Nomograms for predicting overall survival and cancer-specific survival in young patients with pancreatic cancer in the US based on the SEER database

Min Shi [Equal first author, 1], **Biao Zhou** [Equal first author, 1], **Shu-Ping Yang** [Corresp. 1]

[1] Department of Gastroenterology, Liyang Branch of Jiangsu Province Hospital, Liyang, China

Corresponding Author: Shu-Ping Yang
Email address: shupingy@163.com

Background: The incidence of young patients with pancreatic cancer (PC) is on the rise, and there is a lack of models that could effectively predict their prognosis. The purpose of this study was to construct nomograms for predicting the overall survival (OS) and cancer-specific survival (CSS) of young patients with PC.

Methods: PC patients younger than 50 years old from 2004 to 2015 in the Surveillance, Epidemiology, and End Results (SEER) database were selected and randomly divided into training set and validation set. Univariate and forward stepwise multivariate Cox analysis was used to determine the independent factors affecting OS. Fine and Gray competing risk regression model was used to determine the independent factors affecting CSS. We used significant variables in the training set to construct nomograms predicting prognosis. The discrimination and calibration power of models were evaluated by concordance index (C-index), calibration curve and 10-flod cross-validation.

Results: 4,146 patients were selected. Multivariate Cox analysis showed that gender, race, grade, pathological types, AJCC stage and surgery were independent factors affecting OS. The C-index of the nomogram predicting OS in training and validation was 0.733 (average=0.731,95%CI=0.724-0.738) and 0.742 (95%CI=0.725-0.759), respectively. Competing risk analysis showed that primary site, pathological types, AJCC stage and surgery were independent factors affecting CSS. The C-index of the nomogram predicting CSS in training and validation set was 0.792 (average=0.765, 95%CI=0.742-0.788) and 0.776 (95%CI=0.773-0.779), respectively. C-index based on nomogram was better in training and validation set than that based on AJCC stage. Calibration curves showed that these nomograms could accurately predict the 1-, 3- and 5-year OS and CSS both in training set and validation set.

Conclusions: The nomograms could effectively predict OS and CSS in young patients with PC, which help clinicians more accurately and quantitatively judge the prognosis of individual patients.

Nomograms for predicting overall survival and cancer-specific survival in young patients with pancreatic cancer in US based on the SEER database

Min Shi[1, #], Biao Zhou[1, #], Shu-ping Yang[1]

Authors' affiliations:

[1]: Department of Gastroenterology, Liyang Branch of Jiangsu Province Hospital, Jianshe west Road No. 70, Liyang, China

[#]: These authors contributed equally to this work and should be considered co-first authors.

Corresponding author: Shu-ping Yang, E-male: shupingy@163.com, phone number: +86-13813900689

Conflicts of interest: The authors declare they have no conflict of interest

Abstract

Background: The incidence of young patients with pancreatic cancer (PC) is on the rise, and there is a lack of models that could effectively predict their prognosis. The purpose of this study was to construct nomograms for predicting the overall survival (OS) and cancer-specific survival (CSS) of young patients with PC.

Methods: PC patients younger than 50 years old from 2004 to 2015 in the Surveillance, Epidemiology, and End Results (SEER) database were selected and randomly divided into training set and validation set. Univariable and forward stepwise multivariable Cox analysis was used to determine the independent factors affecting OS. Fine and Gray competing risk regression model was used to determine the independent factors affecting CSS. We used significant

variables in the training set to construct nomograms predicting prognosis. The discrimination and calibration power of models were evaluated by concordance index (C-index), calibration curve and 10-flod cross-validation.

Results: 4,146 patients were selected. Multivariable Cox analysis showed that gender, race, grade, pathological types, AJCC stage and surgery were independent factors affecting OS. The C-index of the nomogram predicting OS in training and validation was 0.733 (average=0.731,95%CI=0.724-0.738) and 0.742 (95%CI=0.725-0.759), respectively. Competing risk analysis showed that primary site, pathological types, AJCC stage and surgery were independent factors affecting CSS. The C-index of the nomogram predicting CSS in training and validation set was 0.792 (average=0.765, 95%CI=0.742-0.788) and 0.776 (95%CI=0.773-0.779), respectively. C-index based on nomogram was better in training and validation set than that based on AJCC stage. Calibration curves showed that these nomograms could accurately predict the 1-, 3- and 5-year OS and CSS both in training set and validation set.

