European Review for Medical and Pharmacological Sciences 2020; 24: 10483-10495

Development and validation of an individualized nomogram for predicting pancreatic adenocarcinoma-specific survival: a SEER population analysis of 5,805 patients

J.-H. LIU[1,2], S.-F. WANG[3], B.-R. LI[4], X. ZHU[1,2,5]

[1]Guangdong Key Laboratory for Research and Development of Natural Drugs, The Marine Biomedical Research Institute, Guangdong Medical University, Zhanjiang, China
[2]The Marine Biomedical Research Institute of Guangdong Zhanjiang, Zhanjiang, China
[3]Department of Digestive System Oncology, The Fifth People's Hospital of Ganzhou, Ganzhou, China
[4]Cancer Center, Medical College of Georgia in Augusta University, GA, USA
[5]Southern Marine Science and Engineering Guangdong Laboratory (Zhanjiang), Zhanjiang, China

J.-H. Liu and S.-F. Wang contributed equally to this work

Abstract. – OBJECTIVE: A model to predict the overall survival (OS) of pancreatic adenocarcinoma (PAC) is required in consideration of its inferior prognosis.

PATIENTS AND METHODS: The patients diagnosed with pancreatic cancer between 1975 and 2016 in the Surveillance, Epidemiology, and End Results (SEER) database was used as raw data. A training cohort and a verification cohort were used for internal validation and external validation, respectively. The nomogram model was constructed to predict the OS.

RESULTS: A total of 5,805 patients with PAC from 2010-2015 were analyzed. Most patients were over 65 years old (61.8%), white (81.2%), in stage IIA, IIB (49.0%), and IV (32.4%), less than 50 mm in diameter (80.2%). PAC patients with wide involvement range, no metastasis, and infiltration range more than 300 accounted for 58.2%, 67.6%, 78.2%, respectively. The vast majority of the PAC patients (90.9%) did not receive primary site surgery. Most of the PAC patients (68.1%) received chemotherapy and only 25.8% of PAC patients received radiotherapy. The overall mean survival time, overall median survival time and overall survival rate were 15.1 months, 10.0 months, and 16.7%, respectively.

CONCLUSIONS: Our nomogram that based on age, chemotherapy, grade, Radiation sequence with surgery, Radiation recode, RX Summ-Surg Prim at Site (surgery that removes and/or destroys primary tumor performed as part of the first course of therapy), size, and stage was of well prediction ability.

Key Words:
Nomogram, Pancreatic adenocarcinoma, Prognostic factors, Overall survival, SEER.

Introduction

Pancreatic cancer was one of the deadliest malignancies worldwide. The lack of sensitive and specific early diagnostic indicators caused the late diagnosis of pancreatic cancer. Besides, many patients were refused to perform primary site resection for the arterial blood vessels surrounding the lesion had been violated, although it was the best treatment. What's more, pancreatic cancer was resistant to adjuvant therapy. The above factors caused the 5-year survival rate of pancreatic cancer patients to be less than 5%[1,2]. The mortality rate of pancreatic cancer had been rising for decades[3,4]. Pancreatic cancer ranked as the fourth leading cause of cancer death in the United States[5], and it will be the second leading cause of cancer-related death in decades without improving treatment[6]. PAC was the primary subtype, accounting for about 90% of cases[7]. Consequently, it was necessary to establish a model to predict PAC patients' prognosis to guide the work of clinicians.

Currently, a nomogram proved to be a useful tool[8]. Nomogram based on the independent

Corresponding Authors: Xiao Zhu, MD; e-mail: xzhu@gdmu.edu.cn
Biaoru Li, MD; e-mail: bli@augusta.edu

prognostic factors was created to predict 1-year, 3-year, and 5-year survival for PAC patients. Its predictive value was verified by the concordance index and calibration curve.

The purpose of the present study was to establish a nomogram based on SEER data to better evaluate the prognosis of PAC and understand the independent prognostic factors that influenced prognosis.

Patients and Methods

Data

We used the SEER*Stat version 8.3.5 (https://SEER.Cancer.Gov/) to download patient data from the SEER database. The inclusion criterion was as followed: all patients diagnosed with pancreatic cancer in the SEER database. In total, we downloaded data of 243,418 patients diagnosed with pancreatic cancer between 1975-2016, and the input information included: Age recode with <1 year old, Race recode (White, Black, Other), Sex, Year of diagnosis, Site recode ICD-O-3/WHO 2008, Behavior recode for analysis, Primary Site, Histologic Type ICD-O-3, Grade, Laterality, Derived AJCC Stage Group, seventh ed (2010-2015), Derived AJCC T, seventh ed (2010-2015), Derived AJCC N, seventh ed (2010-2015), Derived AJCC M, seventh ed (2010-2015), RX Summ--Surg Prim Site (1998+), RX Summ--Scope Reg LN Sur (2003+), RX Summ--Surg Oth Reg/Dis (2003+), Radiation sequence with surgery, Reason no cancer-directed surgery, Radiation recode, Chemotherapy recode (yes, no/unk), Tumor Size Summary (2016+), Regional nodes examined (1988+), Regional nodes positive (1988+), CS tumor size (2004-2015), CS extension (2004-2015) (extension of tumor away from the primary site), COD to site recode, SEER cause-specific death classification, SEER other cause of death classification, Survival months, COD to site rec KM, Vital status recode (study cutoff used), Type of follow-up expected, Sequence number, First malignant primary indicator, Record number recode, Total number of in situ/malignant tumors for patient, Total number of benign/borderline tumors for patient, Race recode (W, B, AI, API), Origin recode NHIA (Hispanic, Non-Hisp), Age at diagnosis, Patient ID, Type of Reporting Source, Insurance Recode (2007+), Marital status at diagnosis. However, items with less clinical significance (Type of follow-up expected, Origin recode NHIA (Hispanic, Non-Hisp), Type of Reporting Source, Reason no cancer-directed surgery, Patient ID, Primary Site, Laterality, COD to site recode, COD TO SITE REC KM) and items contained too much incomplete information (RX Summ--Scope Reg LN Sur (2003+), Tumor Size Summary) were deleted to ensure the comprehensiveness of modeling variables and effectively eliminate the statistical interference of too many miscellaneous items.

