Quality Improvement Study

Medicine®

OPEN

Construction and validation of a nomogram for patients with skin cancer

Jizhen Ren, MD[a], Pengfei Sun, MD[b], Yanjin Wang, MD[a], Rui Cao, MD[c], Weina Zhang, MD[a,*]

Abstract

Skin cancer is a common malignant tumor in human beings. At present, the construction of clinical prediction models mainly focuses on malignant melanoma and no researchers have constructed clinical prediction models for all kind of skin cancer to predict the prognosis of skin cancer. We used patient data collected from the surveillance, epidemiology, and end results program database to construct and validate our model for clinical prediction of skin cancer, hoping to provide a reference for clinical treatment of skin cancer.

R software was used for univariate and multivariate Cox regression analysis of variables to screen out factors that have an impact on the survival of skin cancer patients. Then the prognostic model of skin cancer patients was constructed and the nomogram was drawn. Concordance Index (C-index), receiver operating characteristic (ROC) curve and calibration curve were used to evaluate the clinical prediction model.

A total of 3180 skin cancer patients were included in this study. We constructed nomogram, a 3-year and 5-year clinical prediction model for skin cancer patients. We used C-index to evaluate the accuracy of nomogram model, and the result of C-index was 0.728, 95%CI (0.703–0.753). The nomogram model was evaluated by ROC curve. The area under the curve values of the ROC curve for 3-year survival rate and 5-year survival rate were 0.732 and 0.768 respectively. The model calibration diagram of the modeling group also shows that the model exhibits high accuracy.

The nomogram model of postoperative survival of patients with skin cancer, based on the surveillance, epidemiology, and end results program database of patients with skin cancer, has shown good stability and accuracy in multi-method validation.

Abbreviations: AUC = area under the curve, BCC = Basal cell carcinoma, C-index = concordance index, MM = malignant melanoma, NMSC = non-melanoma skin cancer, ROC = receiver operating characteristic, SCC = squamous cell carcinoma, SEER = the surveillance, epidemiology, and end results program.

Keywords: clinical prediction, nomogram, surveillance, epidemiology, and end results program, skin cancer

1. Introduction

Skin cancer is a common malignant tumor in human beings.[1] In recent years, the incidence and mortality of skin cancer have been rising rapidly in all countries in the world, especially in the United States, where the incidence of skin cancer exceeds that of all other human malignant tumors combined.[2] Clinically, skin cancer is mainly divided into Malignant Melanoma (MM) and Non-melanoma Skin Cancer (NMSC). Among them, NMSC mainly includes Basal Cell Carcinoma (BCC) and Squamous Cell Carcinoma (SCC), which are the most common types of skin cancer endangering human health.[3]

The etiology of cutaneous squamous cell carcinoma is still unclear, and its occurrence may be related to the following factors:

1. Environmental factors such as daily exposure and ultraviolet radiation[4];
2. Human papillomavirus can be detected in skin cancers such as BCC and SCC[5];
3. Long-term stimulation of chemical carcinogens such as bitumen and tar derivatives.[6]

Apoptosis is the programmed death of cells, which is the result of inducing apoptosis mechanism in cells induced by internal and external factors. Studies on the pathogenesis of cutaneous squamous cell carcinoma have found that apoptosis inhibition plays an important role in the development of cutaneous squamous cell carcinoma.[7] In the treatment of malignant tumor, radiotherapy and biotherapy are mainly through inducing apoptosis to achieve therapeutic effect.[8] At present, the construction of clinical prediction models mainly focuses on MM and no researchers have constructed clinical prediction models for all kind of skin cancer to predict the prognosis of skin cancer.[9] So we used patient data collected from the Surveillance, Epidemiology, and End Results Program (SEER) database to

Editor: Sorush Niknamian.

RJZ and SPF contributed equally to this work.

The authors report no conflicts of interest.

The datasets generated during and/or analyzed during the current study are publicly available.

[a] Department of Plastic Surgery, Affiliated Hospital of Qingdao University, Qingdao, [b] Department of Auricular Reconstruction, [c] Research Center, Plastic Surgery Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China.

[*] Correspondence: Weina Zhang, Department of Plastic Surgery, Affiliated Hospital of Qingdao University, 266000, Qingdao, China (e-mail: hzhznxt@163.com).

How to cite this article: Ren J, Sun P, Wang Y, Cao R, Zhang W. Construction and validation of a nomogram for patients with skin cancer. Medicine 2021;100:4 (e24489).

Received: 14 October 2020 / Received in final form: 23 December 2020 / Accepted: 28 December 2020

http://dx.doi.org/10.1097/MD.0000000000024489

construct and validate our model for clinical prediction of skin cancer, hoping to provide a reference for clinical treatment of skin cancer. This article was written in accordance with the TRIPOD statement.[10]

2. Materials and methods

2.1. Data acquisition

Based on the SEER database, detailed clinical data were extracted from SEER database for all skin cancer patients from 1973 to 2017 (the latest data is updated to 2017). Because the data for our study came from the public database, our study did not require the approval of the medical ethics committee.

2.2. Inclusion and exclusion of patients

2.2.1. Inclusion criteria for patients.

1. Pathologically confirmed skin cancer.
2. The age, race, gender, survival status, survival time and other data of the patients were not missing.
3. There was no lack of data on TNM staging of skin cancer (staging was based on the TNM staging system of AJCC Cancer Staging Manual, 7th edition, 2010) and pathological type.

2.2.2. Exclusion criteria for patients.

1. Patients with cancer other than skin cancer;
2. Missing data on age, race, sex, survival status, survival time, etc.;
3. Missing data on staging, pathological type TNM skin cancer.

