

# Development and validation of a Nomogram for predicting the Overall Survival of Patients With Pancreatic Ductal Adenocarcinoma of the Head of the Pancreas

**Zhilong Liu**

First Affiliated Hospital of Jinan University

**Haohui Yu**

First Affiliated Hospital of Jinan University

**Mingrong Cao**

First Affiliated Hospital of Jinan University

**Jiexing Li**

First Affiliated Hospital of Jinan University

**Yulin Huang**

First Affiliated Hospital of Jinan University

**Bin Feng**

First Affiliated Hospital of Jinan University

**Liubo Xing**

First Affiliated Hospital of Jinan University

**Ke Cai (✉ caike0811@sina.com)**

First Affiliated Hospital of Jinan University

---

## Research Article

**Keywords:** Pancreatic Ductal Adenocarcinoma of the Head of the Pancreas (PDAC-HP), overall survival (OS)

**Posted Date:** March 17th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-258185/v1>

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

[Read Full License](#)

---

## Abstract

**Background:** The purpose of this study is to develop and validate a nomogram to predict the overall survival (OS) of patients with Pancreatic Ductal Adenocarcinoma of the Head of the Pancreas (PDAC-HP).

**Methods:** Using the Surveillance, Epidemiology, and End Results (SEER) database, we collected patients with PDAC-HP in the United States between 2004 and 2015. Patients were randomly divided into training set and validating set at a ratio of 7:3. The training set is used to develop a nomogram for predicting OS. These indicators such as the C index, the area under curve (AUC) of the receiver operating characteristic (ROC), calibration plots and the net reclassification improvement (NRI) and the integrated discrimination improvement (IDI) were used to evaluate the prediction accuracy of the nomogram.

**Results:** A total of 33,893 patients with PDAC-HP over 20 years old were diagnosed between 2004 and 2015 were collected from the SEER database. Using multivariable Cox regression analysis, we identified eight risk factors that were associated with OS, such as age at diagnosis, sex, marital status at diagnosis, race, AJCC staging, surgery, radiotherapy and chemotherapy. A nomogram was constructed based on these variables. Compared with the AJCC staging system, the nomogram has a better C index and AUC in the training set and validating set. The calibration plots indicated that the nomogram was able to accurately predict the OS of patients with PDAC-HP at 1, 3, and 5 years.

**Conclusions:** We developed and validated a nomogram, and predicted the OS of patients with PDAC-HP at 1, 3, and 5 years. Compared with the AJCC staging system, the nomogram we constructed has better performance. It shows that our nomogram could be served as an effective tool for prognostic evaluation of patients with PDAC-HP.

## Background

Pancreatic cancer is a highly malignant tumor of the digestive tract with a five-year survival rate of approximately 8%.<sup>[1, 2]</sup> It is reported that there were approximately 460,000 new patients with pancreatic cancer globally in 2018, and their incidence is second only to colorectal cancer and rectal cancer in the digestive system, and about 430,000 patients died of the disease.<sup>[1]</sup> There were approximately 55,000 new patients with pancreatic cancer and 44,000 deaths in the United States in 2018.<sup>[3]</sup> It is the tenth most common new tumor and the fourth most common cause of death in the United States.<sup>[4]</sup> Pancreatic ductal adenocarcinoma (PDAC) is the main pathological type of pancreatic cancer, accounting for about 90%, and its degree of malignancy is much higher than other types.<sup>[5]</sup>

The pancreas and its accessory tissues are rich in blood vessels and lymph, and pancreatic cancer is prone to metastasis. Therefore, most patients are diagnosed as advanced pancreatic cancer at the first visit, and their prognosis is poor.<sup>[5, 6]</sup> For PDAC patients, surgery is the first choice of treatment, followed by radiotherapy and chemotherapy.<sup>[7]</sup> Due to the low detection rate of early PDAC patients, most patients

have already developed tumor metastasis after the diagnosis of the disease.<sup>[5, 8]</sup> It was found that although surgery can significantly prolong the survival time of PDAC patients, the 5-year survival rate of patients after surgery is only about 20%, due to the high recurrence rate of patients after surgery.<sup>[9–11]</sup> At present, the tumor staging and prognosis of PDAC patients are mainly based on the AJCC staging system.<sup>[12]</sup> However, the AJCC staging system does not take into account the influence of age, gender, race, surgery and other factors on tumor prognosis.<sup>[13, 14]</sup> Therefore, constructing a prognostic prediction model with high accuracy and specificity is still of great significance for improving the prognosis of PDAC patients and guiding the treatment of patients.

A nomogram is a predictive model that can combine multiple risk factors.<sup>[14]</sup> In recent years, the nomogram has been widely used to predict the overall survival (OS) of different types of cancer patients.<sup>[15–17]</sup> At present, there are few studies using the nomogram to study the OS of patients with pancreatic ductal adenocarcinoma of the head of the pancreas (PDAC-HP). Therefore, the purpose of this study is to explore the potential risk factors that affect the OS of patients with PDAC-HP and construct a nomogram model to predict the OS of the patients.

## Method

### Data source

In this study, the data of patients with PDAC-HP in the United States from 2004 to 2015 were derived from the surveillance, epidemiology and end result (SEER) program of the National Cancer Institute. The SEER database collected a variety of cancer data for about 30% of the population in the United States, including demographics, tumor characteristics, and survival data.<sup>[18]</sup> It provides relevant data to all researchers free of charge.

We obtained the license to use the SEER database by signing the SEER Research Data Agreement.

### Study population and inclusion criteria

Using SEER\*Stat (version 8.3.8) software, we identified patients with PDAC-HP from the SEER 18 database based on the code in the International Taxonomy of Tumors (third edition). The main site code of PDAC-HP is C25.0, and the histology code is 8140. The criteria for patient inclusion were 1) Patients with PDAC-HP diagnosed between 2004 and 2015; 2) The age at diagnosis was over 20 years old. The exclusion criteria included: 1) Age under 20 at diagnosis or unknown age; 2) No follow-up date; 3) Unknown race, 4) Unknown AJCC staging. Finally, we collected 33,893 adult patients with PDAC-HP from the SEER database, who were randomly divided into a training set (70%) and a validating set (30%). The screening process of patients with PDAC-HP is shown in eFigure 1 in the Supplement. The data collected by all patients included sex, marital status at diagnosis, age at diagnosis, race, region, AJCC staging, surgery, radiotherapy, chemotherapy, and survival time, etc.