In order to get better research results, we screened the original data according to the following exclusion criteria. First, we excluded patients who were not the first tumor (n=23,914), as well as those with blank (RX Summ--Surg Prim Site (1998+): n=58,750, RX Summ--Surg Oth Reg/Dis (2003+): n=31,124, CS extension (2004-2015): n=20,263, Insurance Recode (2007+): n=21,646, Derived AJCC Stage Group, seventh ed (2010-2015): n=6,464), unknown (Race recode (White, Black, Other): n=455, RX Summ--Surg Oth Reg/Dis (2003+): n=248, Radiation sequence with surgery: n=32, Radiation recode: n=997, SEER cause-specific death classification: n=879, Insurance Recode (2007+): n=3,747, Marital status at diagnosis: n=2,508, Survival months: n=56, Grade: n=40,735), and defect documented (Site recode ICD-O-3/WHO 2008: n=1, RX Summ--Surg Prim Site (1998+): n=8,179, Insurance Recode (2007+): n=9,431, Derived AJCC Stage Group, seventh ed (2010-2015): n=348, Derived AJCC T, seventh ed (2010-2015): n=876, Derived AJCC N, seventh ed (2010-2015):n =391, CS tumor size (2004-2015): n=338). Then, the patients who did not qualify for the third edition of the international classification of diseases for oncology (ICD-O-3) (c250-c254, c257-c259, histological code 8140, 8141) were also excluded (n=6,231). Ultimately, a total of 5,805 eligible patients were included in the present study (Figure 1). After screening, we also found some unreasonable items. Year of diagnosis only ranged from 2010 to 2015, which had no significance for the study of early prognosis. Race recode (W, B, AI, API) was consistent with Race recode (White, Black, Other). Some items only contained a single variable (Site recode ICD-O-3/WHO 2008, Behavior recode for analysis, Histologic Type ICD-O-3 PANCREAS, First malignant primary indicator). There was also a wide variation in the proportion of variables within the group of some items (Record number recode, Total number of benign/borderline tumors for patients, Regional

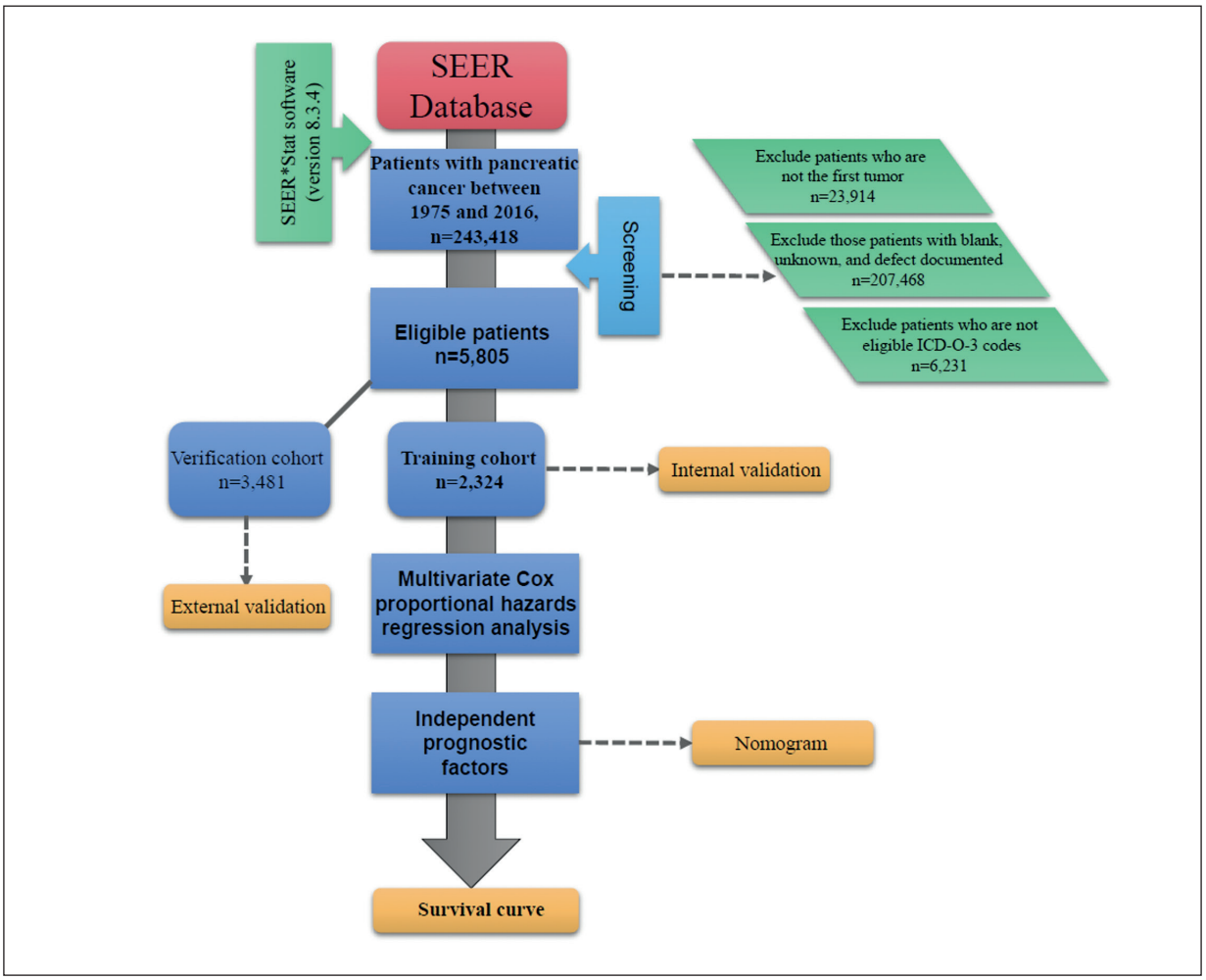

Figure 1. Downloaded 243,418 patients diagnosed with pancreatic cancer between 1975 and 2016 from the SEER database by using the SEER*Stat software (version 8.3.4). Patients with not-first tumor (23,914 cases), incomplete information (207,468 cases), and non-adenocarcinoma (6,231 cases) were excluded. A total of 5,805 eligible patients were included in the study. The patients were randomly divided into a training cohort (2,324 cases) and a verification cohort (3,481 cases) at a ratio of 4/6. Independent prognostic factors obtained in training cohort through multivariate Cox regression analysis will be internally validated in training cohort, and externally validated in verification cohort photos. The survival curve was obtained based on the independent prognostic factors.

nodes examined (1988+), Regional nodes positive (1988+). The above items were excluded for the integrity of the modeling data set.

Statistical Analysis

We randomly divided 5,805 patients into the training cohort and the verification cohort at a ratio of 4/6. Because training cohort and verification cohort were randomly composed of 5,805 patients in proportion, the data proportion between the two groups was basically the same. More details were shown in Table I. Among them, the proportion of white and other, Grade III or IV, primary site surgery, tumor-related death, and patient with a partner in training cohort was larger than that of the verification cohort. In comparison, the proportion of black, Stage I or II, radiotherapy, a small diameter tumor, patients with insurance were larger than those of the training cohort. Significant variables in the training cohort were considered as independent prognostic factors by constructing a multivariate COX regression model. These independent prognostic factors can significantly affect the prognosis of PAC. The concordance index to evaluate the nomogram's authenticity was approximately equal to the area under the ROC curve (AUC)[9], ranging from 0.5 to 1.0. The closer AUC was to 1, the higher the authenticity of the nomogram[10-12]. The calibration curve evaluated the correlation

Table I. Demographic characteristics and grouping of patients.