2.3. Grouping of included patients

We used R software to randomly divide the patient data included in the SEER database into modeling groups and validation groups. The sample size of the modeling group accounts for 70% of the total sample size, and the sample size of the verification group accounts for 30% of the total sample size.

2.4. Statistical analysis

We used R software for statistical analysis. We conducted univariate and multivariate Cox regression analysis on the variables to screen out the factors that had an impact on the survival of skin cancer patients. Then the prognostic model of skin cancer patients was constructed and the nomogram was drawn using the rms package. Concordance Index (C-index), Receiver Operating Characteristic (ROC) curve and calibration curve were used to evaluate the clinical prediction model. Finally, the data from the validation group were used to verify the clinical prediction model.

If the C-index is less than or equal to 0.5, the model is basically unreliable. If the C-index is between 0.51 and 0.7, the model has certain credibility. When the C-index is 0.7 to 0.89, the model has clinical significance. When the C-index is greater than or equal to 0.9, the model has extremely high reliability and extremely high predictive ability.

The Area Under the Curve (AUC) value of ROC curve is less than or equal to 0.5, the model is basically unreliable. If the AUC value is between 0.51 and 0.7, the model has low accuracy. When the AUC value is 0.71 to 0.9, the model has moderate accuracy. Models with an AUC value greater than 0.9 have high accuracy.

The calibration curve for evaluating clinical prediction model based on the correlation between predicted risks and actual results. If the fit is good, the model has reliability and prediction ability. On the contrary, the model has no reliability and prediction ability.

3. Results

3.1. Characteristics of the patients

A total of 3180 skin cancer patients were included in this study, and basic demographic characteristics are shown in Table 1. From the table we can see that the age distribution of patients is mainly over 60 years old, accounting for 75.2% of the total. The mean survival time was 27.9 ± 20.3 months. Caucasian patients accounted for 89.4%. There were 1278 female patients and 1902 male patients. From the TNM stage, early skin cancer patients accounted for the majority of the sample size. In histopathology, adnexal and skin appendage neoplasms account for 77% of the total sample, adenomas and adenocarcinomas account for 8.1%, ductal and lobular neoplasms account for 7.2%.

3.2. COX regression analysis of skin cancer patients

We conducted a COX regression analysis of the related factors in patients with skin cancer, and the univariate COX regression analysis showed (Table 2) that age, race, and NM stage of skin cancer were the influencing factors for the prognosis of patients with skin cancer. Multivariate COX regression analysis showed that age and NM stage of skin cancer affected the prognosis of patients with skin cancer. Although in univariate and multivariate COX regression analyses, skin cancer histological grade had little effect on prognosis ($P > .05$). However, the single-factor analysis for the prognosis of patients with skin cancer was statistically significant ($P < .05$). Therefore, we constructed a Nomogram model using age, race, NM stage, and histopathology.

3.3. Establishment of nomogram model

We used rms package to construct Nomogram (Fig. 1), a 3-year and 5-year clinical prediction model for patients with skin cancer after surgery, to provide a reference for clinical medical workers to predict postoperative survival of patients with skin cancer more accurately and efficiently.

3.4. Verification of nomogram model

We used C-index to evaluate the accuracy of Nomogram model. The evaluation results showed that the C-index of the model was 0.728, 95%CI (0.703–0.753). In the model calibration diagram of the modeling group, the blue line and the reference line (the dashed line in the diagram) basically coincide, indicating the model exhibits high accuracy (Figs. 2 and 3). Validation group data were used to verify the model calibration diagram of the modeling group, and high prediction consistency was found (Figs. 4 and 5). The Nomogram model was evaluated using the ROC curve. The AUC values of the ROC curve for 3-year survival rate and 5-year survival rate were 0.732 and 0.768 respectively, indicating moderate accuracy of the Nomogram model (Figs. 6 and 7).

Table 1

Demographics and clinicopathologic characteristics of patients with skin cancer.