### Statistical analysis

We screened out independent risk factors that affect the OS of patients with PDAC-HP using stepwise COX proportional hazards regression analysis. The Akaike Information Criteria (AIC) was used to screen out the final predictor variables and build a nomogram model. The discrimination of the nomogram was evaluated by Harrell's C index (C statistic), and a bootstrap resampling with 1000 iterations was applied to verify the accuracy of the nomogram. The constructed nomogram model is internally compared with the C index of the AJCC staging system (sixth edition). The area under curve (AUC) of the receiver operating characteristic (ROC) was used to evaluate the accuracy of the nomogram in 1-, 3- and 5-year survival predictions, and the calibration plots was used to evaluate the performance of the nomogram. Using the net reclassification improvement (NRI) and the integrated discrimination improvement (IDI), we evaluated the accuracy of the nomogram which was compared with that of the AJCC staging system. Finally, the decision-curve analysis (DCA) was used to evaluate the clinical validity of the nomogram. All statistical analyses were performed using R software (version 3.6.3) (<http://www.r-project.org/>). All tests are two-sided tests, and  $P < 0.05$  is considered statistically significant.

## Results

### Patient basic characteristics

Table 1 shows the basic characteristics of the patients with PDAC-HP. In this study, we collected 33,893 patients with PDAC-HP over 20 years old were diagnosed between 2004 and 2015. These patients were randomly divided into training set and validating set, including 23,725 patients in the training set and 10,138 patients in the validating set. In the training set, the average age of the patients was  $68.4 \pm 11.6$  years, and the median OS was 11.2 months. In the validating set, the average age of the patients was  $68.8 \pm 11.7$  years, and the median OS was 11.2 months.

Table 1

The basic characteristics of patients with pancreatic ductal adenocarcinoma of the head of the pancreas.

<b>Variables</b>	<b>Overall</b>	<b>Training set</b>	<b>Validating set</b>	<b>p</b>
	<b>N(%)</b>	<b>N(%)</b>	<b>N(%)</b>	
patients	33893	23725(70.0%)	10168(30.0%)	
Age at diagnosis				
mean ± SD(year)	68.7 ± 11.7	68.8 ± 11.7	68.4 ± 11.6	0.017
Sex				
Female	16797 (49.6%)	11821 (49.8%)	4976 (48.9%)	0.137
Male	17096 (50.4%)	11904 (50.2%)	5192 (51.1%)	
Race				
Black	4126 (12.2%)	2862 (12.1%)	1264 (12.4%)	0.601
Other	2294 (6.77%)	1615 (6.81%)	679 (6.68%)	
White	27473 (81.1%)	19248 (81.1%)	8225 (80.9%)	
Marital status				
Married	18836 (55.6%)	13196 (55.6%)	5640 (55.5%)	0.805
Unmarried	15057 (44.4%)	10529 (44.4%)	4528 (44.5%)	
Surgery				
No	27165 (80.1%)	19006 (80.1%)	8159 (80.2%)	0.791
Yes	6728 (19.9%)	4719 (19.9%)	2009 (19.8%)	
Radiation				
No	26562 (78.4%)	18623 (78.5%)	7939 (78.1%)	0.401
Yes	7331 (21.6%)	5102 (21.5%)	2229 (21.9%)	
Chemotherapy				
No	14856 (43.8%)	10492 (44.2%)	4364 (42.9%)	0.027
Yes	19037 (56.2%)	13233 (55.8%)	5804 (57.1%)	
AJCC				
I	2947 (8.70%)	2065 (8.70%)	882 (8.67%)	0.665
II	12848 (37.9%)	8943 (37.7%)	3905 (38.4%)	
III	3974 (11.7%)	2791 (11.8%)	1183 (11.6%)	

Variables	Overall	Training set	Validating set	p
	N(%)	N(%)	N(%)	
IV	14124 (41.7%)	9926 (41.8%)	4198 (41.3%)	
T				
T0	11 (0.03%)	4 (0.02%)	7 (0.07%)	0.998
T1	1259 (3.71%)	869 (3.66%)	390 (3.84%)	
T2	6394 (18.9%)	4463 (18.8%)	1931 (19.0%)	
T3	16651 (49.1%)	11668 (49.2%)	4983 (49.0%)	
T4	6499 (19.2%)	4558 (19.2%)	1941 (19.1%)	
TX	3079 (9.08%)	2163 (9.12%)	916 (9.01%)	
N				
N0	17966 (53.0%)	12582 (53.0%)	5384 (53.0%)	0.625
N1	12735 (37.6%)	8932 (37.6%)	3803 (37.4%)	
NX	3192 (9.42%)	2211 (9.32%)	981 (9.65%)	
M				
M0	19769 (58.3%)	13799 (58.2%)	5970 (58.7%)	0.352
M1	14124 (41.7%)	9926 (41.8%)	4198 (41.3%)	
Region				
Alaska	39 (0.12%)	30 (0.13%)	9 (0.09%)	0.748
East	13443 (39.7%)	9394 (39.6%)	4049 (39.8%)	
Northern Plains	3611 (10.7%)	2519 (10.6%)	1092 (10.7%)	
Pacific Coast	15333 (45.2%)	10739 (45.3%)	4594 (45.2%)	
Southwest	1467 (4.33%)	1043 (4.40%)	424 (4.17%)	
Origin				
Non-Spanish-Hispanic-Latino	30509 (90.0%)	21344 (90.0%)	9165 (90.1%)	0.361
Spanish-Hispanic-Latino	3384 (9.98%)	2381 (10.0%)	1003 (9.86%)	
Tumor size				

Variables	Overall	Training set	Validating set	p
	N(%)	N(%)	N(%)	
Median(1Q, 3Q)	37.0(28.0, 50.0)	37.0(28.0, 50.0)	37.0(28.0, 50.0)	-
Tumor extension				
Median(1Q, 3Q)	440.0(400.0, 600.0)	440.0(400.0, 600.0)	440.0(400.0, 600.0)	-
CS_lymph_nodes				
Median(1Q, 3Q)	0(0, 110.0)	0(0, 110.0)	0(0, 110.0)	-
CS_mets				
Median(1Q, 3Q)	0(0, 40.0)	0(0, 40.0)	0(0, 40.0)	-
Status				
Alive	9685(28.6%)	6738(28.4%)	7366(96.4%)	0.613
Dead	24208(71.4%)	16987(71.6%)	275(3.6%)	
Survival time				
Median(month)	6.0(2.0, 14.0)	6.0(2.0, 14.0)	6.0(2.0, 14.0)	-