Variable	Total cohort, N = 5,805 (%)	Training cohort, N = 2,324 (%)	Verification cohort, N = 3,481 (%)
Age			
< 65 years	2,218 (38.2)	918 (39.5)	1,300 (37.3)
≥ 65 years	3,587 (61.8)	1,406 (60.5)	2,181 (62.7)
Race			
Black	638 (10.9)	268 (11.5)	370 (10.6)
White	4,714 (81.2)	1,856 (79.8)	2,858 (82.1)
Other (American Indian/AK Native, Asian/Pacific Islander)	453 (7.9)	200 (8.7)	253 (7.3)
Sex			
Male	3,047 (52.4)	1,219 (52.4)	1,828 (52.5)
Female	2,758 (47.6)	1,105 (47.6)	1,653 (47.5)
Grade			
Well differentiated; Grade I	609 (10.4)	248 (10.6)	361 (10.3)
Moderately differentiated; Grade II	2,570 (44.2)	1,010 (43.4)	1,560 (44.8)
Poorly differentiated; Grade III	2,544 (43.8)	1,028 (44.2)	1,516 (43.5)
Undifferentiated; anaplastic; Grade IV	82 (1.6)	38 (1.8)	44 (1.4)
Stage			
IA, IB	441 (7.6)	175 (7.5)	266 (7.6)
IIA, IIB	2,843 (49.0)	1,130 (48.6)	1,713 (49.3)
III	644 (11.0)	269 (11.6)	375 (10.8)
IV	1,877 (32.4)	750 (32.3)	1,127 (32.3)
Stage of T			
T0/T1	238 (4.0)	95 (4.0)	143 (4.1)
T2	1,047 (18.1)	419 (18.1)	628 (18.1)
T3	3,378 (58.2)	1,336 (57.5)	2,042 (58.6)
T4	1,142 (19.7)	474 (20.4)	668 (19.2)
Stage of N			
N0	2,872 (49.4)	1,164 (50.1)	1,708 (49.1)
N1	2,933 (50.6)	1,160 (49.9)	1,773 (50.9)
Stage of M			
M0	3,928 (67.6)	1,574 (67.7)	2,354 (67.6)
M1	1,877 (32.4)	750 (32.3)	1,127 (32.4)
RX Summ--Surg Prim Site			
No surgical procedure of primary site	3,119 (53.7)	1,240 (53.4)	1,879 (54.0)
Resection	2,686 (46.3)	1,084 (46.6)	1,602 (46.0)
RX Summ--Surg Oth Reg/Dis			
No surgical	5,278 (90.9)	2,109 (90.7)	3,169 (91.1)
Surgical	527 (9.1)	215 (9.3)	312 (8.9)
Radiation sequence with surgery			
None	4,795 (82.6)	1,929 (83.0)	2,866 (82.3)
Surgery with radiation	1,010 (17.4)	395 (17.0)	615 (17.7)
Radiation recode			
No radiation	4,313 (74.2)	1,735 (74.6)	2,578 (74.0)
Radiation	1,492 (25.8)	589 (25.4)	903 (26.0)
Chemotherapy			
No/Unknown	1,854 (31.9)	724 (31.2)	1,130 (32.5)
Yes	3,951 (68.1)	1,600 (68.8)	2,351 (67.5)
Size			
≤ 30 mm	2,125 (36.6)	851 (36.6)	1,274 (36.5)
> 30 mm, ≤ 50 mm	2,530 (43.6)	998 (42.9)	1,532 (44.1)
> 50 mm, ≤ 100 mm	1,092 (18.9)	450 (19.4)	642 (18.5)
> 100 mm	58 (0.9)	25 (1.1)	33 (0.9)
Extension			
≤ 100	757 (13.0)	306 (13.1)	451 (12.9)
> 100, ≤ 300	515 (8.8)	203 (8.7)	312 (9.0)
> 300, ≤ 500	2,321 (40.0)	930 (40.1)	1,391 (39.9)
> 500, ≤ 999	2,212 (38.2)	885 (38.1)	1,327 (38.2)

(Table Continued)

Table I *(Continued)*. Demographic characteristics and grouping of patients.

Variable	Total cohort, N = 5,805 (%)	Training cohort, N = 2,324 (%)	Verification cohort, N = 3,481 (%)
SEER cause-specific death classification			
Alive or dead of other cause	1,178 (20.3)	465 (20.1)	713 (20.5)
Dead (attributable to this cancer dx)	4,627 (79.7)	1,859 (79.9)	2,768 (79.5)
SEER other cause of death classification			
Alive or dead due to cancer	5,597 (96.4)	2,241 (96.4)	3,356 (96.4)
Dead (attributable to causes other than this cancer dx)	208 (3.6)	83 (3.6)	125 (3.6)
Sequence number			
One primary only	5,621 (96.8)	2,247 (96.6)	3,374 (96.9)
1st of 2 or more primaries	184 (3.2)	77 (3.4)	107 (3.1)
Total number of in situ/malignant tumors for patient			
1	5,647 (97.2)	2,258 (97.1)	3,389 (97.3)
2,3	158 (2.8)	66 (2.9)	92 (2.7)
Age at diagnosis			
< 65 years	2,218 (38.2)	918 (39.5)	1,300 (37.3)
≥ 65 years	3,587 (61.8)	1,406 (60.5)	2,181 (62.7)
Insurance Recode			
Uninsured	205 (3.5)	87 (3.7)	118 (3.4)
Insured	5,600 (96.5)	2,237 (96.3)	3,363 (96.6)
Marital status at diagnosis			
Single (never married)	689 (11.8)	275 (11.8)	414 (11.9)
Unmarried or Domestic Partner/Married (including common law)	3,702 (63.7)	1,509 (64.9)	2,193 (62.9)
Separated/Divorced/Widowed	1,414 (24.5)	540 (23.3)	874 (25.2)

between nomogram and real value. $p < 0.05$ was believed to be statistically significant. The above statistics were analyzed using R (version 3.5.3) software.

Results

Demographic Characteristics

We included a total of 5,805 PAC patients from the SEER database. These patients were diagnosed with PAC from 2010 to 2015, which guaranteed follow-up for five years or more. All patients were randomly divided into a training cohort (2,324 cases) and a verification cohort (3,481 cases). In the total cohort, the majority of patients were over 65 years old (61.8%), white (81.2%), not a single (63.7%), diagnosed over 65 years old (61.8%), insured (96.5%), and they had only one primary malignant tumor (97.2%). A large proportion of the PAC patients were in stage IIA, IIB (49.0%), and IV (32.4%), less than 50 mm in diameter (80.2%), in stage of T3 (58.2%). No metastasis (67.6%) and CS extension (2004-2015) more than 300 (78.2%) also accounted for a large proportion. The PAC did not invade regional lymph nodes (49.4%) or only affected a few nearby lymph nodes (50.6%). Grade II and III tumors accounted for 44.2% and 43.8%, respectively. In terms of treatment, less than half of the patients underwent primary site resection (46.3%). The majority of patients also did not receive excision outside the primary site (90.9%). Chemotherapy (68.1%) was more frequent than radiotherapy (25.8%). PAC-specific death accounted for the vast majority (79.7% and 96.4%). Since the training cohort and verification cohort were composed of the same data set randomly assigned, there was little difference between them and the total group. The details were shown in Table I.

Univariate and Multivariate Cox Proportional Hazards Regression Analyses

By performing the univariate and multivariate Cox proportional hazards regression analyses, we identified age, chemotherapy, grade, Radiation sequence with surgery, Radiation recode, RX Summ--Surg Prim Site, size, and American Joint Committee for Cancer (AJCC) stage as the independent prognostic factors that significantly affected the OS of PAC. Among them, the TNM staging system based on AJCC mainly describes the development of the primary tumor (stage of T), regional lymph node involvement (stage of N) and distant metastasis (stage of M). After the

Table II. Univariate Cox proportional hazards regression analysis for patients with pancreatic adenocarcinoma in training cohort.