Variable	0–49 years	50–59 years	60–69 years	70–79 years	≥80 years	Total
	341	449	735	778	877	3180
Survival Time (months)	32.6 ± 21.7	29.6 ± 20.9	28.7 ± 20.6	28.3 ± 20.3	24.3 ± 18.4	27.9 ± 20.3
Race						
Black	33 (9.7%)	43 (9.6%)	38 (5.2%)	30 (3.9%)	17 (1.9%)	161 (5.1%)
White	279 (81.8%)	378 (84.2%)	658 (89.5%)	710 (91.3%)	817 (93.2%)	2842 (89.4%)
Other	29 (8.5%)	28 (6.2%)	39 (5.3%)	38 (4.9%)	43 (4.9%)	177 (5.6%)
Sex						
Female	148 (43.4%)	176 (39.2%)	271 (36.9%)	303 (38.9%)	380 (43.3%)	1278 (40.2%)
Male	193 (56.6%)	273 (60.8%)	464 (63.1%)	475 (61.1%)	497 (56.7%)	1902 (59.8%)
Stage_T						
T0	1 (0.3%)	1 (0.2%)	0 (0.0%)	0 (0.0%)	1 (0.1%)	3 (0.09%)
T1	124 (36.4%)	147 (32.7%)	230 (31.3%)	218 (28.0%)	245 (27.9%)	964 (30.3%)
T2	66 (19.4%)	79 (17.6%)	132 (18.0%)	115 (14.8%)	135 (15.4%)	527 (16.6%)
T3	5 (1.5%)	0 (0.0%)	4 (0.5%)	3 (0.4%)	3 (0.3%)	15 (4.7%)
T4	3 (0.9%)	8 (1.8%)	7 (1.0%)	6 (0.8%)	7 (0.8%)	31 (1%)
TX	142 (41.6%)	214 (47.7%)	362 (49.3%)	436 (56.0%)	486 (55.4%)	1640 (51.6%)
Stage_N						
N0	291 (85.3%)	373 (83.1%)	597 (81.2%)	611 (78.5%)	658 (75.0%)	2530 (79.6%)
N1	11 (3.2%)	9 (2.0%)	11 (1.5%)	9 (1.2%)	8 (0.9%)	48 (1.5%)
N2a	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.1%)	2 (0.2%)	3 (0.1%)
N2b	3 (0.9%)	4 (0.9%)	12 (1.6%)	9 (1.2%)	10 (1.1%)	38 (1.2%)
N2c	0 (0.0%)	1 (0.2%)	4 (0.5%)	0 (0.0%)	0 (0.0%)	5 (0.2%)
N3	2 (0.6%)	1 (0.2%)	1 (0.1%)	0 (0.0%)	0 (0.0%)	4 (0.1%)
NX	34 (10.0%)	61 (13.6%)	110 (15.0%)	148 (19.0%)	199 (22.7%)	552 (17.4%)
Stage_M						
M0	338 (99.1%)	438 (97.6%)	725 (98.6%)	772 (99.2%)	868 (99.0%)	3141 (98.8%)
M1	3 (0.9%)	11 (2.4%)	10 (1.4%)	6 (0.8%)	9 (1.0%)	39 (1.2%)
Status						
Alive	327 (95.9%)	416 (92.7%)	660 (89.8%)	642 (82.5%)	563 (64.2%)	2608 (82%)
Dead	14 (4.1%)	33 (7.3%)	75 (10.2%)	136 (17.5%)	314 (35.8%)	572 (18%)
Histology						
Transitional cell papillomas and carcinomas	0 (0.0%)	3 (0.7%)	5 (0.7%)	1 (0.1%)	5 (0.6%)	14 (0.4%)
Adenomas and adenocarcinomas	35 (10.3%)	43 (9.6%)	51 (6.9%)	60 (7.7%)	67 (7.6%)	256 (8.1%)
Adnexal and skin appendage neoplasms	278 (81.5%)	362 (80.6%)	574 (78.1%)	586 (75.3%)	649 (74.0%)	2449 (77%)
Mucoepidermoid neoplasms	4 (1.2%)	0 (0.0%)	1 (0.1%)	1 (0.1%)	3 (0.3%)	9 (0.3%)
Cystic, mucinous and serous neoplasms	13 (3.8%)	11 (2.4%)	36 (4.9%)	23 (3.0%)	22 (2.5%)	105 (3.3%)
Ductal and lobular neoplasms	2 (0.6%)	14 (3.1%)	47 (6.4%)	82 (10.5%)	84 (9.6%)	229 (7.2%)
Acinar cell neoplasms	1 (0.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.1%)	2 (0.1%)
Complex epithelial neoplasms	2 (0.6%)	9 (2.0%)	11 (1.5%)	14 (1.8%)	24 (2.7%)	60 (1.9%)
Complex mixed and stromal neoplasms	6 (1.8%)	7 (1.6%)	10 (1.4%)	11 (1.4%)	22 (2.5%)	56 (1.8%)

Table 2

COX regression analysis of patients with skin cancer.

Variable	Univariate Analysis HR	Univariate Analysis 95% CI	Univariate Analysis *P*-value	Multivariate Analysis HR	Multivariate Analysis 95% CI	Multivariate Analysis *P* value
Age						
<50	1 (reference)					
50–59	11.50	(6.73,19.65)	< .05	11.22	(6.55,19.24)	< .05
60–69	1.95	(1.05,3.65)	< .05	1.69	(0.90,3.17)	.10
70–79	2.80	(1.58,4.96)	< .05	2.64	(1.49,4.67)	< .05
>=80	4.87	(2.81,8.44)	< .05	4.83	(2.78,8.40)	< .05
Sex						
Female	1 (reference)					
Male	1.10	(0.93,1.31)	.26	1.13	(0.96,1.35)	.15
Race						
Black	1 (reference)					
White	1.31	(0.69,2.48)	.41	1.13	(0.59,2.16)	.72
Other	1.87	(1.13,3.07)	< .05	1.40	(0.84,2.35)	.20
Stage_T						
T0	1 (reference)					
T1	3.724e+05	–	.99	2.004e+06	–	.99
T2	6.252e+05	–	.99	3.437e+06	–	.99
T3	5.026e+05	–	.99	2.057e+06	–	.99
T4	1.014e+06	–	.99	4.214e+06	–	.99
TX	4.453e+05	–	.99	1.954e+06	–	.99
Stage_N						

(*continued*)

Table 2

(continued).

