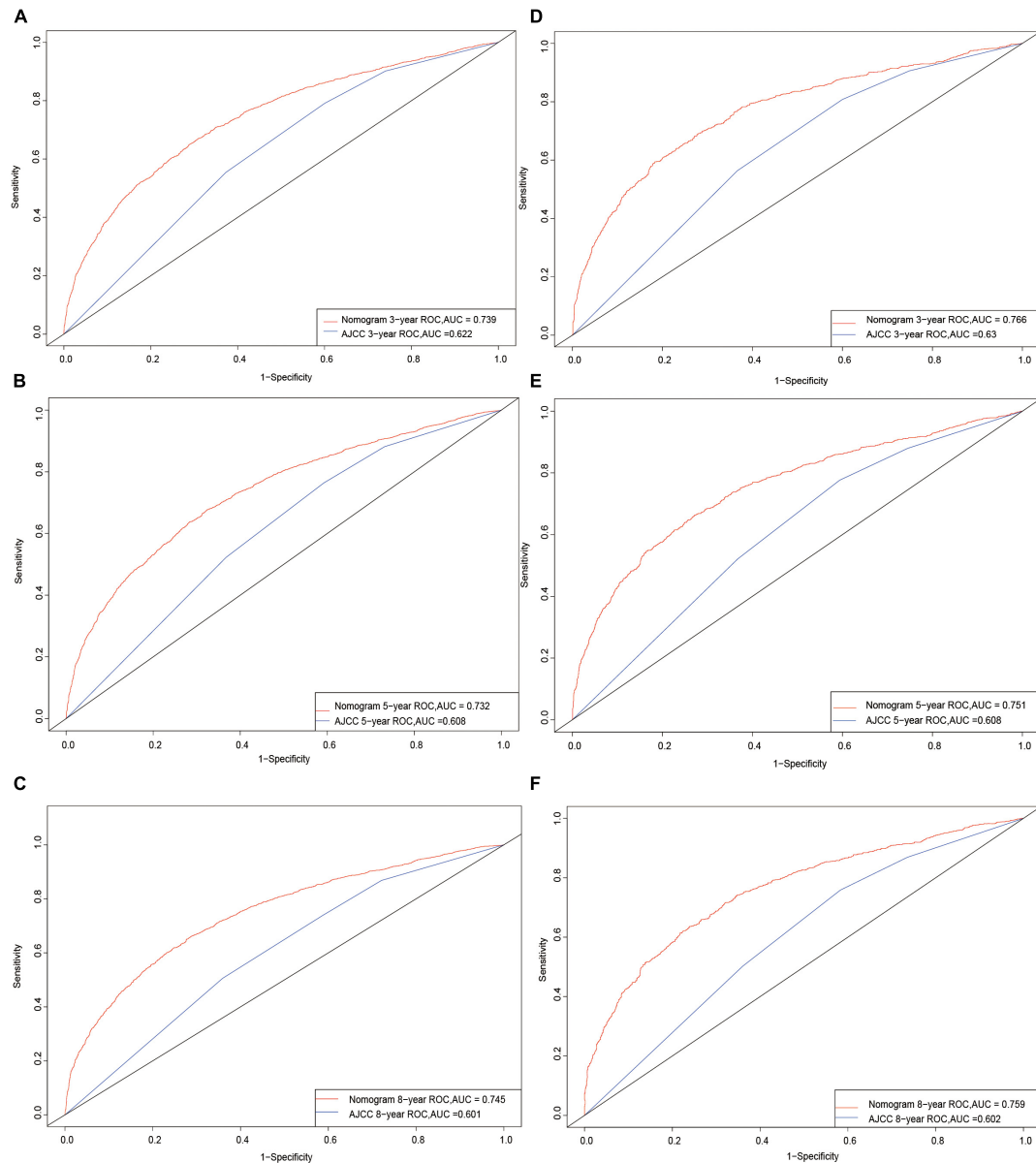


FIGURE 3 Receiver operating characteristic curves. The area under the ROC curve (AUC) for 3-, 5-, and 8-year CSS probability of the training cohort (A–C) and validation cohort (D–F).

the first time. Most of the published research studies have considered prognosis factors as an independent item (19, 21). Although most cSCCs can be successfully eradicated by surgical resection, cSCCs are somewhat characterized by high probabilities of recurrence, metastasis, and death. cSCC is most common in Caucasians and is more common in males than in females (with a 3:1 ratio) (22). Incidence increases with age, with an average age at diagnosis in the mid-60s (23), which differs from our results. All of these attributes indicate the need to develop a clinical prediction nomogram specific to NKLCSCC to assist doctors in making informed decisions. We were successful in creating a predictive nomogram that utilized the SEER database based on an integrated examination of demographic and clinicopathological factors. We then ascertained if the new model was superior to the AJCC staging system by comparing them.