Variable	HR	95.0% CI	p
Age			
< 65 years	1	Ref	
≥ 65 years	1.33	1.22-1.46	< 0.001
Race			
Black	1	Ref	
White	0.96	0.84-1.11	0.600
Other (American Indian/AK Native, Asian/Pacific Islander)	0.96	0.79-1.18	0.721
Sex			
Male	1	Ref	
Female	0.97	0.88-1.06	0.45
Grade			
Well differentiated; Grade I	1	Ref	
Moderately differentiated; Grade II	0.99	0.85-1.16	0.942
Poorly differentiated; Grade III	1.45	1.24-1.70	< 0.001
Undifferentiated; anaplastic; Grade IV	1.42	0.98-2.05	0.065
Stage			
IA, IB	1	Ref	
IIA, IIB	1.07	0.89-1.29	0.482
III	1.79	1.45-2.23	< 0.001
IV	3.39	2.80-4.11	< 0.001
Stage of T			
T0/T1	1	Ref	
T2	2.58	1.96-3.40	< 0.001
T3	1.56	1.20-2.03	0.001
T4	2.75	2.10-3.62	< 0.001
Stage of N			
N0	1	Ref	
N1	0.80	0.73-0.87	< 0.001
Stage of M			
M0	1	Ref	
M1	2.91	2.65-3.21	< 0.001
RX Summ--Surg Prim Site			
No surgical procedure of primary site	1	Ref	
Resection	0.29	0.26-0.32	< 0.001
RX Summ--Surg Oth Reg/Dis			
No surgical	1	Ref	
Surgical	0.62	0.53-0.74	< 0.001
Radiation sequence with surgery			
None	1	Ref	
Surgery with radiation	0.98	0.33-0.43	< 0.001
Radiation recode			
No radiation	1	Ref	
Radiation	0.51	0.46-0.56	< 0.001
Chemotherapy			
No/Unknown	1	Ref	
Yes	0.44	0.40-0.48	< 0.001
Size			
≤ 30 mm	1	Ref	
> 30 mm, ≤ 50 mm	1.42	1.29-1.58	< 0.001
> 50 mm, ≤ 100 mm	2.13	1.88-2.41	< 0.001
> 100 mm	2.42	1.60-3.67	< 0.001
Extension			
≤ 100	1	Ref	
> 100, ≤ 300	1.28	1.06-1.55	0.010
> 300, ≤ 500	0.75	0.65-0.86	< 0.001
> 500, ≤ 999	1.22	1.06-1.40	0.007
SEER cause-specific death classification			
Alive or dead of other cause	1	Ref	
Dead (attributable to this cancer dx)	14.78	11.80-18.51	< 0.001

(Table Continued)

Table II *(Continued)*. Univariate Cox proportional hazards regression analysis for patients with pancreatic adenocarcinoma in training cohort.

Variable	HR	95.0% CI	*p*
SEER other cause of death classification			
Alive or dead due to cancer	1	Ref	
Dead (attributable to causes other than this cancer dx)	1.38	1.11-1.72	0.004
Sequence number			
One primary only	1	Ref	
1st of 2 or more primaries	0.49	0.37-0.65	< 0.001
Total number of in situ/malignant tumors for patient			
1	1	Ref	
2,3	0.51	0.38-0.68	< 0.001
Age at diagnosis			
< 65 years	1	Ref	
≥ 65 years	1.33	1.22-1.46	< 0.001
Insurance Recode			
Uninsured	1	Ref	
Insured	1.00	0.79-1.27	0.992
Marital status at diagnosis			
Single (never married)	1	Ref	
Unmarried or Domestic Partner/Married (including common law)	0.92	0.80-1.06	0.238
Separated/Divorced/Widowed	1.10	0.94-1.29	0.253

Ref = Reference.

determination of T, N and M in TNM stages, the corresponding total stages can be obtained, namely, stage I, stage II, stage III, stage IV. Details are as follows: IA: T1 N0 M0, IB: T2 N0 M0, IIA: T3 N0 M0, IIB: T1-T3 N1 M0, III: T4 N-any M0, IV: T-any N-any M1. More details were shown in Tables II and III.

The Construction, Verification, and Predictive Value of Nomogram

Nomogram based on the training cohort's independent prognostic factors was shown in Figure 2. Each level of each variable was projected up to the points to get the corresponding score. The total of each score was the total points. The 1-year, 3-year, and 5-year survival rates of the patient can be obtained by projecting the total points downward. The higher the total points, the worse the prognosis. Nomogram was internally verified in the training cohort and was externally verified in the verification cohort. The concordance index can be applied to evaluate the difference between the model and the actual value. The higher the concordance index, the more accurate nomogram's prediction. The concordance index (training cohort: 0.762, verification cohort: 0.760) indicated our nomogram's good accuracy value. In a wholly calibrated model, the prediction should fall on the diagonal slope of 1 in the calibration curve[13]. Calibration curves for predicting 1-, 3-and 5-year survival rates in the training cohort (Figure 3 A, B, C) and verification cohort (Figure 3 D, E, F) showed right consistency. The survival area under the ROC curve (AUC) ranged from 0.71 to 0.90, suggesting the survival model was of moderate accuracy. The survival AUC (0.828, 0.811, 0.819) of the training cohort (Figure 4 A, B, C) and the survival AUC (0.822, 0.815, 0.802) of the verification cohort (Figure 4 D, E, F) were all of good predictive value.

Survival Analysis

As shown in Figure 5 A-C, young age, low differentiation tumor, small diameter tumor, AJCC stage IA, IB, radiotherapy, chemotherapy, and primary site resection contributed to the OS. Mean survival time was 12.5 months in the age group younger than 65 and was 8.6 months in the age group equal or older than 65. The mean survival time of Grade I, II, III and IV were 13.4 months, 12.5 months, 7.7 months, and 7 months, respectively. The mean survival time of Stage I, II, III and IV were 15.1 months, 15.8 months, 9.2 months, and 3.7 months, respectively. The mean survival time after surgery, radiotherapy, and chemotherapy was 20.2 months, 17.7 months, and 13.5 months, respectively. However, the mean survival time without surgery, radiotherapy or chemotherapy was 5.0 months, 7.6 months, and 2.8 months, respectively. Also, the survival curve of the risk score of multivar-

Table III. Multivariate Cox proportional hazards regression analysis for patients with pancreatic adenocarcinoma in training cohort.