Variable	Univariate Analysis HR	Univariate Analysis 95% CI	Univariate Analysis *P*-value	Multivariate Analysis HR	Multivariate Analysis 95% CI	Multivariate Analysis *P* value
N0	1 (reference)					
N1	1.49	(0.84,2.64)	.17	1.20	(0.64,2.24)	.56
N2a	13.47	(3.35,54.27)	< .05	7.80	(1.92,31.62)	< .05
N2b	2.86	(1.68,4.87)	< .05	2.00	(1.16,3.47)	< .05
N2c	13.11	(4.19,41.01)	< .05	3.17	(0.90,11.23)	.07
N3	2.098e–06	–	.99	2.282e–06	–	.99
NX	1.67	(1.38,2.04)	< .05	1.53	(1.23,1.89)	< .05
Stage_M						
M0	1 (reference)					
M1	5.94	(3.87,9.11)	< .05	5.56	(3.36,9.18)	< .05
Histology						
Transitional cell papillomas and carcinomas	1 (reference)					
Adenomas and adenocarcinomas	1.04	(0.33,3.32)	.94	1.21	(0.37,3.89)	.75
Adnexal and skin appendage neoplasms	0.74	(0.24,2.32)	.61	0.90	(0.29,2.84)	.86
Mucoepidermoid neoplasms	0.98	(0.16,5.87)	.98	1.03	(0.17,6.32)	.97
Cystic, mucinous and serous neoplasms	0.24	(0.06,0.94)	< .05	0.34	(0.09,1.33)	.12
Ductal and lobular neoplasms	0.80	(0.25,2.58)	.71	0.70	(0.21,2.28)	.55
Acinar cell neoplasms	1.694e–06	–	.99	3.711e–07	–	.99
Complex epithelial neoplasms	1.31	(0.39,4.42)	.67	1.16	(0.34,3.98)	.81
Complex mixed and stromal neoplasms	1.36	(0.39,4.69)	.63	1.49	(0.43,5.23)	.53

CI = confidence interval; HR = hazard ratio.

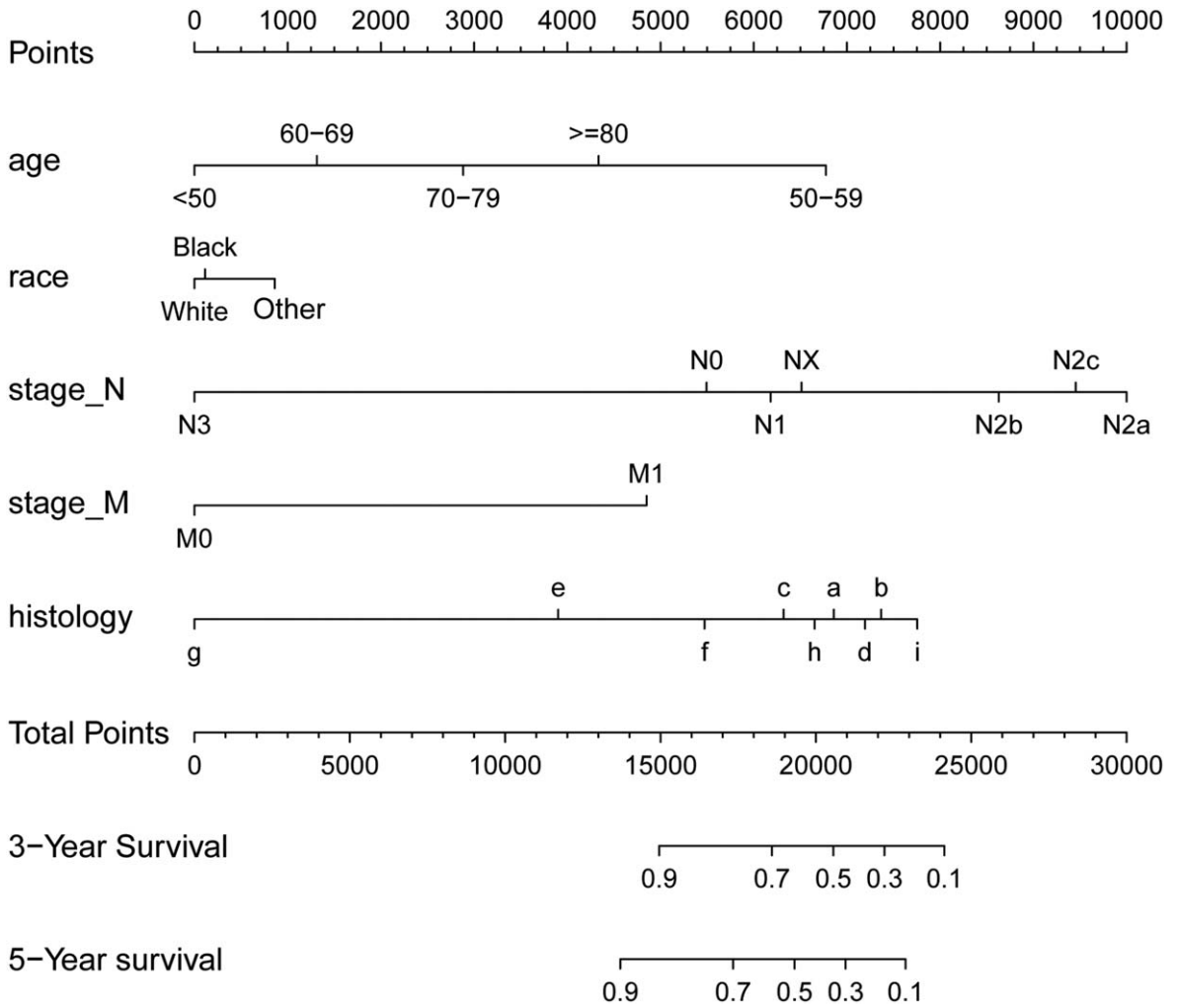

Figure 1. Skin cancer survival nomogram. (To use the nomogram, an individual patient's value is located on each variable axis, and a line is drawn upward to determine the number of points received for each variable value. The sum of these numbers is located on the Total Points axis, and a line is drawn downward to the survival axes to determine the likelihood of 3- or 5-year survival). A: Transitional cell papillomas and carcinomas; B: Adenomas and adenocarcinomas; C: Adnexal and skin appendage neoplasms; D: Mucoepidermoid neoplasms; E: Cystic, mucinous and serous neoplasms; F: Ductal and lobular neoplasms; G: Acinar cell neoplasms; H: Complex epithelial neoplasms; I: Complex mixed and stromal neoplasms.