## Predictors of Survival and Survival curve

The hazard ratio (HR) and 95% confidence interval (CI) of each factor related to the OS of patients with PDAC-HP are shown in Table 2. Using Cox regression analysis, we screened out independent risk factors that affect the OS of patients with PDAC-HP, such as age at diagnosis (HR = 1.01, 95% CI = 1.01–1.02, P < 0.001), sex (female vs male: HR = 1.07, 95% CI = 1.05–1.10, P < 0.001), marital status at diagnosis (unmarried vs married: HR = 1.95, 95% CI = 1.66–2.28, P < 0.001), race (white vs black: HR = 1.12, 95% CI = 1.09–1.15, P < 0.001), AJCC staging (II vs I: HR = 1.27, 95% CI = 1.19–1.35, P < 0.001; III vs I: HR = 1.30, 95% CI = 1.19–1.41, P < 0.001; IV vs I: HR = 1.80, 95% CI = 1.60–2.02, P < 0.001), surgery (yes vs no: HR = 0.41, 95% CI = 0.40–0.41, P < 0.001), radiotherapy (yes vs no: HR = 0.91, 95% CI = 0.88–0.95, P < 0.001), and chemotherapy (yes vs no: HR = 0.46, 95% CI = 0.44–0.47, P < 0.001). The above 8 variables are incorporated into the final prediction model through AIC criteria. Subsequently, we performed COX survival regression analysis on patients with PDAC-HP to analyze the impact of the selected predictors on the OS of patients. As shown in eFigure 2 in the Supplement, the selected predictors had an impact on the OS of patients with PDAC-HP.

Table 2

Univariable and multivariable Cox proportional hazard regression analysis of patients with pancreatic ductal adenocarcinoma of the head of the pancreas.

Variables	Univariable analysis		Multivariable analysis	
	HR(95%CI)	P	HR(95%CI)	P
Age	1.01(1.01–1.01)	< 0.001***	1.01(1.01–1.02)	< 0.001***
Sex				
Female	Reference		Reference	
Male	1.02(1.00–1.05)	0.059	1.07(1.05–1.10)	< 0.001***
Race				
Black	Reference		Reference	
Other	0.87(0.84–0.90)	< 0.001***	0.93(0.89–0.96)	< 0.001***
White	0.94(0.88–0.99)	0.031*	0.97(0.91–1.03)	0.282
Marital status				
Married	Reference		Reference	
Unmarried	1.23(1.20–1.26)	< 0.001***	1.12(1.09–1.15)	< 0.001***
Surgery				
No	Reference		Reference	
Yes	0.33(0.32–0.34)	< 0.001***	0.41(0.40–0.41)	< 0.001***
Radiation				
No	Reference		Reference	
Yes	0.54(0.53–0.56)	< 0.001***	0.91(0.88–0.95)	< 0.001***
Chemotherapy				
No	Reference		Reference	
Yes	0.46(0.44–0.47)	< 0.001***	0.46(0.44–0.47)	< 0.001***

\* P < 0.05, \*\* P < 0.01, \*\*\* P < 0.001.

Variables	Univariable analysis		Multivariable analysis	
	HR(95%CI)	P	HR(95%CI)	P
AJCC				
I	Reference		Reference	
II	0.97(0.93–1.03)	0.326	1.27(1.19–1.35)	< 0.001***
III	1.30(1.23–1.38)	< 0.001***	1.30(1.19–1.41)	< 0.001***
IV	2.39(2.27–2.51)	< 0.001***	1.80(1.60–2.02)	< 0.001***
T				
T0	Reference		Reference	
T1	0.33(0.17–0.65)	0.001**	0.63(0.32–1.22)	0.17
T2	0.54(0.28–1.04)	0.064	0.92(0.48–1.78)	0.803
T3	0.42(0.22–0.80)	0.009	0.86(0.45–1.66)	0.658
T4	0.55(0.29–1.06)	0.074	0.89(0.46–1.72)	0.734
TX	0.96(0.50–1.85)	0.900	0.93(0.48–1.79)	0.83
N				
N0	Reference		Reference	
N1	0.91(0.88–0.93)	< 0.001***	1.14(1.1–1.18)	< 0.001***
NX	1.81(1.73–1.89)	< 0.001***	1.10(0.96–1.27)	0.171
M				
M0	Reference		Reference	
M1	2.30(2.24–2.36)	< 0.001***	NA	NA
Region				

\* P < 0.05, \*\* P < 0.01, \*\*\* P < 0.001.

Variables	Univariable analysis		Multivariable analysis	
	HR(95%CI)	P	HR(95%CI)	P
Alaska	Reference		Reference	
East	0.71(0.50-1)	0.051	0.89(0.63–1.27)	0.528
Northern Plains	0.74(0.52–1.04)	0.087	0.93(0.65–1.32)	0.674
Pacific Coast	0.74(0.52–1.04)	0.085	0.86(0.6–1.22)	0.385
Southwest	(0.79 0.55–1.12)	0.184	0.91(0.64–1.3)	0.606
Origin				
Non-Spanish-Hispanic-Latino	Reference		Reference	
Spanish-Hispanic-Latino	1.07(1.02–1.11)	0.003**	0.98(0.94–1.02)	0.331
Tumor size	1.00(1.00–1.00)	<0.001***	1.00(1.00–1.00)	0.395
Tumor extension	1.00(1.00–1.00)	<0.001***	1.00(1.00–1.00)	0.412
CS_lymph_nodes	1.00(1.00–1.00)	<0.001***	1.00(1.00–1.00)	0.626
CS_mets	1.02(1.02–1.02)	<0.001***	1.00(1.00–1.00)	0.621

\* P < 0.05, \*\* P < 0.01, \*\*\* P < 0.001.

## Nomogram development

We used the training set to construct the survival prediction nomogram of patients with PDAC-HP through the selected predictors. The C index of the nomogram was 0.736, which was higher than the C index of the AJCC staging system which was 0.625. The nomogram was used to calculate the OS of patients at 1, 3 and 5 years after diagnosis (Fig. 1).