Variable	HR	95.0% CI	p
Age			
< 65 years	1	Ref	
≥ 65 years	1.12	1.02-1.23	0.019
Grade			
Well differentiated; Grade I	1	Ref	
Moderately differentiated; Grade II	1.09	0.93-1.28	0.289
Poorly differentiated; Grade III	1.35	1.15-1.57	< 0.001
Undifferentiated; anaplastic; Grade IV	1.38	0.95-2.00	0.091
Stage			
IA, IB	1	Ref	
IIA, IIB	1.26	0.99-1.61	0.056
III	1.11	0.83-1.49	0.481
IV	1.77	1.40-2.24	< 0.001
Stage of T			
T0/T1	1	Ref	
T2	0.97	0.72-1.31	0.847
T3	0.60	0.25-1.47	0.267
T4	0.56	0.23-1.36	0.200
Stage of N			
N0	1	Ref	
N1	0.97	0.88-1.07	0.521
Stage of M			
M0	1	Ref	
M1	NA	NA	NA
RX Summ--Surg Prim Site			
No surgical procedure of primary site	1	Ref	
Resection	0.45	0.39-0.52	< 0.001
RX Summ--Surg Oth Reg/Dis			
No surgical	1	Ref	
Surgical	0.99	0.84-1.18	0.938
Radiation sequence with surgery			
None	1	Ref	
Surgery with radiation	1.35	1.08-1.69	0.008
Radiation recode			
No radiation	1	Ref	
Radiation	0.73	0.61-0.86	< 0.001
Chemotherapy			
No/Unknown	1	Ref	
Yes	0.42	0.37-0.46	< 0.001
Size			
≤ 30 mm	1	Ref	
> 30 mm, ≤ 50 mm	1.14	1.02-1.26	0.022
> 50 mm, ≤ 100 mm	1.52	1.33-1.74	< 0.001
> 100 mm	1.63	1.06-2.48	0.025
Extension			
≤ 100	1	Ref	
> 100, ≤ 300	0.98	0.81-1.19	0.855
> 300, ≤ 500	1.49	0.59-3.81	0.400
> 500, ≤ 999	1.63	0.64-4.12	0.302
SEER cause-specific death classification			
Alive or dead of other cause	1	Ref	
Dead (attributable to this cancer dx)	1.44*108	1.12*10-315-Inf	0.961
SEER other cause of death classification			
Alive or dead due to cancer	1	Ref	
Dead (attributable to causes other than this cancer dx)	1.58*108	1.22*10-315-Inf	0.960
Sequence number			
One primary only	1	Ref	
1st of 2 or more primaries	0.55	0.25-1.24	0.148
Total number of *in situ*/malignant tumors for patient			
1	1	Ref	
2,3	1.19	0.50-2.79	0.696
Age at diagnosis			
<65 years	1	Ref	
≥65 years	NA	NA	NA

Ref = Reference; Inf = Infinite; NA = No Answer.

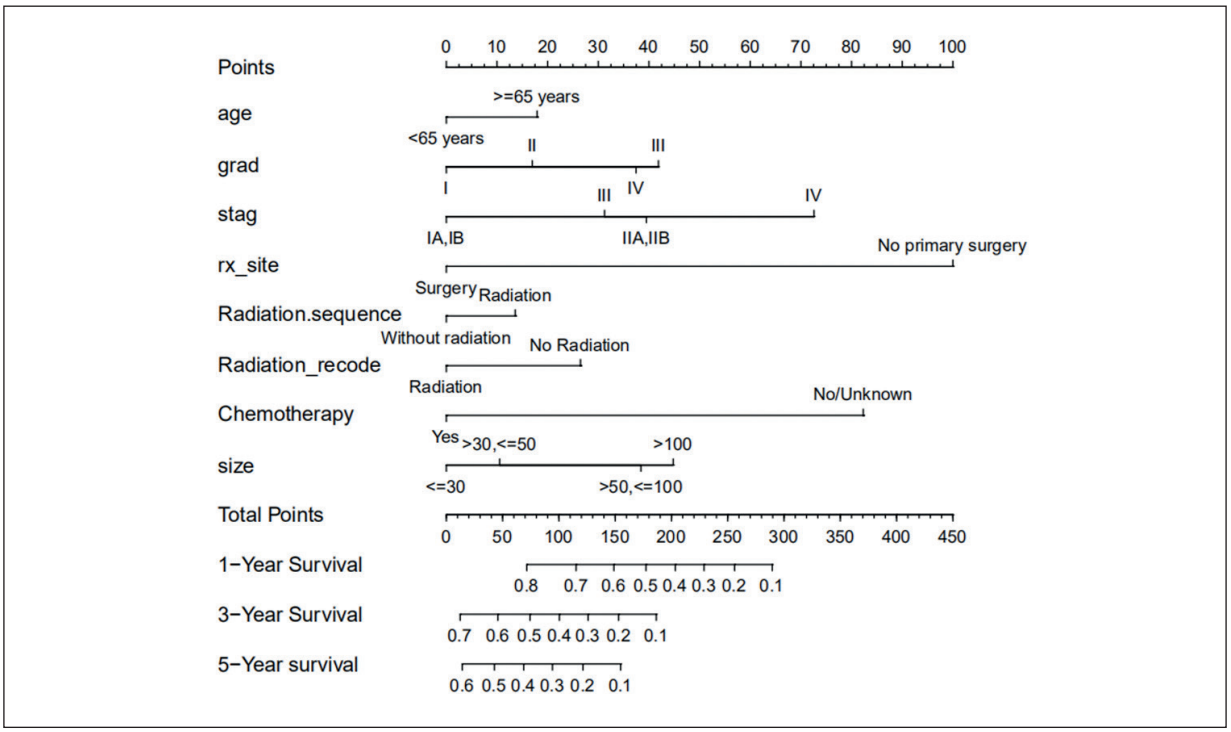

Figure 2. Nomogram for predicting 1-year, 3-year and 5-year survival of patients with pancreatic adenocarcinoma based on eight factors, including age, chemotherapy, grade (grad), Radiation sequence with surgery (Radiation.sequence), Radiation recode (Radiation_recode), RX Summ--Surg Prim Site (rx_site), size and AJCC(American Joint Committee for Cancer) stage (stag).

iate Cox proportional hazards regression analysis in the training cohort was shown in Figure 6. The overall mean survival time was 15.1 months. The overall median survival time was 10.0 months, and the overall survival rate was 16.7%. For high-risk patients, the mean survival time was 6.8 months, the median survival time was 3.8 months, and the survival rate was 0. For