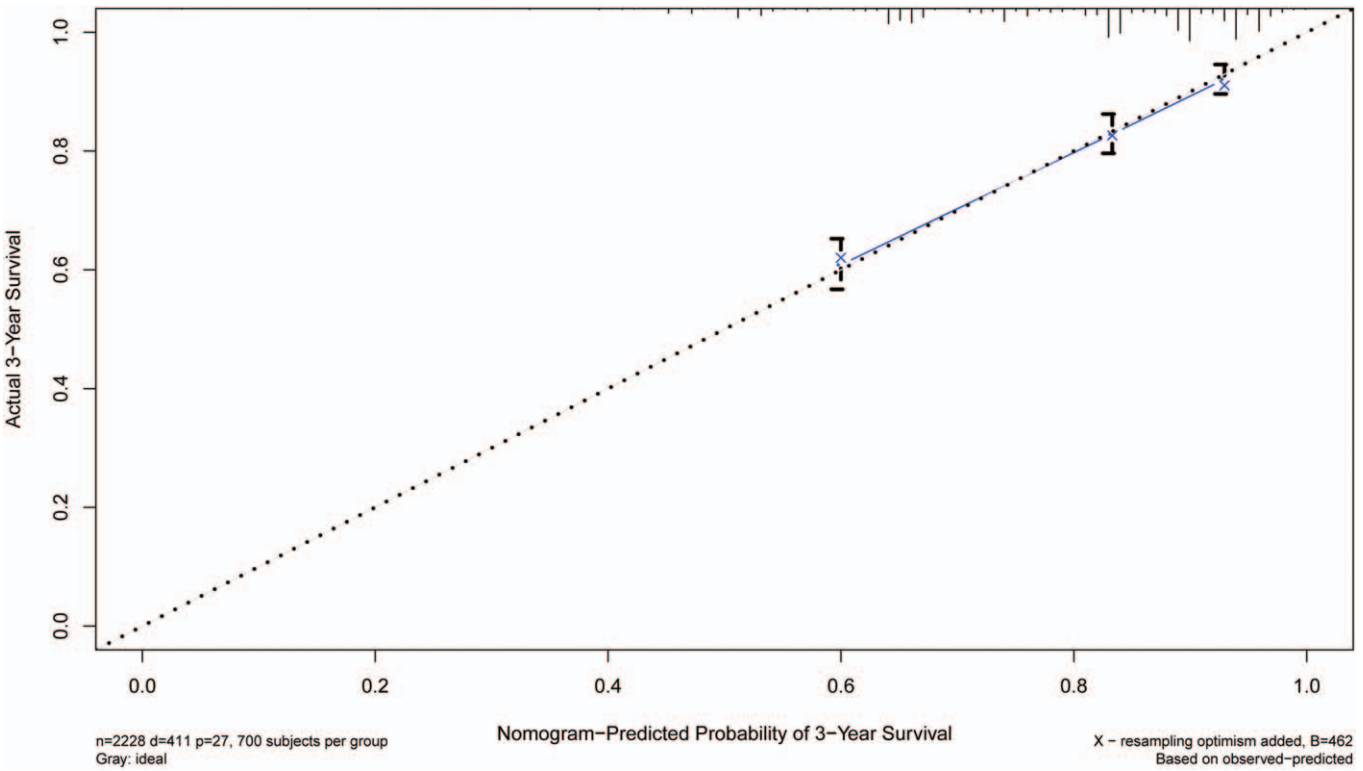

Figure 2. Nomogram-Predicted Probability of 3-Year Survival (the modeling group). In the model calibration diagram, the abscissa of the diagram is the prediction probability, indicating the probability of the occurrence of the predicted events. The vertical coordinate of the diagram is the actual probability, which represents the actual probability of the events. The blue line is the fitting line, representing the actual value corresponding to the predicted value. If the prediction model has high accuracy, the blue line and the reference line (the dashed line in the diagram) basically coincide. On the contrary, the blue line and the reference line do not coincide.