Next, we compared the AUCs when using the nomogram and the AJCC staging system to predict the OS of patients with PDAC-HP at 1, 3, and 5 years after diagnosis. The result is shown in Fig. 2. The nomogram predicted that the AUCs of the OS at 1, 3, and 5 years after diagnosis were 0.791(Fig. 2A), 0.792 (Fig. 2B), and 0.790 (Fig. 2C), respectively. For the AJCC staging system, the AUCs at 3, 5, and 10 years after diagnosis were 0.669 (Fig. 2A), 0.688 (Fig. 2B), and 0.694 (Fig. 2C), respectively.

Finally, we constructed calibration plots to verify the predictive ability of the nomogram. The calibration curves of patients with PDAC-HP at 1, 3, and 5 years showed that the 1, 3, and 5 year survival predictions by the nomogram are very close to the actual survival (Fig. 3).

### Nomogram validation

We used the validating set to verify the nomogram, and its index was 0.732, which was higher than the C index of the AJCC staging system (0.623). Then, we compared the AUCs of the both models in the validating set to predict the OS of patients at 1, 3 and 5 years after diagnosis. In the validating set, We found that the AUCs of nomogram for predicting the OS at 1, 3 and 5 years after diagnosis were 0.792(Fig. 4A), 0.792 (Fig. 4B) and 0.780 (Fig. 4C), respectively. It was higher than the AUCs in the AJCC staging system which were 0.670(Fig. 4A), 0.680 (Fig. 4B) and 0.675 (Fig. 4C), respectively (Fig. 4).

In addition, we constructed calibration plots on the nomogram for patients in the validating set 1, 3, and 5 years after diagnosis. The results show that the nomogram's 1-, 1- and 5-year survival predictions were very close to the actual survival (Fig. 5).

Finally, we further analyzed the accuracy of the nomogram to predict the OS of patients with PDAC-HP. Compared with the AJCC staging system, the nomogram has better NRI and IRI values. In the nomogram, the NRIs for 1, 3 and 5 years were 0.578 (95% CI = 0.551–0.598), 0.398 (95% CI = 0.364–0.435), and 0.386 (95% CI = 0.338–0.451), respectively; the IDIs for 1, 3 and 5 years were 0.011 ( $P = 0.003$ ), 0.014 ( $P < 0.001$ ), and 0.014 ( $P < 0.001$ ), respectively. Meanwhile, the DCA curves of the nomogram for the validating set at 1, 3, and 5 years after diagnosis are shown in Fig. 6. The results show that compared with the AJCC staging system, the nomogram has better accuracy and clinical validity in predicting the OS of patients with PDAC-HP.

## Discussion

In recent years, there have been about 18 million new cancer cases worldwide each year, and about 9.6 million patients died of cancer.<sup>[1]</sup> There are approximately 460,000 new cases of pancreatic cancer each year, accounting for 2.5% of all tumors and ranking 14th among new cancers worldwide.<sup>[1]</sup> Pancreatic cancer includes pancreatic ductal adenocarcinoma (PDAC) and other types, which was the digestive system cancer with the highest mortality rate, with a 5-year survival rate of about 8%.<sup>[19, 20]</sup> Because PDAC is prone to metastasis, various treatments such as surgery, radiotherapy and chemotherapy are little effective, which often has a serious impact on the survival of patients.<sup>[21, 22]</sup> Therefore, early detection of cancer and effective treatment can effectively improve the treatment effect of patients. At present, the TNM staging system is the most widely used tool for evaluating the prognosis of cancer patients, but due to its own limitations, it is unable to make individualized predictions for cancer patients themselves.<sup>[13, 23, 24]</sup> Therefore, the development of an individualized predictive model that integrates multiple predictive factors is of great significance for improving the treatment effect of PDAC patients and prolonging the survival period of patients.

A nomogram is a new type of prediction model that can predict the survival rate of a specific outcome.<sup>[14]</sup> It can combine a variety of predictive factors, such as demographic and tumor characteristics, and graphically display the survival rate of each patient.<sup>[25]</sup> Accumulating studies have shown that compared with AJCC staging system, nomogram has better predictive ability.<sup>[26, 27]</sup> At present, nomogram is widely used to predict the survival outcome of patients with various tumors, such as lung cancer, breast cancer, and liver cancer.<sup>[28-30]</sup>

In this study, we constructed a nomogram of the survival outcomes of patients with PDAC-HP based on 33,893 American patients with PDAC-HP in the SEER database. The results of the study showed that eight variables proved to be independent prognostic factors, including age at diagnosis, sex, race, marital status at diagnosis, AJCC staging, surgery, radiotherapy and chemotherapy. Through the AIC criteria, we finally determined these eight variables as the predictors of the final nomogram model. In the training set, the C index of the nomogram we constructed is 0.736, which was higher than the C index of the AJCC staging system (0.625). Compared with the AJCC staging system, the nomogram model predicts a higher AUCs for the OS of patients at 1, 3 and 5 years. The results of the calibration plots also showed that the nomogram predicted the expected survival rate of patients with PDAC-HP was very close to the actual survival rate. It shows that the nomogram can predict the survival outcome of patients with PDAC-HP well, and the predictive ability is better than the AJCC staging system.

In the validating set, We validated the nomogram for patients with PDAC-HP survival. The C index of the nomogram for the validating set (0.732) was similar to the C index of the training set, but was higher than that of the AJCC staging system (0.623). The results of the AUCs and calibration curve of the nomogram showed that the nomogram of the validating set could also predict the survival outcome of patients with PDAC-HP well. Then, in order to further evaluate the predictive ability and clinical significance of the nomogram, we analyzed the NRI, IDI and DCA of the nomogram. The NRI and IDI are evaluation indicators of model effectiveness.<sup>[31-33]</sup> Compared with the AJCC staging system, the NRI and IDI of the nomogram were higher at 1, 3 and 5 years after diagnosis. DCA is generally considered to be useful for verifying the benefits and clinical effectiveness of the model.<sup>[34-36]</sup> In our study, the nomogram has better DCA results than the AJCC staging system at 1, 3 and 5 years after diagnosis. This shows that compared with the AJCC staging system, the nomogram is more clinically effective and accurate in predicting the OS of patients with PDAC-HP. In short, in predicting the OS of patients with PDAC-HP, the nomogram we constructed is better than the AJCC staging system, and provides a reference for patient treatment strategies.