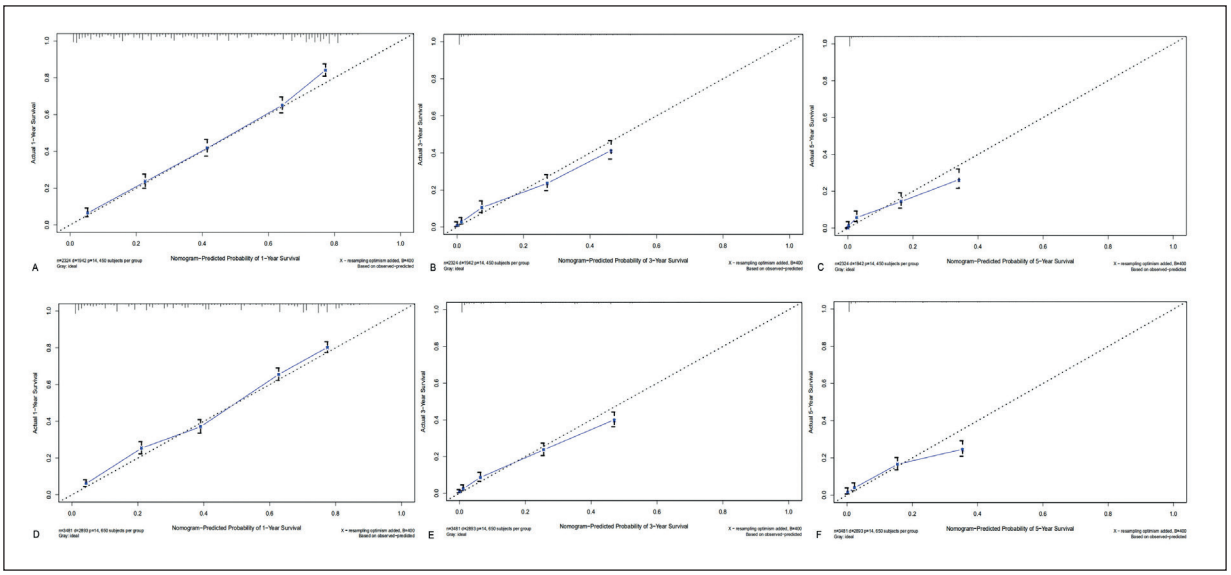

Figure 3. Calibration plots for predictions of the 1(**A,**)-, 3(**B,**)-, 5(**C,**)-year survivals in training cohort. Calibration plots for predictions of the 1(**D,**)-, 3(**E,**)-, 5(**F,**)-year survivals in verification cohort.

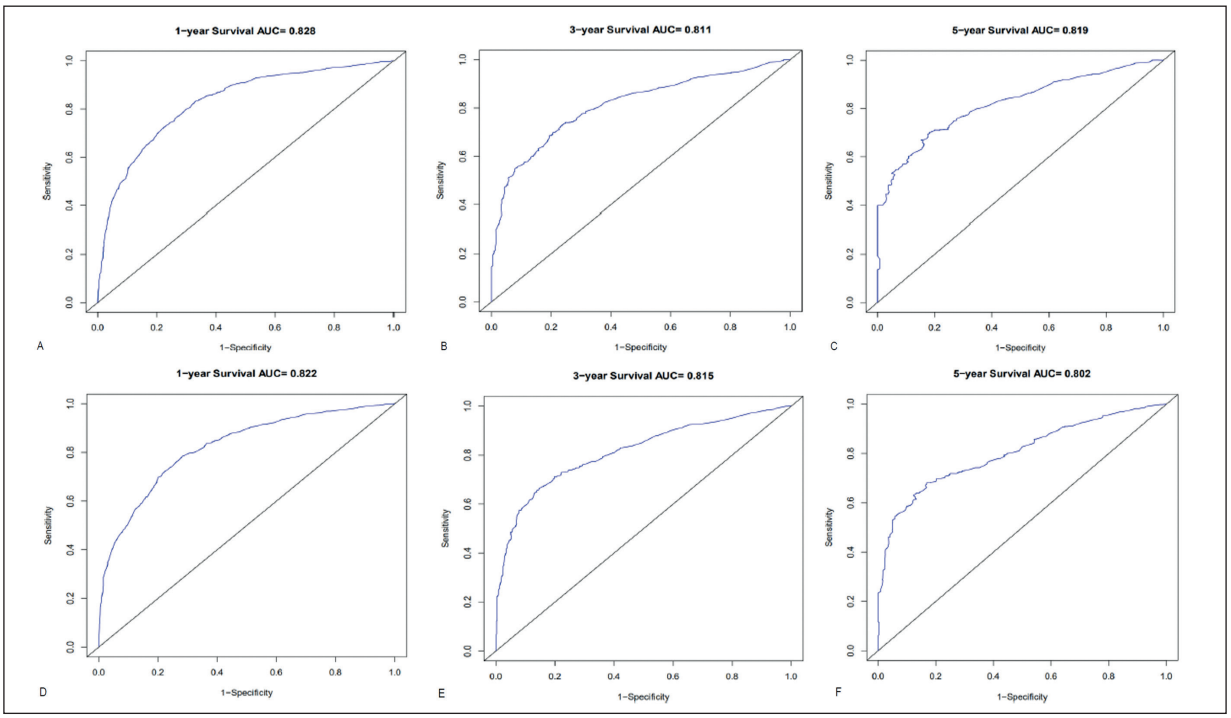

Figure 4. ROC curve to evaluate the nomogram. The 1(**A,** 0.828)-, 3(**B,** 0.811)-, 5(**C,** 0.819)-year AUC of training cohort and 1(**D,** 0.822)-, 3(**E,** 0.815)-, 5(**F,** 0.802)-year AUC of VERIFICATION cohort in the present study all indicated that the nomogram had good predictive performance. Note: AUC=0.5 means no prediction ability, AUC=0.51-0.7 means low accuracy, AUC=0.71-0.9 means medium accuracy, AUC>0.9 means high accuracy. AUC, area under the ROC curve.

low-risk patients, the mean survival time was 23.3 months, the median survival time was 21.4 months, and the survival rate was 33.5%.

Discussion

The PAC was a rapidly fatal malignancy with a mortality rate significantly close to its incidence[14]. Therefore, it was necessary to use a model to predict the prognosis of PAC patients in order to guide the work of clinicians. At present, nomograms have been applied to predict the OS of different types of cancers for its good predictive value[8]. We developed a PAC nomogram based on the SEER database. The nomogram was built on the eight independent prognostic factors in the training cohort, including age, chemotherapy, grade, Radiation sequence with surgery, Radiation recode, RX Summ-Surg Prim at Site, the size, and AJCC stage.

Our nomogram showed RX Summ-Surg Prim Site had the most significant influence on OS, followed by chemotherapy, AJCC stage, the size, grade, Radiation recode, age, Radiation sequence with surgery. 1- year survival was less than 0.4 and 3-year survival was less than 0.1 of PAC patients without primary site surgery, radiotherapy, or chemotherapy. Aged more than sixty-five, high-grade tumor, and large diameter were risk factors for the prognosis of PAC patients.

Presently, we have conducted several surveys. The PAC was a cancer of the elderly, most commonly occurring between the ages of 60 and 80[15,16]. Agedness was one reason why patients were refused surgery by surgeons[17]. If age-related complications can be effectively monitored, the prognoses of older patients in surgery could be similar to that of younger patients[18]. Tumor grade was a risk factor for OS after tumor resection[19]. The TNM staging system was a frequently used tool. Some researchers reported that the TNM staging system occurred defects, and the combination of grade and TNM to form a new TNMG system can enhance the predictive ability of the tool[20]. Primary site resection was beneficial for the prognosis of pancreatic cancer[21]. Surgery remained the only cure[22], which mainly involved local excision of tumor, partial pancreatectomy, total pancreatectomy, total pancreatectomy and