The Cox regression results included in the nomogram indicate that besides summary stage, AJCC stage was the factor with the

strongest effect on CSS probability. This was mostly due to the AJCC staging system including information on regional lymph node and distant metastases, which are very important prognostic factors in NKLCSCC (24). Summary stage as a factor only represents the metastasis of the patient, which had the highest HR. Among demographic characteristics, age has always been an important prognostic factor for tumors, and the present results were no exception. Black race also presented a worse prognosis than white or other race, and being married had a better prognosis than being single or divorced, separated, or widowed (DSW), which was consistent with the findings of other studies (25, 26).

While the incidence of NKLCSCC was previously found to be higher in males than in females (27), the sex distribution in the present study was relatively even. A particularly interesting aspect of this study was that previous studies have not investigated the impact of marital status on NKLCSCC prognosis; this study found

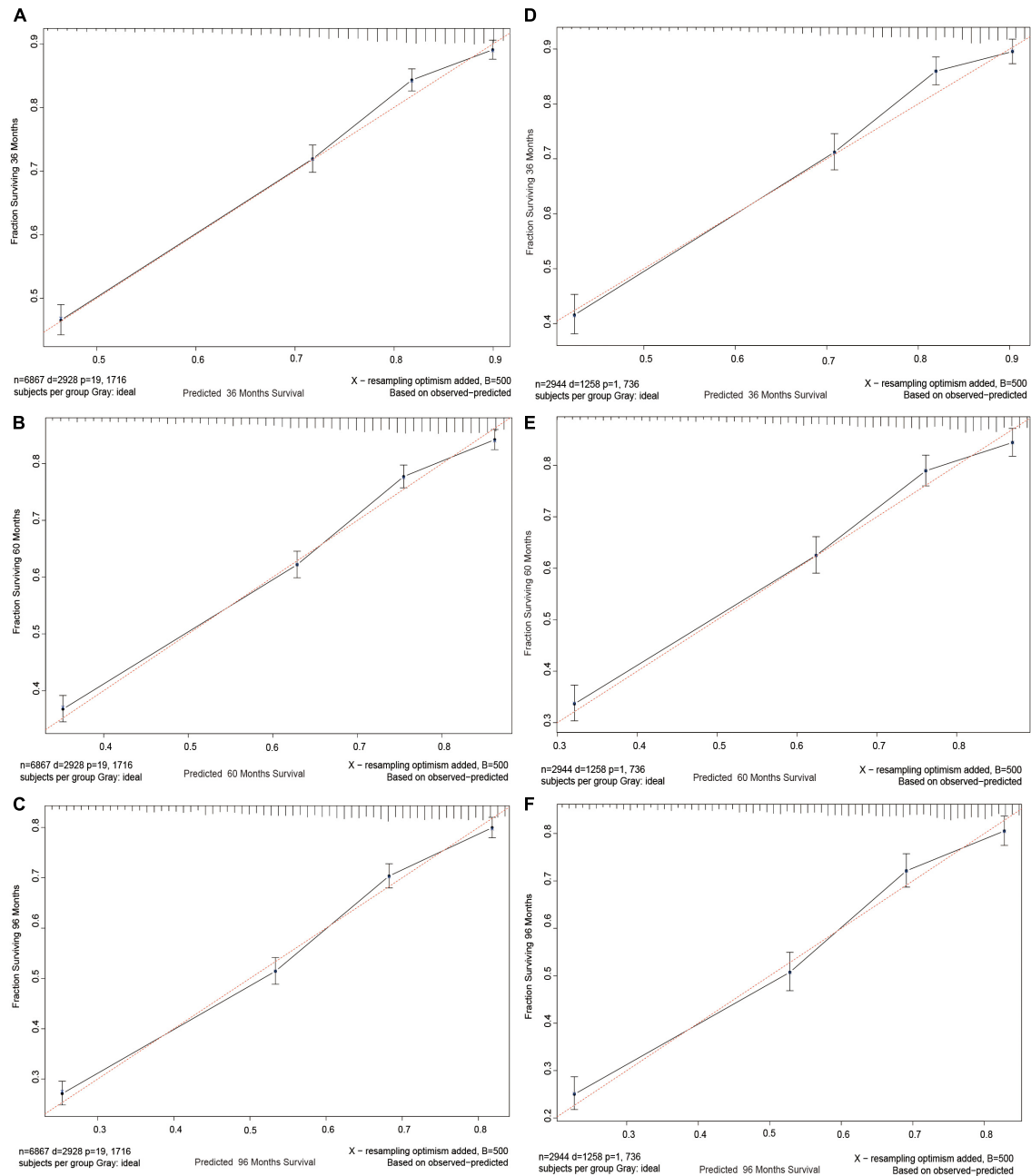


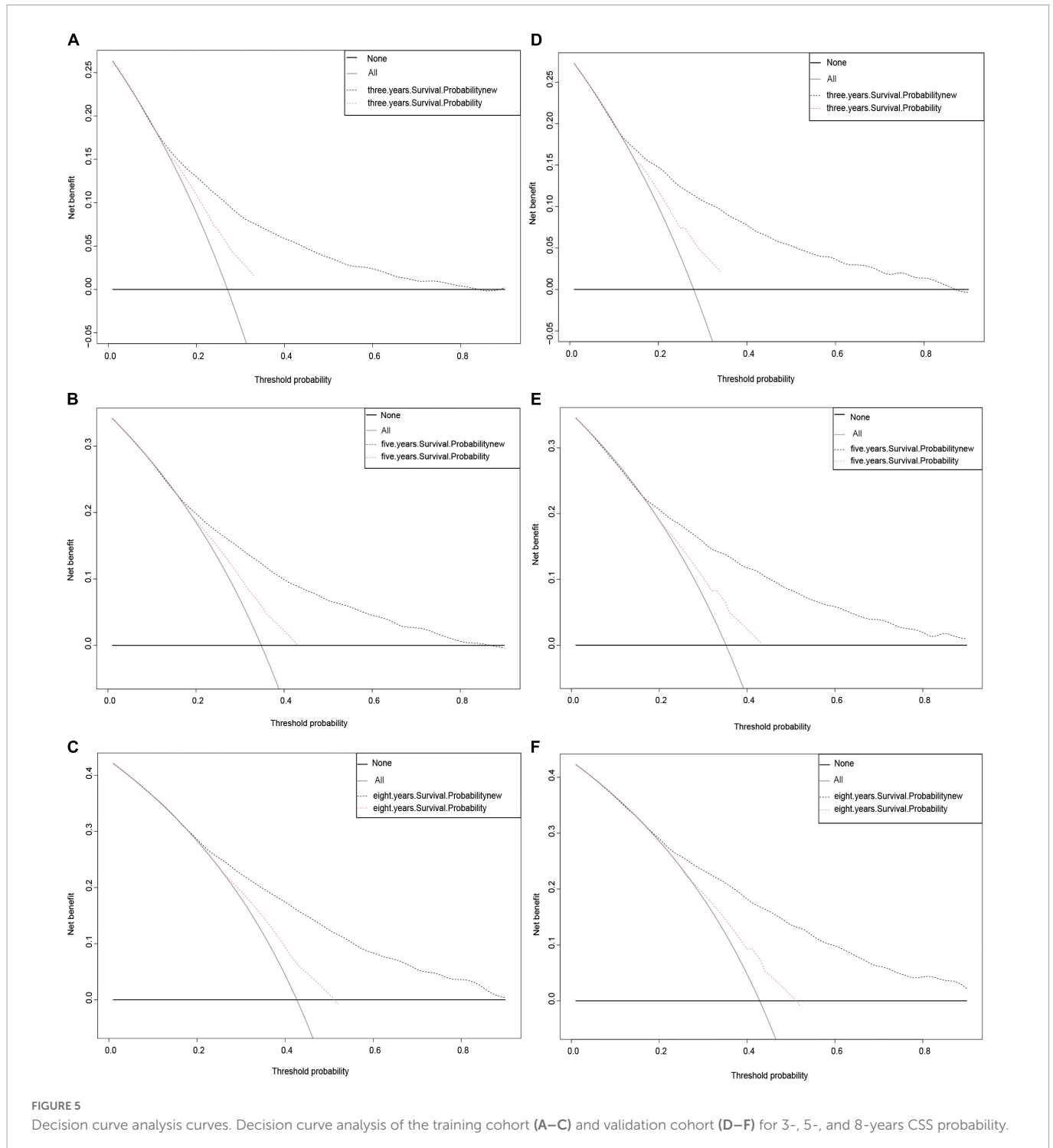
FIGURE 4

Calibration curves. Calibration curves for 3-, 5-, and 8-years CSS probability depict the calibration of each model in terms of the agreement between the predicted probabilities and observed outcomes of the training cohort (A–C) and validation cohort (D–F).

that being single or DSW were risk factors for the prognosis. Regarding clinicopathological features, AJCC staging is known to influence treatment selection, outcome, and prognosis. We found that summary stage significantly affected the NKLCSCC CSS probability in a similar way to RNP, and a larger RNP increased the effect on the prognosis. As can be seen from Figure 2, surgery, radiotherapy, and chemotherapy treatments were also significant prognostic factors.