This study still has some limitations that should be noted. First of all, the study is a retrospective study based on the SEER database. Some factors that may affect the OS of patients are not included in the nomogram, such as religious beliefs, education level, lymphovascular invasion, drug treatments. Second, the retrospective research has its own limitations, such as selection and information bias in the selection process of the research set. In addition, the nomogram only includes some predictors, and there may be some deviations when doctors predict the OS of patients. Finally, the data of this study only included the

PDAC-HP population in some parts of the United States, and it was concluded that more large independent sets should be added for verification when it is promoted.

## Conclusion

In summary, we screened and identified eight predictors related to the OS of patients with PDAC-HP, such as age at diagnosis, sex, race, marital status at diagnosis, AJCC staging, surgery, radiotherapy, and chemotherapy. We established a nomogram of patients with PDAC-HP based on a large research set to predict the OS of patients at 1, 3, and 5 years after diagnosis. The performance of the nomogram we constructed is better than that of the AJCC staging system, and can predict the OS of patients with PDAC-HP well.

## Declarations

### Ethical approval:

The data of this study comes from the SEER database. The SEER database is a tumor-related database developed by the National Cancer Institute of the United States, providing research data for researchers free of charge. All patients participating in the study received the ethical approval sought by the National Cancer Institute. The informed consent was obtained from all patients or, if patients are under 18, from a parent and/or legal guardian.

### Consent for publication:

Consent for publication was obtained from all participants.

### Availability of data and materials:

We obtained permission to access the database after signing and submitting the SEER Research Data Agreement form via email. The data that support the findings of this study are available from SEER database but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of SEER database.

### Data sharing:

The datasets generated and analyzed during the current study are available in the SEER database repository (<https://seer.cancer.gov>) .

### Conflicts of Interest:

None

### Funding:

None.

### Author Contributions:

All authors had full access to all of the data in the study.

Doc. Zhilong Liu, Haohui Yu and Ke Cai take responsibility for the integrity of the data and the accuracy of the data analysis.

*Concept and design:* Ke Cai, Zhilong Liu.

*Acquisition, analysis, or interpretation of data:* Zhilong Liu, Haohui Yu, Yulin Huang, Bin Feng, Liubo Xing.

*Drafting of the manuscript:* Zhilong Liu, Haohui Yu, Mingrong Cao, Ke Cai.

*Critical revision of the manuscript for important intellectual content:* All authors.

*Statistical analysis:* Zhilong Liu, Haohui Yu, Yulin Huang, Bin Feng.

*Administrative, technical, or material support:* Ke Cai, Mingrong Cao, Jiebing Li, Liubo Xing.

*Supervision:* Ke Cai.

All authors contributed to writing of the manuscript and approved the final version.

#### **Acknowledgements:**

For the completion of my research, first, I wish to express my deepest gratitude to Prof. Cai, who has given me the most valuable suggestions and advices, and made necessary corrections. Then I am greatly indebted to Doc. Yu, Doc. Cao, and Doc. Huang for his advice on the format and help with my computerization. Finally, I would like to express my thanks to my colleagues Doc. Li, Doc. Feng, and Mr Xing, who have generously offered their help with my study.

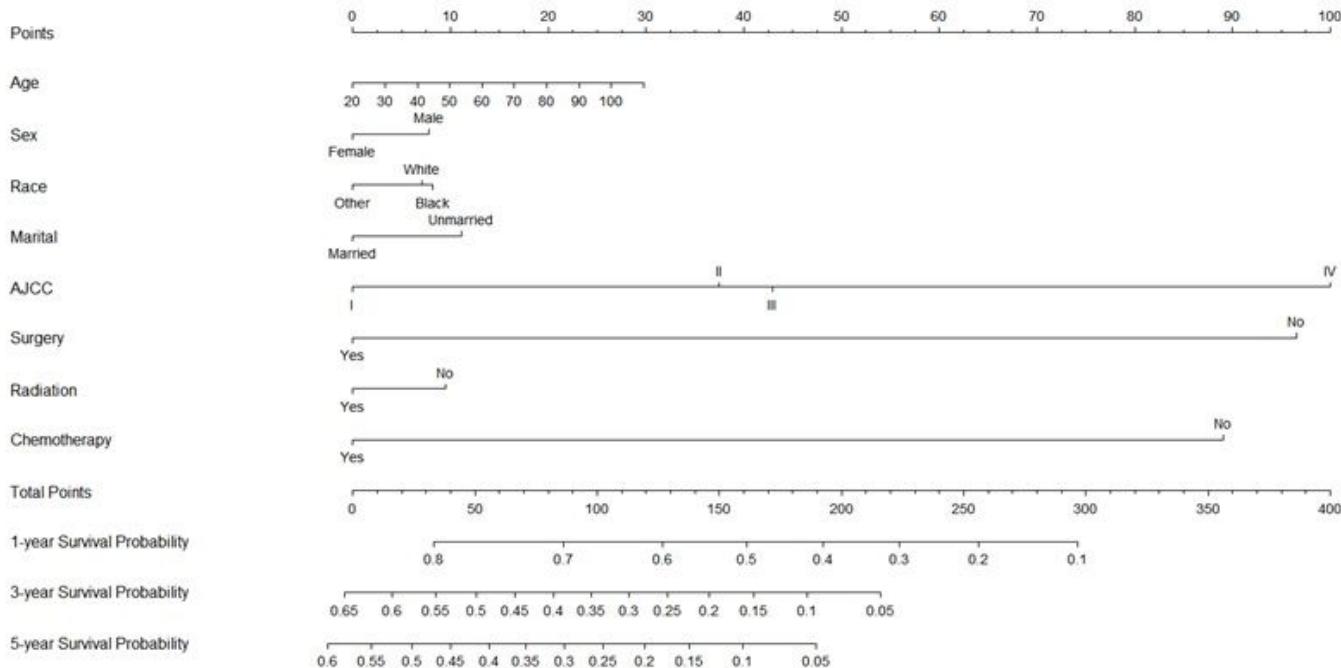
## **References**

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.
2. World Health Organization. Global Health Observatory. Geneva: World Health Organization; 2018. [who.int/gho/database/en/](http://who.int/gho/database/en/).
3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin. 2018;68(1):7–30.
4. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69(1):7–34.
5. McGuigan A, Kelly P, Turkington RC, et al. Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes. World J Gastroenterol. 2018;24(43):4846–4861.
6. Puleo F, Nicolle R, Blum Y, et al. Stratification of Pancreatic Ductal Adenocarcinomas Based on Tumor and Microenvironment Features. Gastroenterology. 2018;155(6):1999–2013.