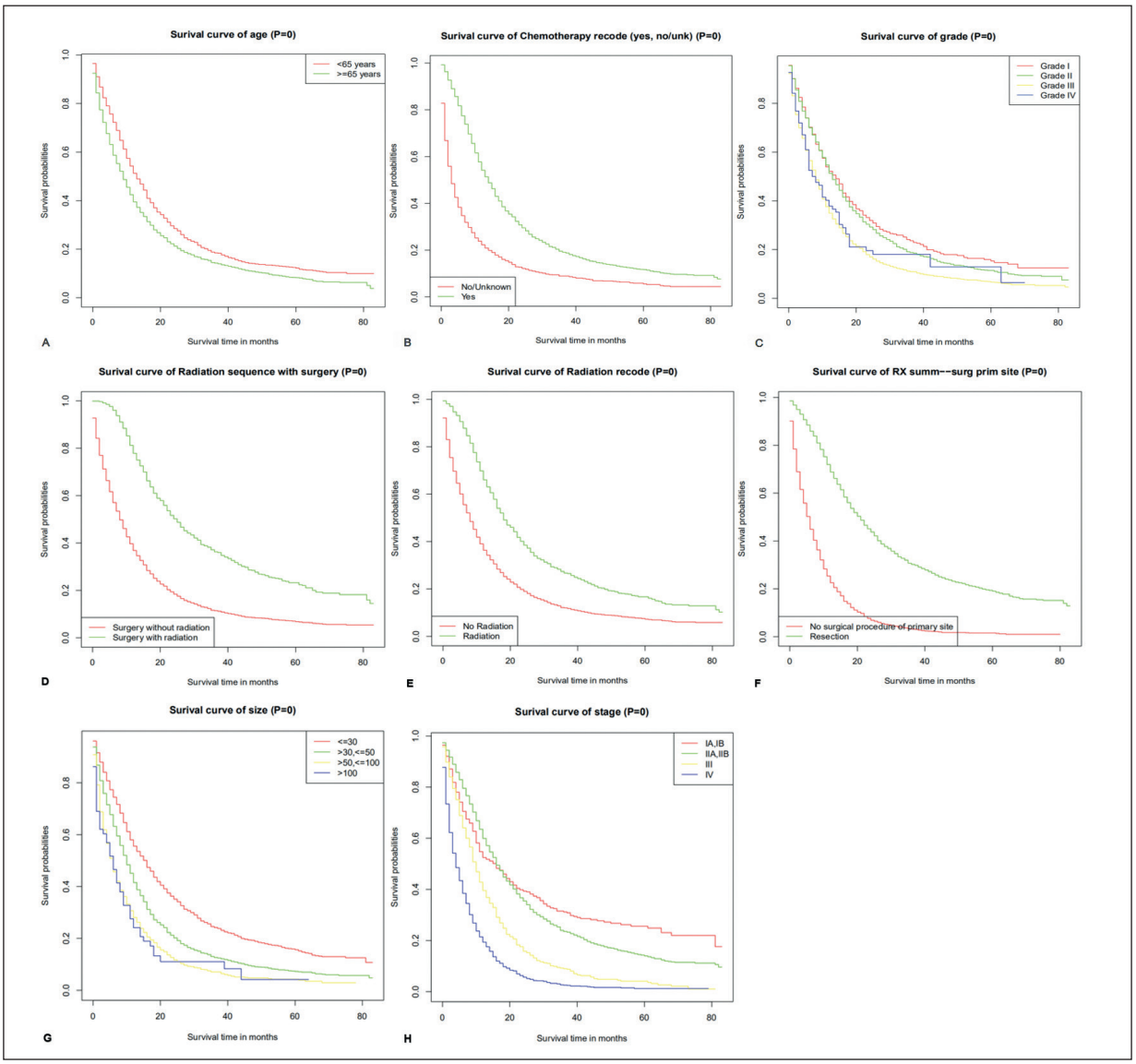

Figure 5. The prognosis and survival stratification analyses of PAC patients with **A,** (age), **B,** (Chemotherapy recode (yes, no/unk)), **C,** (grade), **D,** (Radiation sequence with surgery), **E,** (Radiation recode), **F,** (RX Summ--Surg Prim Site), G (size) and **H,** (stage).

subtotal gastrectomy or duodenectomy, and extended pancreatoduodenectomy. Among them, pancreatectomy was a fantastic choice[21]. Still, many patients were not ready for surgery, depending to a large extent on the degree of arterial invasion. It was reported[23] that the survival rates were higher in patients with arterectomy. Size had turned out to be an independent prognostic factor in OS after pancreatic cancer resection[24]. Mayo et al[25] had shown that the diameter of three centimeter was the boundary line greatly affecting prognosis. The tumor diameter of two to three centimeters was an appropriate range for tumor resection. Radiotherapy and chemotherapy were favorable factors to improve OS in pancreatic cancer[26]. However, the effect of surgical alone or adjuvant therapy (radiotherapy or chemotherapy) alone was not ideal. The combination of adjuvant therapy (radiotherapy and chemotherapy) can significantly improve patient survival[27]. These reports were in agreement with our results.

The biggest innovation of our research was the selection and processing of data. Under the tendency of individualized treatment, our

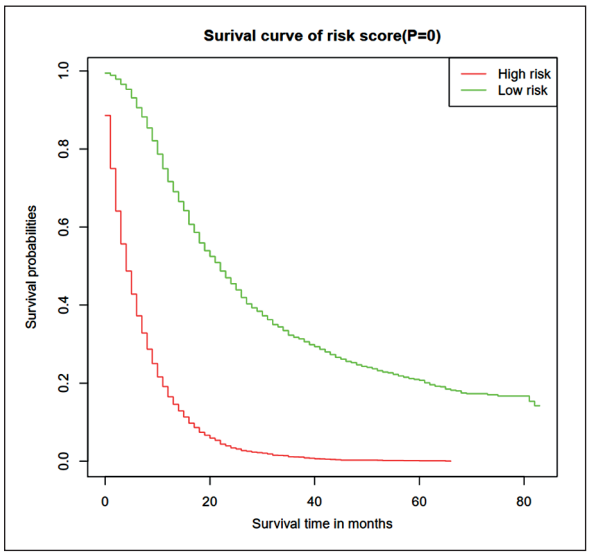

Figure 6. According to the risk scores, the overall survival curves of PAC patients were obtained.

work was one of the few studies to include the whole SEER database pancreatic cancer patients (243,418 cases), making the results largely representative of the survival of pancreatic cancer patients with current treatments. In addition, our processing of the raw data further enhanced the application value of the consequences. First of all, due to the long span (1975-2016) of the raw data, changes in medical technology during this period had significantly changed the survival rates of PAC patients. However, the rigorous screening made the included patients only range from 2010 to 2015, which avoided the above differences and ensured the timeliness of the data. Second, 5,805 patients were randomly divided into training cohort and the verification cohort, and used for internal and external validation, respectively. To some extent, this avoided interference with the results due to the lack of external database validation. In general, our selection and processing of the data allowed the results to serve as a reference standard for clinicians to use when formulating individualized treatment plans for PAC patients.

The present study also had some limitations. First, only 5,805 (2.4%) of the original 243,418 patients complied with the inclusion criteria, causing the selection bias. Second, white (81.2%) were significantly more than other races, leading to uncertainty about whether the results could be applied to other countries. Third, coding choices can lead to biasing results.

Conclusions

We established a nomogram to predict 1-year, 3-year, and 5-year survival in PAC patients. The nomogram was based on the independent prognostic factors, including age, chemotherapy, grade, Radiation sequence with surgery, Radiation recode, RX Summ--Surg Prim Site, size, and AJCC stage. Our nomogram was of good predictive power and practicality.