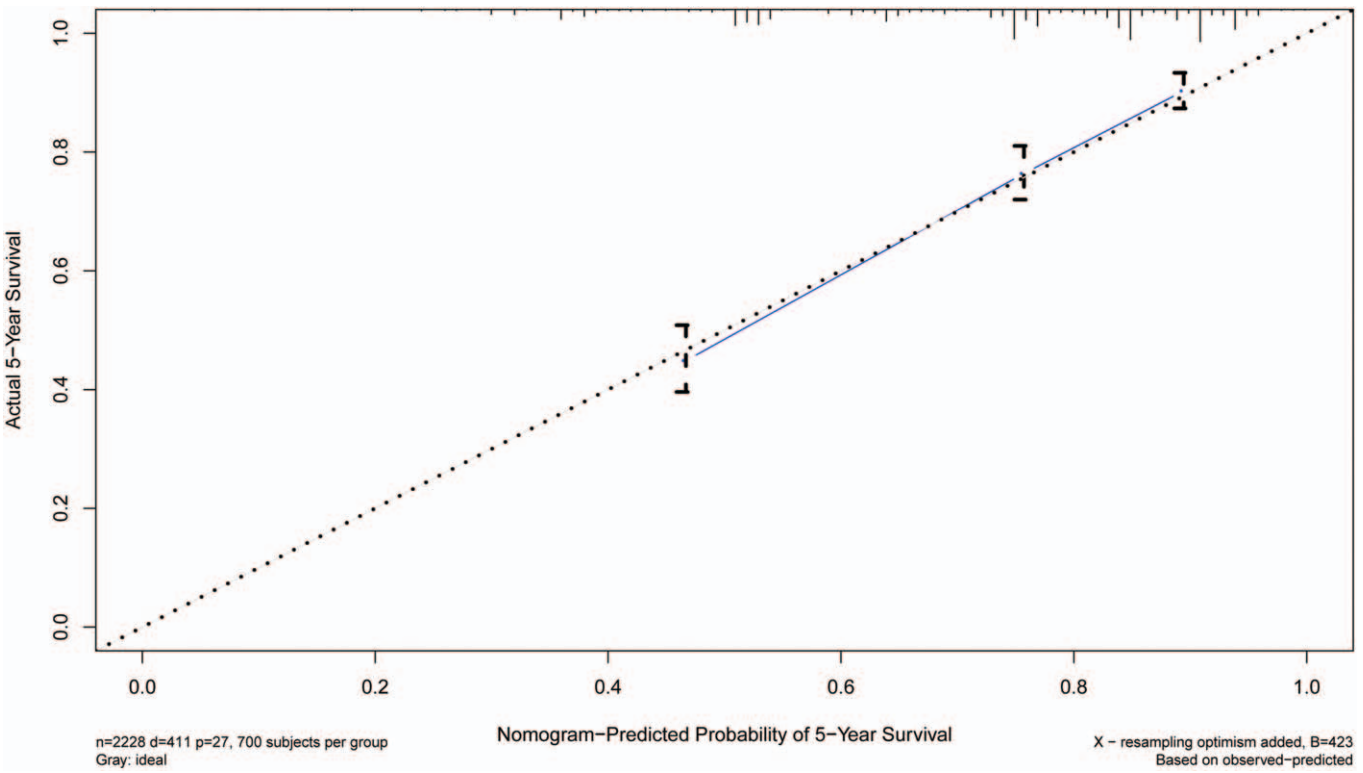

Figure 3. Nomogram-Predicted Probability of 5-Year Survival (the modeling group).

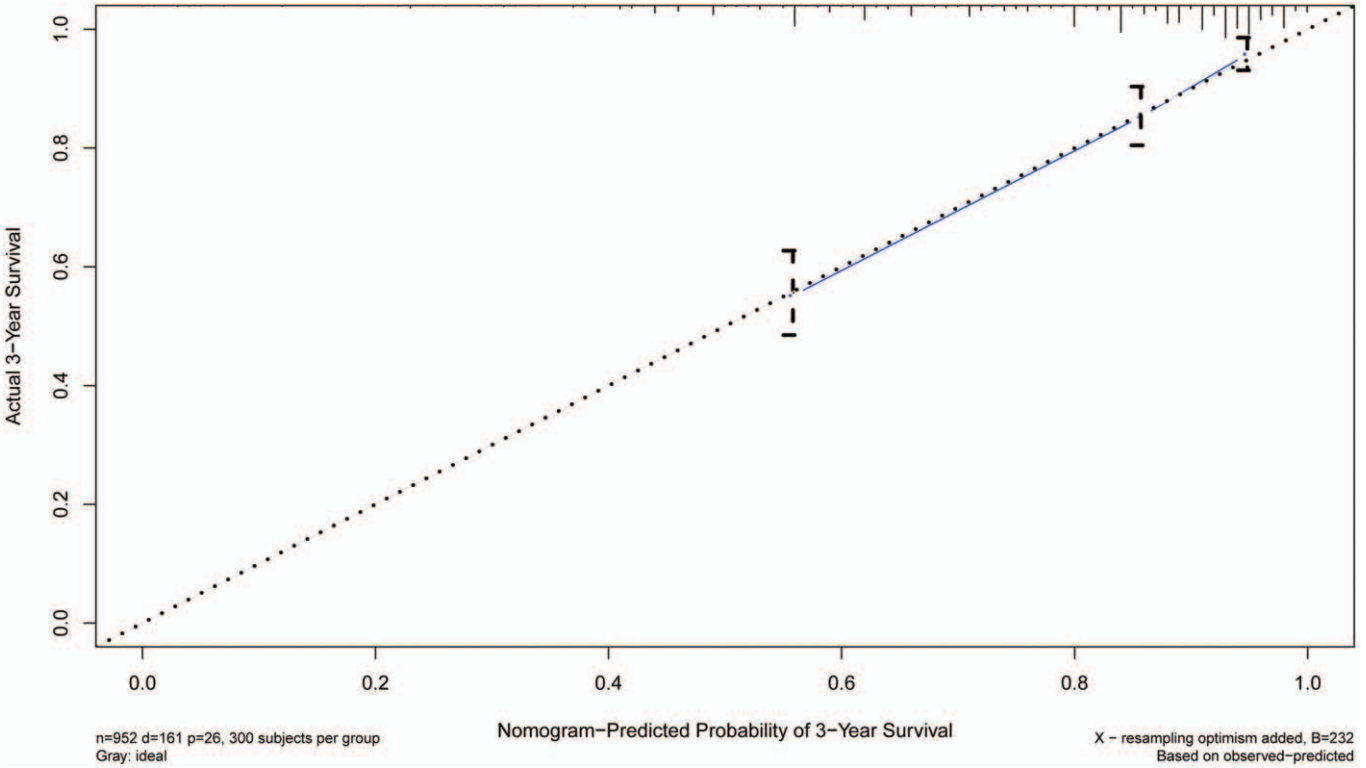

Figure 4. Nomogram-Predicted Probability of 3-Year Survival (the validation group).