Conclusions: The nomograms could effectively predict OS and CSS in young patients with PC, which help clinicians more accurately and quantitatively judge the prognosis of individual patients.

Keywords: younger pancreatic cancer, SEER, nomograms, prognosis

Introduction

PC is a highly malignant tumor. Global cancer data show that PC ranks 14th and 7th among all cancers in terms of incidence and mortality, respectively.[1] It is estimated that by 2030, the

mortality rate of PC will be the second highest in the world.[2] According to the American Cancer Society, in 2019, the United States will have 56,770 new cases of PC and 45,750 will die of PC. The 5-year survival rate of PC is the lowest of all cancers, with only 9%.[3]

The incidence of PC in young patients is on the rise. A survey based on the SEER database found that the incidence of PC among young whites and blacks increased by 57% between 2001 and 2015.[4] In addition, another epidemiological study found that from 1992 to 2013, the annual percentage change in the incidence of PC in women aged 25-34 was more than 2.5%.[5] The median age of diagnosis of PC is 71 years old.[6] Due to ethnic and regional differences, the definition of " young patients " with PC has not been unified. Previous studies have generally classified young PC at the age of 45[7-9] or 50[10-12]. In this study, we screened PC patients younger than 50 years old from the SEER database in order to enable more people to benefit from our study.

American Joint Committee on Cancer (AJCC) stage has always provided a reference for judging the progress of the disease and for the choice of clinical decision-making. However, the researchers found that the staging system is insufficient .[13-15] It only incorporates some features of the tumors into the staging system, and the prognostic factors are far more than these. Previous studies have found that age, sex, grade, different treatment and even marital status are important independent factors affecting the prognosis of patients with PC.[16-19]

A nomogram is a graphical calculation or algorithm that combines several continuous variables to predict a specific end point using traditional statistical methods (for example, Logistic or Cox

61 regression model).[20] It has been constructed and verified in a variety of cancers to help
62 clinicians quickly and accurately judge the prognosis of patients.[17, 21, 22] Many studies[7-12,
63 23] have confirmed that the clinical characteristics and prognosis of young patients with PC are
64 different from older patients, but the nomogram of PC in previous studies[17, 24, 25] cannot
65 accurately predict the prognosis of young patients with PC.
66 The aim of this study was to establish and validate nomograms to predict 1-, 3-, and 5-year OS
67 and CSS based on data from younger PC patients in the SEER database between 2004 and 2015.
68 Patients and methods
69 Selection of patients
70 PC patients younger than 50 years old who were diagnosed from January 1, 2004 to December 31,
71 2015 were selected from the Incidence-SEER 18 Registries Custom Data (with additional
72 treatment fields), released April 2019, based on the November 2018 submission. The patient meted
73 the following conditions: (1) site recode ICD-0-3/WHO 2008:"pancreas"; (2) positive histology;
74 (3) active follow-up; (4) one primary only. The exclusion criteria were as follows: (1) age, race,
75 surgery unknown; (2) AJCC stage NOS, N/A or unknown; (3) age at diagnosed $\geqslant$50 years old;
76 (4) survival time unknown and less than 1 month. The detailed flow chart was shown in Figure 1.
77 For this study, we signed the SEER research data agreement to access SEER information with the
78 username10067-Nov2018. The SEER database is publicly available and the data for all patients
79 are de-identified, so the approval and informed consent of the institutional review committee were
80 not required in the current study.

Variable Classification

clinical variables including race, gender, primary site, grade, pathological types, AJCC stage, status of surgery, survival time, status of survival and cause for death were extracted from the SEER database. The degree of differentiation of tumors was divided into three groups: grade I (well differentiated) and grade II (moderately differentiated) were high differentiated, grade III (poorly differentiated) and grade IV (undifferentiated) were low differentiated, and unknown. The pathological types were divided into ductal adenocarcinoma and non-ductal adenocarcinoma. The International Classification of Diseases for Oncology, Third Edition (ICD-0-3) codes of ductal adenocarcinoma were 8500 and 8140.[26] The staging of cancer is based on the 6th edition of AJCC stage, which adapted to patients in the SEER database with a diagnosis time of 2004-2015.