Prior to actually applying the model, we performed a number of assessments that are essential for clinical prediction models after constructing the nomogram and accounting for the indicated prognostic factors. First, we verified the discrimination ability of

the model. The conventional ROC curve is a simple approach (28), and Figure 3 displays a nomogram with an AUC of >0.7 . This illustrates the good overall discrimination capability of the nomogram. Comparing the ROC of the new model with that of the AJCC model also revealed that the new model outperformed it. Second, the C-index is a more generalized measure of the discrimination ability of prediction models between different outcomes for survival data (29). The current findings further demonstrate the outstanding discrimination ability of our new model. NRI is frequently used to compare the accuracy of the predictive capability of two models, and in contrast to AUC and the C-index, it quantifies the number of subjects correctly



characterized by two models at a given set of cut-off points (30). In the training cohort, NRI values indicated increases in the proportion of correctly classifying CSS probabilities at 3, 5, and 8 years of 56.8, 58.1, and 62.4%, respectively, while those for the validation cohort were 70.8, 68.9, and 74.8% ($p = 0.001$). IDI is another metric that accounts for different cut-off points and can be used to reflect the overall improvement of a model, complementing NRI to a certain extent (31). The present IDI values demonstrated that, when compared with the AJCC model, the new model had better predictive power for 3-, 5-, and 8-year CSS probabilities by 12.5, 14.2, and 15.1% in the training cohort,