7. Adamska A, Domenichini A, Falasca M. Pancreatic Ductal Adenocarcinoma: Current and Evolving Therapies. *Int J Mol Sci.* 2017;18(7):1338.
8. Luchini C, Capelli P, Scarpa A. Pancreatic Ductal Adenocarcinoma and Its Variants. *Surg Pathol Clin.* 2016;9(4):547–560.
9. Ayres Pereira M, Chio IIC. Metastasis in Pancreatic Ductal Adenocarcinoma: Current Standing and Methodologies. *Genes (Basel).* 2019;11(1):6.
10. Feng Q, Li C, Zhang S, et al. Recurrence and survival after surgery for pancreatic cancer with or without acute pancreatitis. *World J Gastroenterol.* 2019;25(39):6006–6015.
11. Lof S, Korrel M, van Hilst J, et al. Impact of Neoadjuvant Therapy in Resected Pancreatic Ductal Adenocarcinoma of the Pancreatic Body or Tail on Surgical and Oncological Outcome: A Propensity-Score Matched Multicenter Study. *Ann Surg Oncol.* 2020 Jun;27(6):1986–1996.
12. Guillén-Ponce C, Blázquez J, González I, et al. Diagnosis and staging of pancreatic ductal adenocarcinoma. *Clin Transl Oncol.* 2017;19(10):1205–1216.
13. Balachandran VP, Gonen M, Smith JJ, et al. Nomograms in oncology: more than meets the eye. *Lancet Oncol.* 2015;16(4):e173-e180.
14. David A Grimes. The nomogram epidemic: resurgence of a medical relic. *Ann Intern Med.* 2008;149(4):273–275.
15. Hwang WL, Tendulkar RD, Niemierko A, et al. Comparison Between Adjuvant and Early-Salvage Postprostatectomy Radiotherapy for Prostate Cancer With Adverse Pathological Features. *JAMA Oncol.* 2018;4(5):e175230.
16. Tang XR, Li YQ, Liang SB, et al. Development and validation of a gene expression-based signature to predict distant metastasis in locoregionally advanced nasopharyngeal carcinoma: a retrospective, multicentre, cohort study. *Lancet Oncol.* 2018;19(3):382–393.
17. Wang S, Yang L, Ci B, et al. Development and Validation of a Nomogram Prognostic Model for SCLC Patients. *J Thorac Oncol.* 2018;13(9):1338–1348.
18. Yu H, Cai K, Huang Y, et al. Risk factors associated with suicide among leukemia patients: A Surveillance, Epidemiology, and End Results analysis. *Cancer Med.* 2020;9(23):9006–9017.
19. Mizrahi JD, Surana R, Valle JW, et al. Pancreatic cancer. *Lancet.* 2020;395(10242):2008–2020.
20. Heller DR, Nicolson NG, Ahuja N, et al. Association of Treatment Inequity and Ancestry With Pancreatic Ductal Adenocarcinoma Survival. *JAMA Surg.* 2020;155(2):e195047.
21. Christenson ES, Jaffee E, Azad NS. Current and emerging therapies for patients with advanced pancreatic ductal adenocarcinoma: a bright future. *Lancet Oncol.* 2020;21(3):e135-e145.
22. Grossberg AJ, Chu LC, Deig CR, et al. Multidisciplinary standards of care and recent progress in pancreatic ductal adenocarcinoma. *CA Cancer J Clin.* 2020;70(5):375–403.
23. Van Roessel S, Kasumova GG, Verheij J, et al. International Validation of the Eighth Edition of the American Joint Committee on Cancer (AJCC) TNM Staging System in Patients With Resected Pancreatic Cancer. *JAMA Surg.* 2018;153(12):e183617.

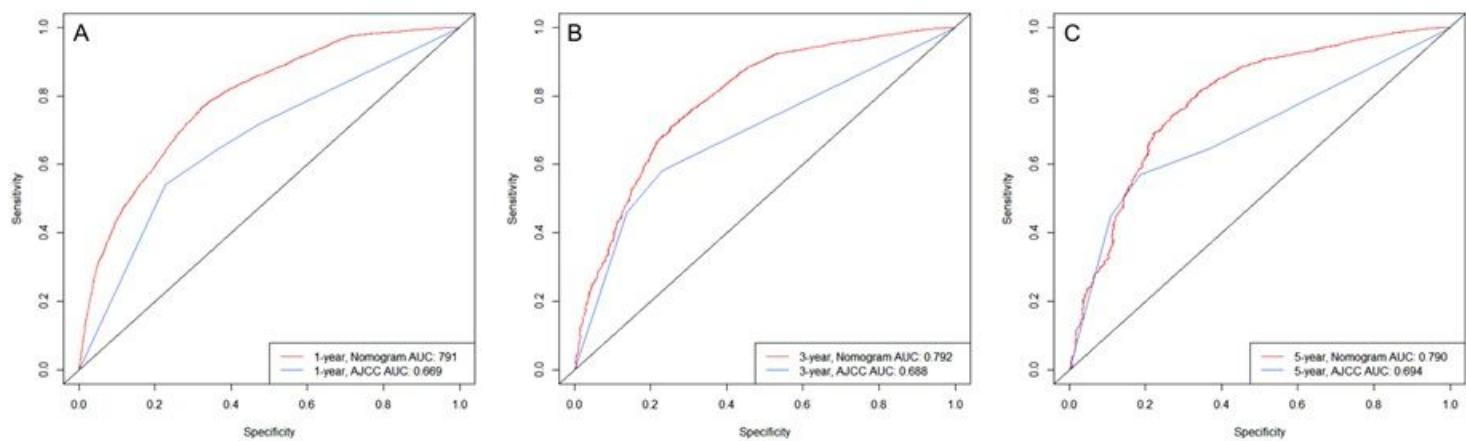
24. Amin MB, Greene FL, Edge SB, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin.* 2017;67(2):93–99.
25. Iasonos A, Schrag D, Raj GV, et al. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol.* 2008;26(8):1364–1370.
26. He Y, Zhu Z, Chen Y, et al. Development and Validation of a Novel Diagnostic Nomogram to Differentiate Between Intestinal Tuberculosis and Crohn's Disease: A 6-year Prospective Multicenter Study. *Am J Gastroenterol.* 2019;114(3):490–499.
27. Callegaro D, Miceli R, Bonvalot S, et al. Development and external validation of two nomograms to predict overall survival and occurrence of distant metastases in adults after surgical resection of localised soft-tissue sarcomas of the extremities: a retrospective analysis. *Lancet Oncol.* 2016;17(5):671–680.
28. Liang W, Zhang L, Jiang G, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. *J Clin Oncol.* 2015;33(8):861–869.
29. Wang X, Feng Z, Huang Y, et al. A Nomogram To Predict The Overall Survival Of Breast Cancer Patients And Guide The Postoperative Adjuvant Chemotherapy In China. *Cancer Manag Res.* 2019;11:10029–10039.
30. Du Q, Wang Y, Guan S, et al. The diagnostic nomogram of platelet-based score models for hepatic alveolar echinococcosis and atypical liver cancer. *Sci Rep.* 2019;9(1):19403.
31. Leening MJ, Vedder MM, Witteman JC, et al. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. *Ann Intern Med.* 2014;160(2):122–131.
32. McKearnan SB, Wolfson J, Vock DM, et al. Performance of the Net Reclassification Improvement for Nonnested Models and a Novel Percentile-Based Alternative. *Am J Epidemiol.* 2018;187(6):1327–1335.
33. Hayashi K, Eguchi S. The power-integrated discriminant improvement: An accurate measure of the incremental predictive value of additional biomarkers. *Stat Med.* 2019;38(14):2589–2604.
34. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* 2006;26(6):565–574.
35. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res.* 2019;3:18.
36. Van Calster B, Wynants L, Verbeek JFM, et al. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. *Eur Urol.* 2018;74(6):796–804.