Statistical Analysis

In this study, the whole cohort was randomly divided into two groups, 2,904 (70%) were training set and 1,242 (30%) were validation set. Chi-square test was used to compare the clinicopathological characteristics between the training set and the validation set.

OS referred to the duration from diagnosis to any original death or last follow-up. variables associated with OS in univariable analysis ($p < 0.05$) were included into multivariable analysis. The method of forward stepwise selection in a multivariable regression model was applied to the training cohort to select variables. The independent prognostic factors on multivariable analysis were used to construct nomogram for OS.

Considering that death from other causes is a competitive event of pancreatic cancer death, we constructed a competing risk nomogram to predict CCS. CSS referred to the duration from diagnosis to death from PC, patients who were alive at the point of last follow-up were considered as censored events. Variables related with CSS ($p < 0.05$) in the univariable analysis or with important clinical value were included into a multivariable analysis based on proportional subdistribution hazard models and those independent variables were selected to build a nomogram for CSS.

Discrimination of the nomograms was measured through the concordance index (C-index) with its respective 95% confidence interval (CI), which quantifies the level of concordance between probabilities of prediction and the actual chance of having the event of interest.[27] The larger the C-index is, the more accurate the nomogram is to predict the prognosis.[28]10-fold cross-validation was applied to verify the stability of the model and calculate the average value of C index in the training group. In order to reduce the overfit bias, calibration was evaluated by comparing the actual probabilities and the plot of the nomogram using 1000 bootstrap samples. In the calibration curve, the vertical axis is the actual value and the horizontal axis is the predicted value. If the actual / predicted value passes through the origin along the 45 °line, it shows that the nomogram has been well calibrated.[27]

For all analyses, p-value <0.05 was considered statistically significant. All date was obtained through SEER*Stat software version 8.3.5. Fine and Gray competing risk Statistical analyses

120 were performed using R work, and the others performed by SPSS (IBM, NY). The nomograms
121 and calibration curves were draw using R version 3.6.0 (http://www.r-project.org)
122 Results
123 Patients characteristic
124 Finally, there were a total of 4,146 cases of young patients with PC diagnosed in the SEER
125 database between January 1, 2004 and December 31, 2015. In the entire cohort, in terms of
126 demography, mainly whites (74.1%) and males (53.7%); In terms of tumor characteristics,
127 pancreatic head cancer (46.1%) was the most common, more highly differentiated tumors
128 (30.7%), mainly ductal adenocarcinoma (68.0%), in addition, more advanced tumors in diagnosis
129 (55.7%); In terms of treatment, the proportion of patients undergoing surgery was low (37.5%).
130 Detailed patient clinical characteristics were summarized in Table 1.
131 Construction and validation of nomogram for predicting OS of young patients with PC
132 The results of univariable and multivariable Cox regression models for OS were shown in Table
133 2. Univariable analysis showed that race, gender, primary site, grade, pathological types, AJCC
134 stage and surgery were correlated with OS ($p<0.05$). Multivariable analysis suggested that
135 gender, grade, pathological types, AJCC stage and surgery were independent risk factors for OS.
136 In details, female (HR:0.91, 95%CI=0.83-0.98; p=0.019), non-ductal adenocarcinoma
137 (HR:0.40,95%CI=0.36-0.45; $p<0.001$) and receiving surgery (HR:0.43,95%CI=0.38-0.50;
138 $p<0.001$) had a better OS. Black (HR:1.15, 95%CI=1.03-1.28; p=0.016), low differentiated
139 tumor (HR:1.59, 95%CI=1.40-1.80; $p<0.001$) and advanced stage (HR:4.63, 95%CI=3.53-6.07;

p<0.001) had a worse OS. The Nomogram was constructed of the above six variables in training set (Figure 2.A). It can be seen from the nomogram that AJCC stage, which range of risk score is from 0 to 100, had the greatest contribution on OS (including 1- ,3- and 5-years), followed by pathological subtypes, surgery, grade, race and gender. The detailed steps for the application of the nomogram are as follows: Draw a vertical line to the horizontal axis marked "points" at the top of the nomogram according to the classification (e.g., gender is divided into male and female) of each prognostic variable (gender, grade, pathological type, AJCC stage, and surgery). At the position where the vertical line passes through the "Points" axis, each prognostic variable can obtain a score. Add the scores of the five variables to get the total score, find the position of the total score on the horizontal axis marked as "total points", and draw a vertical line from the total score position marked on the horizontal axis of "Total Points" to the 1 -, 3-and 5-years OS axis. Where the vertical line intersects the 1-year OS axis is the % probability of the 1-year overall survival.