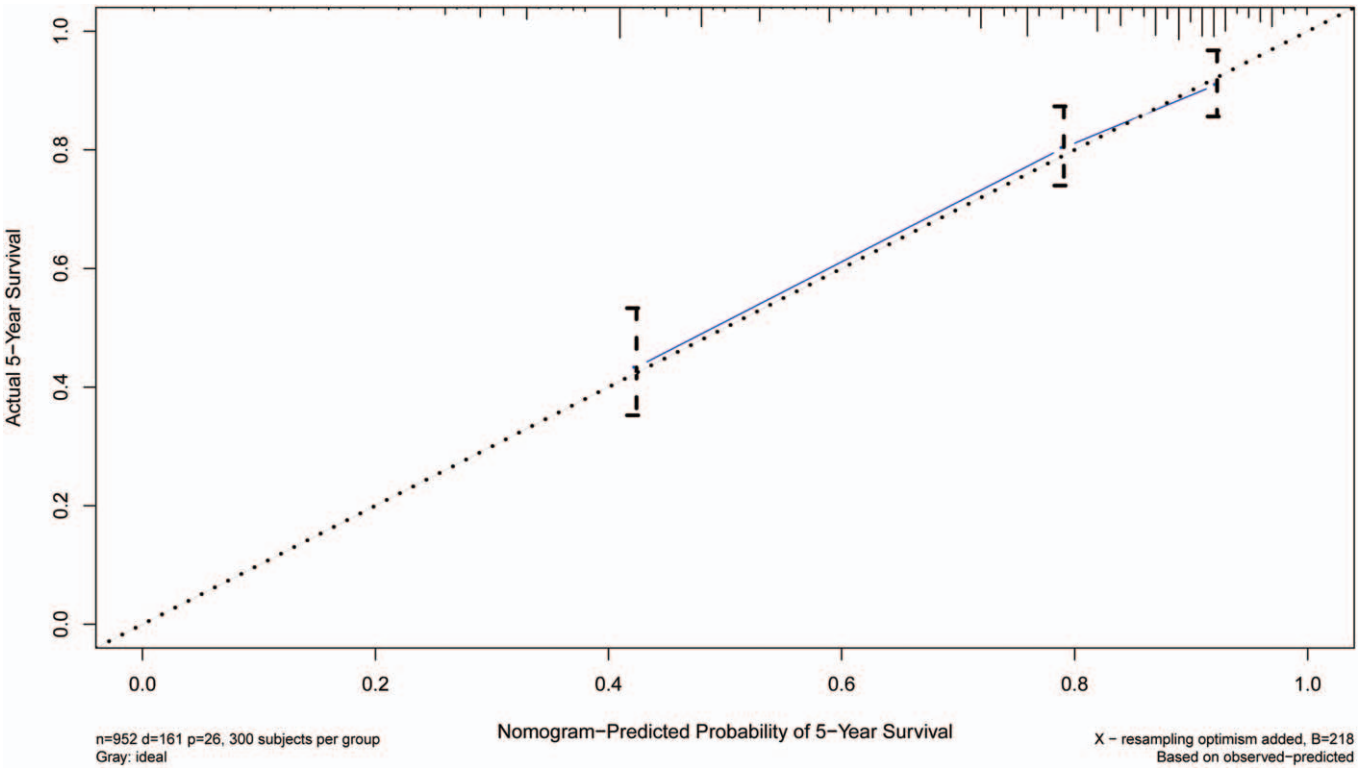

Figure 5. Nomogram-Predicted Probability of 5-Year Survival (the validation group).

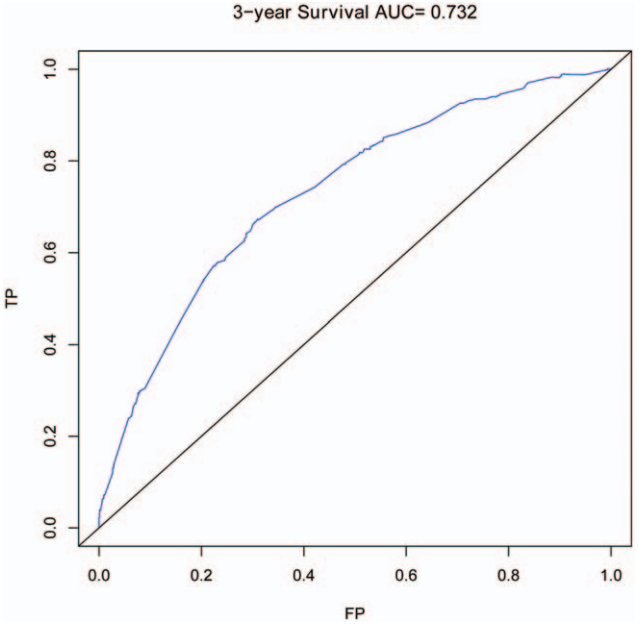

Figure 6. The ROC curve for 3-year survival rate.