respectively, and by 15.2, 15.8, and 16.8% in the validation cohort ($p = 0.001$).

The aforementioned four indicators unmistakably demonstrate the strong discrimination ability of the nomogram and that the model is capable of accurately predicting the survival rate of patients with NKLCSCC. We also drew calibration plots to assess the calibration accuracy of the model. In Figure 4, the calibration curve of the model demonstrates a uniform distribution and is fairly near to the standard line. This suggested that the model-predicted incidence was very close to the actual incidence, indicating that the model has good conformity. The model demonstrated a good performance level

for estimating 3-, 5-, and 8-year CSS probabilities in patients with NKLCSCC when combined with the examination of discrimination and calibration abilities.

Finally, we assessed how effectively the model performed in healthcare situations. DCA is increasingly being used by researchers to evaluate the overall value of medical treatments for patients (17, 32). Our new approach had a better overall net benefit than the AJCC staging system and greater tolerance of survival probability, as seen in Figure 5. This indicates that the new model will provide patients with a greater overall benefit and assist physicians in making more informed treatment decisions.

Naturally, there were certain limitations to this study. First, we used the SEER database to collect analytic data in a retrospective manner, which could have induced some information bias. Second, several genetic markers, biomarkers, behavioral patterns, and other characteristics were not included in the study. Future cohort studies should more precisely pinpoint key prognostic markers and concentrate on including more prognostic factors and evaluating the model in external cohorts to provide more accurate findings.