Compared with C-index based on AJCC stage in training set (0.677, 95%CI=0.666-0.688) and validation set (0.672, 95%CI=0.656-0.688), The C-index of our model in the training set and validation set was 0.733 (average=0.731,95%CI=0.724-0.738) and 0.742 (95%CI=0.725-0.759) respectively, showing a better degree of discrimination. The calibration curve of training set and validation set showed good consistency between prediction and observation in the probability of 1 -, 3 -, and 5-year OS, respectively (Figure 3).

Construction and validation of nomogram for predicting CSS of young patients with PC

160 The results of univariable and multivariable competing risks models for CSS were shown in
161 Table 3. Fine and Gray analysis showed that primary site, grade, pathological types, AJCC stage
162 and surgery were correlated with CSS. Multivariable analysis suggested that primary site,
163 pathological types, AJCC stage and surgery were independent risk factors for CSS. In details,
164 non-ductal adenocarcinoma (SHR:0.57, 95%CI=0.50-0.65; p<0.001), receiving surgery
165 (SHR:0.77, 95%CI=0.68-0.87; p<0.001) and tumors in others site (SHR:0.87, 95%CI=0.77-0.99;
166 p=0.037) had a better CSS. Advanced stage (SHR:1.53, 95%CI=1.22-1.93; p<0.001) had a worse
167 CSS. The Nomogram was constructed of the above four variables in training set (Figure2.B). It
168 also can be seen from the nomogram that AJCC stage had the greatest effect on CSS (including
169 1- ,3- and 5-years), followed by pathological subtypes, surgery and primary site. Finally, similar
170 to the above, we could also predict 1 -, 3 -, and 5-year CSS in patients with PC.
171 Compared with C-index based on AJCC stage in training set (0.706, 95%, CI=0.704-0.707) and
172 validation set (0.692, 95%CI=0.695-0.689), The C-index of our model in the training set and
173 validation set was 0.792 (average=0.765, 95%CI=0.742-0.788) and 0.776 (95%CI=0.773-0.779),
174 respectively, showing a better degree of discrimination. The calibration curve of training set and
175 validation set also showed good consistency between prediction and observation in the
176 probability of 1 -, 3 -, and 5-year CSS, respectively (Figure 4).
177 Discussion
178 In this study, we determined that race, gender, grade, pathological types, surgery and AJCC stage
179 were independent factors affecting the OS of younger patients with PC by univariable and

multivariable regression analysis based on the SEER database, and competing risk analysis was used to determine that surgery, pathological types and AJCC stage and primary site were independent factors affecting CSS in younger patients with PC. We integrated the factors and draw nomograms that could effectively predict 1 -, 3 -, and 5-year OS and CSS in younger patients with PC.

As far as we know, this was the first time that a nomogram based on a large multicenter dataset has been constructed to effectively predict the prognosis of younger patients with PC. Previous studies[17, 24, 25, 29] found that age was an independent factor affecting the OS or CSS of patients with PC, and most of them divided age into < 60 years (or 65 years) and ≥ 60 years (or 65 years). There were differences in clinicopathological characteristics and prognosis between young patients with PC and elderly patients.[7-12, 23] Our data showed that primary site was an independent factor affecting CSS in young patients with PC, which was consistent with previous study.[24] Another study[29] used five variables, including age, differentiation, TNM staging, surgery and lymph node surgery, to construct a nomogram for predicting the OS rate of patients with PC, of which age was divided into 25-39, 40-59, 60-79 and 80+. But the study still had some limitations. They included only patients with non-metastasis and ductal adenocarcinoma. Our study found that in young patients with PC, patients with distant metastasis and non-ductal adenocarcinoma accounted for about 55.7% and 32.0% of the total population, respectively. Therefore, it was necessary to establish a nomogram that could predict the prognosis of young patients with PC.