Conflict of Interest

The Authors declare that they have no conflict of interests.

Contributorship Statement

JL extracted the data, wrote the main manuscript and performed most statistical analysis. XZ assisted the statistical analysis. SW, BL and XZ edited the format of the manuscript, tables and figures. XZ performed quality control as a senior export in statistics and funded the manuscript. All authors have read and approved the manuscript.

Funding

This work was supported partly by Guangdong Science and Technology Department (2016B030309002 and 2019B090905011); Guangdong Province Natural Science Foundation (2018A030310155); the Fund of Southern Marine Science and Engineering Guangdong Laboratory (Zhanjiang) (ZJW-2019-007); the Science and Technology Program of Zhanjiang (2017A06012); and The Public Service Platform of South China Sea for R&D Marine Biomedicine Resources (GDMUK201808).

Ethics Approval and Consent to Participate

The work was approved by Guangdong Medical University Ethics Committee (YS2019131). Informed consent forms are not required for patient data extracted from SEER database.

References

1) Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015; 136: E359-386.
2) Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA Cancer J Clin 2015; 65: 5-29.
3) Malvezzi M, Carioli G, Bertuccio P, Rosso T, Boffetta P, Levi F, La Vecchia C, Negri E. European cancer mortality predictions for the year 2016 with focus on leukaemias. Ann Oncol 2016; 27: 725-731.

4) Qiu D, Katanoda K, Marugame T, Sobue T. A Joinpoint regression analysis of long-term trends in cancer mortality in Japan (1958-2004). Int J Cancer 2009; 124: 443-448.

5) Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA Cancer J Clin 2017; 67: 7-30.

6) Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. Cancer Res 2014; 74: 2913-2921.

7) Goral V. Pancreatic cancer: pathogenesis and diagnosis. Asian Pac J Cancer Prev 2015; 16: 5619-5624.

8) Wang Y, LI J, Xia Y, Gong R, Wang K, Yan Z, Wan X, Liu G, Wu D, Shi L, Lau W, Wu M, Shen F. Prognostic nomogram for intrahepatic cholangiocarcinoma after partial hepatectomy. J Clin Oncol 2013; 31: 1188-1195.

9) Chen Y, Cao D, LI C, Zhang P, Wang X, LI N, Han Q, Liu Z. A nomogram for discrimination of non-alcoholic fatty liver disease in patients with chronic hepatitis B. Eur J Gastroenterol Hepatol 2020 Feb 27. doi: 10.1097/MEG.0000000000001691. Online ahead of print.

10) Huitzil-Melendez FD, Capanu M, O'reilly EM, Duffy A, Gansukh B, Saltz LL, Abou-Alfa GK. Advanced hepatocellular carcinoma: which staging systems best predict prognosis? J Clin Oncol 2010; 28: 2889-2895.

11) Li H, He Y, Huang L, Luo H, Zhu X. The nomogram model predicting overall survival and guiding clinical decision in patients with glioblastoma based on the SEER database. Front Oncol 2020; 10: 1051.

12) Liang B, Yu H, Huang L, Luo H, Zhu X. A prognostic model guides surgical resection in cervical squamous cell carcinoma. Transl Cancer Res 2020; 9: 1711-1731.

13) Zhang Y, Zhang J, Zeng H, Zhou XH, Zhou HB. Nomograms for predicting the overall and cancer-specific survival of patients with classical Hodgkin lymphoma: a SEER-based study. Oncotarget 2017; 8: 92978-92988.

14) Zhang J, Dhakal I, Yan H, Phillips M, Kesteloot H, Registries SC. Trends in pancreatic cancer incidence in nine SEER Cancer Registries, 1973-2002. Ann Oncol 2007; 18: 1268-1279.

15) Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394-424.

16) Bosetti C, Bertuccio P, Negri E, La Vecchia C, Zeegers MP, Boffetta P. Pancreatic cancer: overview of descriptive epidemiology. Mol Carcinog 2012; 51: 3-13.

17) Burmeister EA, Waterhouse M, Jordan SJ, O'Connell DL, Merrett ND, Goldstein D, Wyld D, Beesley V, Gooden H, Janda M, Neale RE. Determinants of survival and attempted resection in patients with non-metastatic pancreatic cancer: An Australian population-based study. Pancreatology 2016; 16: 873-881.

18) Li X, Liu Z, Ye Z, Gou S, Wang C. Impact of age on survival of patients with pancreatic cancer after surgery: Analysis of SEER data. Pancreatology 2018; 18: 133-138.

19) Winter JM, Cameron JL, Campbell KA, Arnold MA, Chang DC, Coleman J, Hodgin MB, Sauter PK, Hruban RH, Riall TS, Schulick RD, Choti MA, Lillemoe KD, Yeo CJ. 1423 pancreaticoduodenectomies for pancreatic cancer: a single-institution experience. J Gastrointest Surg 2006; 10: 1199-1210; discussion 1210-1191.

20) Rochefort MM, Ankeny JS, Kadera BE, Donald GW, Isacoff W, Wainberg ZA, Hines OJ, Donahue TR, Reber HA, Tomlinson JS. Impact of tumor grade on pancreatic cancer prognosis: validation of a novel TNMG staging system. Ann Surg Oncol 2013; 20: 4322-4329.

21) Gurusamy KS, Kumar S, Davidson BR, Fusai G. Resection versus other treatments for locally advanced pancreatic cancer. Cochrane Database Syst Rev 2014; (2): CD010244.

22) Hidalgo M. Pancreatic cancer. N Engl J Med 2010; 362: 1605-1617.

23) Mollberg N, Rahbari NN, Koch M, Hartwig W, Hoeger Y, Büchler MW, Weitz J. Arterial resection during pancreatectomy for pancreatic cancer: a systematic review and meta-analysis. Ann Surg 2011; 254: 882-893.

24) Li Y, Tang CG, Zhao Y, Cao WY, Qu GF. Outcomes and prognostic factors of patients with stage IB and IIA pancreatic cancer according to the 8th edition American Joint Committee on Cancer criteria. World J Gastroenterol 2017; 23: 2757-2762.

25) Mayo SC, Nathan H, Cameron JL, Olino K, Edil BH, Herman JM, Hirose K, Schulick RD, Choti MA, Wolfgang CL, Pawlik TM. Conditional survival in patients with pancreatic ductal adenocarcinoma resected with curative intent. Cancer 2012; 118: 2674-2681.

26) Chen-Zhao X, Hernando O, López M, Sánchez E, Montero A, García-Aranda M, Ciérvide R, Valero J, Alonso R, Cárdenas-Rebollo JM, Vicente E, Quijano Y, Cubillo A, Álvarez R, Prados S, Plaza C, García J, Zucca D, Fernández-Letón P, Rubio C. A prospective observational study of the clinical and pathological impact of stereotactic body radiotherapy (SBRT) as a neoadjuvant strategy of chemoradiation in pancreatic cancer. Clin Transl Oncol 2020.

27) Parmar AD, Vargas GM, Tamirisa NP, Sheffield KM, Riall TS. Trajectory of care and use of multimodality therapy in older patients with pancreatic adenocarcinoma. Surgery 2014; 156: 280-289.