Conclusion

Based on a reasonably large retrospective population, we have developed the first nomogram for predicting 3-, 5-, and 8-year CSS probabilities in patients with NKLCSCC. This nomogram incorporates demographic and clinicopathological characteristics, and thorough validation and assessment make this model helpful and simple for doctors to utilize when making clinical decisions for specific patients. It was also demonstrated to provide beneficial recommendations. In the future, we hope to construct more thorough nomograms based on a larger variety of data sources.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: SEER database.

References

- Waldman A, Schmults C. Cutaneous squamous cell carcinoma. *Hematol Oncol Clin North Am.* (2019) 33:1–12. doi: 10.1016/j.hoc.2018.08.001
- Staples M, Elwood M, Burton R, Williams J, Marks R, Giles G. Non-melanoma skin cancer in Australia: the 2002 national survey and trends since 1985. *Med J Austr.* (2006) 184:6–10. doi: 10.5694/j.1326-5377.2006.tb00086.x
- Nguyen K, Han J, Li T, Qureshi A. Invasive cutaneous squamous cell carcinoma incidence in US health care workers. *Arch Dermatol Res.* (2014) 306:555–60. doi: 10.1007/s00403-014-1469-3
- Brewster D, Bhatti L, Inglis J, Nairn E, Doherty V. Recent trends in incidence of nonmelanoma skin cancers in the East of Scotland, 1992–2003. *Br J Dermatol.* (2007) 156:1295–300. doi: 10.1111/j.1365-2133.2007.07892.x
- Andersson E, Paoli J, Wastensson G. Incidence of cutaneous squamous cell carcinoma in coastal and inland areas of Western Sweden. *Cancer Epidemiol.* (2011) 35:e69–74. doi: 10.1016/j.canep.2011.05.006
- Miller D, Weinstock M. Nonmelanoma skin cancer in the United States: incidence. *J Am Acad Dermatol.* (1994) 30:774–8. doi: 10.1016/S0190-9622(08)1509-5
- Stern R. Prevalence of a history of skin cancer in 2007: results of an incidence-based model. *Arch Dermatol.* (2010) 146:279–82. doi: 10.1001/archdermatol.2010.4
- Muzic J, Schmitt A, Wright A, Alniemi D, Zubair A, Olazagasti Lourido J, et al. Incidence and trends of basal cell carcinoma and cutaneous squamous cell carcinoma: a population-based study in Olmsted County, Minnesota, 2000 to 2010. *Mayo Clin Proc.* (2017) 92:890–8. doi: 10.1016/j.mayocp.2017.02.015
- Ildstad S, Tollerud D, Bigelow M, Remensnyder JP. A multivariate analysis of determinants of survival for patients with squamous cell carcinoma of the head and neck. *Ann Surg.* (1989) 209:237–41. doi: 10.1097/0000658-198902000-00016
- Balachandran V, Gonen M, Smith J, DeMatteo R. Nomograms in oncology: more than meets the eye. *Lancet Oncol.* (2015) 16:e173–80. doi: 10.1016/S1470-2045(14)71116-7
- Li C, Yang J, Zheng S, Xu F, Han D, Bai L, et al. Establishment and validation of a nomogram for tonsil squamous cell carcinoma: a retrospective study based on the SEER database. *Cancer Control.* (2020) 27:1073274820960481. doi: 10.1177/1073274820960481
- Xu F, Feng X, Zhao F, Huang Q, Han D, Li C, et al. Competing-risks nomograms for predicting cause-specific mortality in parotid-gland carcinoma: a population-based analysis. *Cancer Med.* (2021) 10:3756–69. doi: 10.1002/cam4.3919
- Xu F, Yang J, Han D, Huang Q, Li C, Zheng S, et al. Nomograms for estimating cause-specific death rates of patients with inflammatory breast cancer: a competing-risks analysis. *Technol Cancer Res Treat.* (2021) 20:15330338211016371. doi: 10.1177/15330338211016371

Ethics statement

Approximately 34.6% of the U.S. population is represented in the population-based cancer registries used by SEER to gather data on cancer incidence. Each spring, based on information provided by the registry in November of the previous year, SEER releases a standardized set of research data. The database utilized for this study did not contain any personal identities and only comprised publically available data, therefore institutional review board permission was not necessary. These were acquired *via* the SEER*Stat software after receiving further clearances.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Conflict of interest

KC was employed by the Biotech Company.