## Figures



**Figure 1**

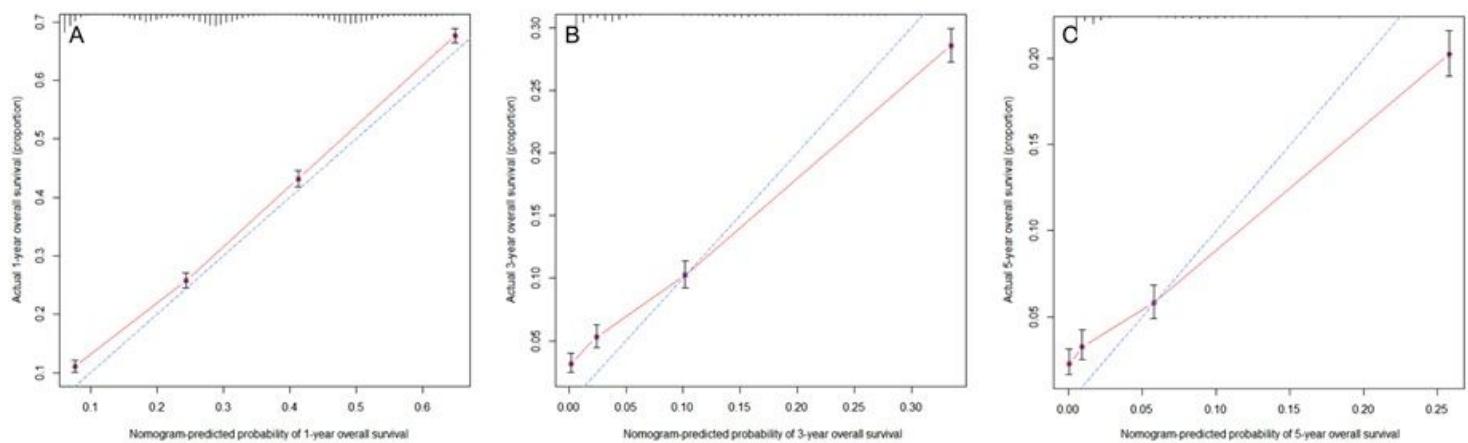
Nomogram predicting the 1-year, 3-year and 5-year overall survival of PDAC-HP. Using the nomogram, we determine the position of each variable of the patient on the corresponding axis, and calculate the corresponding score on the points axis. Then calculate the total score of all variables of the patient. Finally, draw a straight line on the score corresponding to the total points axis to determine the patient's survival rate at 1-year, 3-year and 5-year.



**Figure 2**

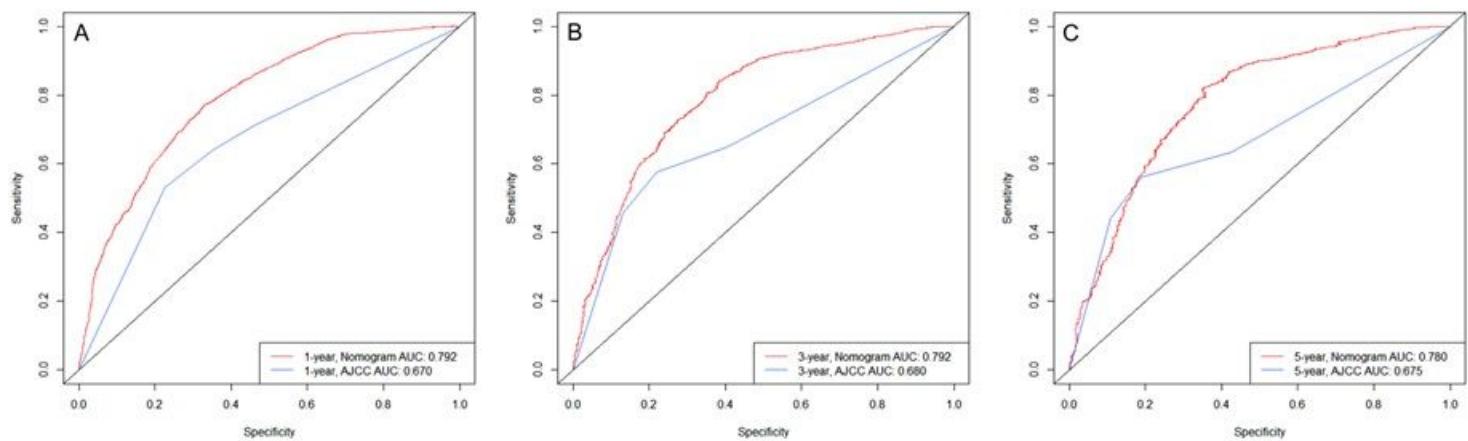
Comparison of the AUC of nomogram and AJCC staging system in the training set. The AUC is used to predict the overall survival of PDAC-HP at 1-year (A), 3-year (B) and 5-year (C). The red line represents the

overall survival predicted by the nomogram, and the blue line represents the overall survival predicted by AJCC staging system.