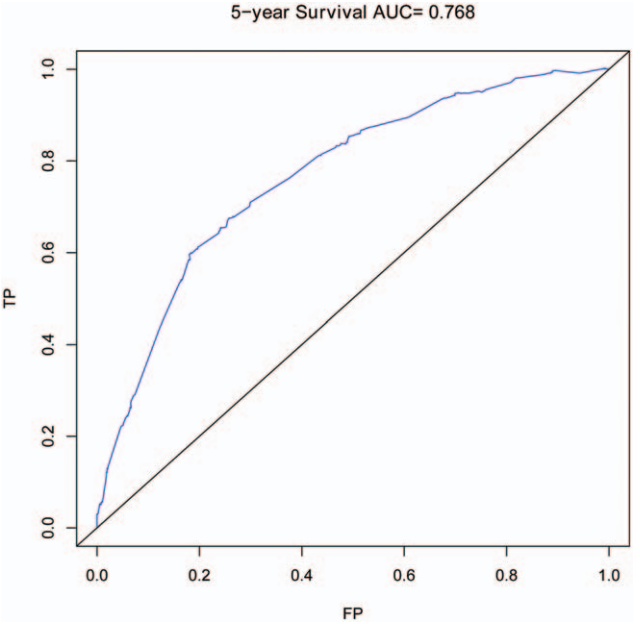

Figure 7. The ROC curve for 5-year survival rate.

4. Discussion

Although the fatality rate of skin cancer is much lower than that of other malignancies, the morbidity and mortality rates have also been increasing in recent years, which are of great harm to human health.[11] Therefore, it is of great significance to construct the prognostic prediction model of skin cancer patients. At present, there is no accurate prediction model for the prognosis of skin cancer patients, and little is known about the prognostic factors of skin cancer.[12] This study collected data of patients with skin cancer from SEER database, analyzed prognostic factors of patients with skin cancer after surgery from a large sample and multi-center perspective, and established a more accurate prediction model to predict the prognosis of patients, so as to provide certain reference and guidance for clinicians in clinical work.

In this study, we used SEER data on skin cancer patients to construct a skin cancer clinical prediction model for 3180 skin cancer patients, a large sample size and a comprehensive population, which is widely used to predict the prognosis of skin cancer patients. The Nomogram model was verified by c-Index, ROC curve and calibration diagram, and the clinical prediction model for skin cancer patients constructed in this study was of moderate accuracy.

This study also has limitations:

1. The data used in this study were obtained from SEER database, and the SEER database did not provide specific treatment plans such as surgery, chemotherapy, and follow-up treatment for skin cancer patients, nor did it record patients' disease history, complications, tumor recurrence, systemic diseases, etc. These missing values may affect the prognosis prediction of patients.
2. Due to the limited histopathological classification of skin cancer in SEER database, the Nomogram model we constructed did not incorporate traditional MM, BCC, SCC.
3. The Nomogram model was validated using SEER database patient data, and although the results suggest moderate accuracy, further validation of the Nomogram model with clinical data is needed.

5. Conclusions

The nomogram model of postoperative survival of patients with skin cancer, based on SEER database of patients with skin cancer, has shown good stability and accuracy in multi-method validation, suggesting that this model may provide clinicians with potential value in assessing postoperative survival of patients with skin cancer.

Author contributions

Conceptualization: Jizhen Ren, Weina Zhang.
Data curation: Jizhen Ren.
Formal analysis: Jizhen Ren, Pengfei Sun.
Investigation: Yanjin Wang.
Methodology: Yanjin Wang, Rui Cao, Weina Zhang.
Software: Pengfei Sun, Rui Cao.
Supervision: Weina Zhang.
Validation: Pengfei Sun, Weina Zhang.
Writing – original draft: Jizhen Ren, Pengfei Sun.
Writing – review & editing: Yanjin Wang, Weina Zhang.

References

[1] Ternov NK, Vestergaard T, Hölmich LR, et al. Reliable test of clinicians' mastery in skin cancer diagnostics. Arch Dermatol Res 2020; doi: 10.1007/s00403-020-02097-8.
[2] Tschandl P, Rinner C, Apalla Z, et al. Human-computer collaboration for skin cancer recognition. Nat Med 2020;26:1229–34.
[3] Ferrari B, Reggiani C, Francomano M, et al. Clinical factors influencing the outcomes of an acellular dermal matrix for skin cancer treatment: a retrospective study. Adv Skin Wound Care 2020;33:367–74.
[4] Vaverková E, Neradová Richterová M, Adamcová D, et al. Environmental changes and their impact on human behaviour - Case study of the incidence of skin cancer. Sci Total Environ 2020;738:139788.

[5] Hochheiser K, Gyorki DE, Gebhardt T. Skin colonization with beta papilloma virus drives tissue immunity and resistance to squamous cell cancer. Immunol Cell Biol 2020;98:9–11.
[6] Chen L, Hayden MS, Gilmore ES, et al. PKK deletion in basal keratinocytes promotes tumorigenesis after chemical carcinogenesis. Carcinogenesis 2018;39:418–28.
[7] Parekh V, Seykora JT. Cutaneous squamous cell carcinoma. Clin Lab Med 2017;37:503–25.
[8] Abdel Fadeel DA, Kamel R, Fadel M. PEGylated lipid nanocarrier for enhancing photodynamic therapy of skin carcinoma using curcumin: in-vitro/in-vivo studies and histopathological examination. Sci Rep 2020;10:10435.
[9] Bertolli E, Franke V, Calsavara VF, et al. Validation of a nomogram for non-sentinel node positivity in melanoma patients, and its clinical implications: a Brazilian-Dutch study. Ann Surg Oncol 2019;26:395–405.
[10] Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BJOG 2015;122:434–43.
[11] Ji AL, Rubin AJ, Thrane K, et al. Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. Cell 2020;182:497–514. e22.
[12] Bui AN, LeBoeuf NR, Nambudiri VE. Skin cancer risk in CHEK2 mutation carriers. J Eur Acad Dermatol Venereol 2020; doi: 10.1111/jdv.16729.