The remaining 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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

14. Yang J, Li Y, Liu Q, Li L, Feng A, Wang T, et al. Brief introduction of medical database and data mining technology in big data era. *J Evid Based Med.* (2020) 13:57–69. doi: 10.1111/jebm.12373
15. Wu W, Li Y, Feng A, Li L, Huang T, Xu A, et al. Data mining in clinical big data: the frequently used databases, steps, and methodological models. *Milit Med Res.* (2021) 8:44. doi: 10.1186/s40779-021-00338-z
16. Parikh C, Coca S, Thiessen-Philbrook H, Shlipak M, Koyner J, Wang Z, et al. Postoperative biomarkers predict acute kidney injury and poor outcomes after adult cardiac surgery. *J Am Soc Nephrol.* (2011) 22: 1748–57.
17. Steyerberg E, Vickers A, Cook N, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology.* (2010) 21:128–38.
18. Vickers A, Cronin A, Elkin E, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med Inform Decis Mak.* (2008) 8:53. doi: 10.1186/1472-6947-8-53
19. Kim J, Kozlow J, Mittal B, Moyer J, Olenecki T, Rodgers P. Guidelines of care for the management of cutaneous squamous cell carcinoma. *J Am Acad Dermatol.* (2018) 78:560–78.
20. Farberg A, Goldenberg G. New guidelines of care for the management of nonmelanoma skin cancer. *Cutis.* (2018) 101:319–21.
21. Newlands C, Currie R, Memon A, Whitaker S, Woolford T. Non-melanoma skin cancer: United Kingdom national multidisciplinary guidelines. *J Laryngol Otol.* (2016) 130:S125–32. doi: 10.1017/S0022215116000554
22. Que S, Zwald F, Schmultz C. Cutaneous squamous cell carcinoma: incidence, risk factors, diagnosis, and staging. *J Am Acad Dermatol.* (2018) 78: 237–47.
23. Xiang F, Lucas R, Hales S, Neale R. Incidence of nonmelanoma skin cancer in relation to ambient UV radiation in white populations, 1978–2012: empirical relationships. *JAMA Dermatol.* (2014) 150:1063–71. doi: 10.1001/jamadermatol.2014.762
24. Karia P, Morgan F, Califano J, Schmultz C. Comparison of tumor classifications for cutaneous squamous cell carcinoma of the head and neck in the 7th vs 8th edition of the AJCC cancer staging manual. *JAMA Dermatol.* (2018) 154:175–81. doi: 10.1001/jamadermatol.2017.3960
25. Yang J, Pan Z, Zhao F, Feng X, Liu Q, Li Y, et al. A nomogram for predicting survival in patients with nodular melanoma: a population-based study. *Medicine.* (2019) 98:e16059. doi: 10.1097/MD.00000000000016059
26. Yin T, Zhao Y, Yang Y, Xu H, Zheng D, Lyu J, et al. Nomogram for predicting overall survival in acral lentiginous melanoma: a population-based study. *Int J Gen Med.* (2021) 14:9841–51. doi: 10.2147/IJGM.S336443
27. Vazquez A, Khan M, Govindaraj S, Baredes S, Eloy J. Nasopharyngeal squamous cell carcinoma: a comparative analysis of keratinizing and nonkeratinizing subtypes. *Int Forum Allergy Rhinol.* (2014) 4:675–83.
28. Hanley J, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology.* (1983) 148:839–43.
29. Wang Y, Li J, Xia Y, Gong R, Wang K, Yan Z, et al. Prognostic nomogram for intrahepatic cholangiocarcinoma after partial hepatectomy. *J Clin Oncol.* (2013) 31: 1188–95.
30. Iki M, Fujita Y, Tamaki J, Kouda K, Yura A, Sato Y, et al. Trabecular bone score may improve FRAX® prediction accuracy for major osteoporotic fractures in elderly Japanese men: the fujiwara-kyo osteoporosis risk in men (FORMEN) cohort study. *Osteoporos Int.* (2015) 26:1841–8. doi: 10.1007/s00198-015-3092-3
31. Chambless L, Cumiskey C, Cui G. Several methods to assess improvement in risk prediction models: extension to survival analysis. *Stat Med.* (2011) 30:22–38.
32. Vickers A, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res.* (2019) 3:18. doi: 10.1186/s41512-019-0064-7