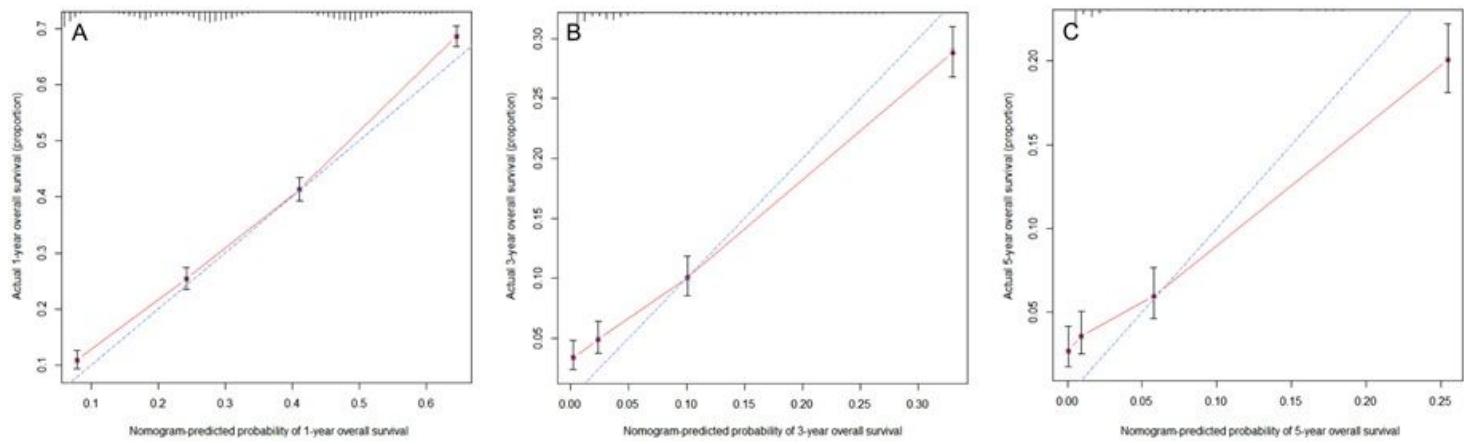
**Figure 3**

The calibration plots of the nomogram using the training set. The x-axis represents the overall survival predicted by the nomogram, and the y-axis represents the actual survival. (A) 1-year overall survival. (B) 3-year overall survival. (C) 5-year overall survival.



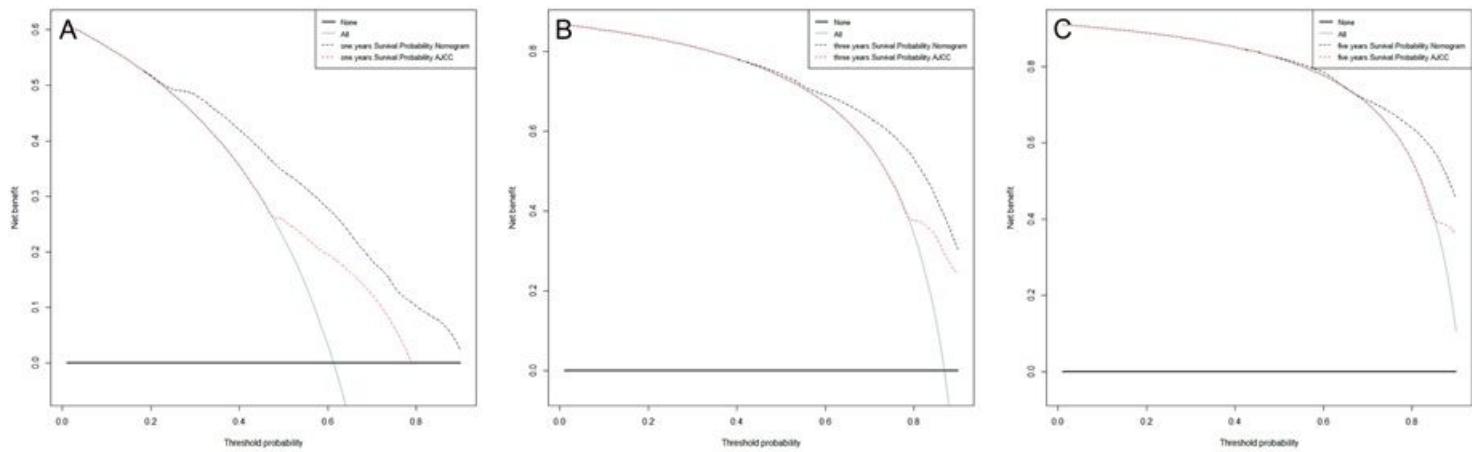
**Figure 4**

Comparison of the AUC of nomogram and AJCC staging system in the validating set. The AUC is used to predict the overall survival of PDAC-HP at 1-year (A), 3-year (B) and 5-year (C). The red line represents the overall survival predicted by the nomogram, and the blue line represents the overall survival predicted by AJCC staging system.



**Figure 5**

The calibration plots of the nomogram using the validating set. The x-axis represents the overall survival predicted by the nomogram, and the y-axis represents the actual survival. (A) 1-year overall survival. (B) 3-year overall survival. (C) 5-year overall survival.



**Figure 6**

The decision curve analysis of the nomogram and the AJCC staging system in the validating set. The decision curve analysis for predicting the prognosis of PDAC-HP at 1-year (A), 3-year (B) and 5-year (C). The black dotted line represents the nomogram, and the red dotted line represents the AJCC staging system.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplement.docx](#)