In our study, AJCC stage and surgery were independent factors affecting OS and CSS in young patients with PC, which was consistent with the results of previous studies[17, 24, 29] in patients with PC. AJCC stage had the greatest influence on OS and CSS in young patients with PC, followed by pathological type, surgery and grade. Additionally, it was worth noting that the impact of pathological types on prognosis was even greater than that of surgery. Mohamed E. Mostafa et al[30] pointed out that although ductal adenocarcinoma was the most common pathological type of PC, there were many other pathological types of PC in the real world, such as solid pseudopapillary tumors, neuroendocrine tumors, and so on. They had different clinicopathological characteristics and biological behavior. Therefore, more research was needed to focus on non-ductal adenocarcinoma. There were racial differences in the OS of young patients with PC, and the risk of death was higher in blacks. This might be due to the higher incidence of distant pancreatic cancer in black patients.[4]

Compared with the traditional AJCC stage, all the C-index of our model were more than 0.72, indicating a better discrimination and the ability to provide individualized prediction for patients. For example, two patients with stage II PC after surgery, one was a male with low differentiated ductal adenocarcinoma, and the other was a female with high differentiated non-ductal adenocarcinoma. According the Table S1, the 1-year OS of the two patients could be calculated to be about 56% and 88%, respectively. However, according to the AJCC stage, both of them were patients with stage II PC, and it was difficult to compare the differences in prognosis between them.

Our research had some advantages. First of all, our study was based on the SEER database, which covered 28% of the population of the United States, so the nomograms were more applicable. Secondly, compared with the previous nomograms used to evaluate the prognosis of PC patients, our models were more targeted to evaluate the prognosis of PC patients younger than 50 years old. Finally, the calibration curves of our training set and validation set showed good consistency, indicating that our nomograms had good calibration power.

Even so, some limitations still existed in our study. First, at present, surgery was still the main treatment for patients with PC, but studies have found that radiotherapy and chemotherapy can also effectively improve the prognosis of postoperative patients with PC.[31, 32] Detailed radiotherapy or chemotherapy regimens were not available from the SEER database. In addition, other known risk factors for prognosis, such as family history[33], tobacco or alcohol[34] were also hard to obtain from the SEER database. Moreover, in this study, patients were randomly divided into two groups, 70% of them were used to construct and other 30% were used to validate the nomograms. Although both the C index and the calibration curve performed well, external validation was still needed in other populations to evaluate the accuracy of our models.

Conclusion

We developed and validated nomograms that could effectively predict the prognosis of PC patients younger than 50 years old. These nomograms could help clinicians more accurately and conveniently predict the 1-, 3-and 5-year OS and CSS of individual patients.

Abbreviation:

PC: pancreatic cancer

OS: overall survival

CSS: cancer-specific survival

SEER: Surveillance, Epidemiology, and End Results

C-index: concordance index

AJCC: American Joint Committee on Cancer

Figure legends

Figure 1. Flowchart of patient selection.

Figure1 contains detailed selection of PC patients in 2004-2015 from SEER database.

Figure 2. Nomogram for predicting OS and CSS of young patients with PC.

Figure 2 contains: A. Nomogram for predicting 1-, 3- and 5-year OS for young patients with PC; B. Nomogram for predicting 1-, 3- and 5-year CSS for young patients with PC

Figure 3. Calibration curves for 1-, 3- and 5-year OS in training set and validation set.

Figure 3 contains: A. Calibration curves for 1-year OS in training set; B. Calibration curves for 3-year OS in training set; C. Calibration curves for 5-year OS in training set; D. Calibration curves for 1-year OS in validation set; E. Calibration curves for 3-year OS in validation set; F. Calibration curves for 5-year OS in validation set.

Figure 4. Calibration curves for 1-, 3- and 5-year CSS in training set and validation set.

Figure 4 contains: A. Calibration curves for 1-year CSS in training set; B. Calibration curves for 3-year CSS in training set; C. Calibration curves for 5-year CSS in training set; D. Calibration

curves for 1-year CSS in validation set; E. Calibration curves for 3-year CSS in validation set; F. Calibration curves for 5-year CSS in validation set.

Acknowledgments

The authors acknowledge the efforts of the Surveillance, Epidemiology, and End Results Program tumor registries in the creation of the SEER database.

Reference

1. Bray, F., 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): p. 394-424.
2. Rahib, L., et al., *Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States.* Cancer Res, 2014. **74**(11): p. 2913-21.
3. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2019.* CA Cancer J Clin, 2019. **69**(1): p. 7-34.
4. Tavakkoli, A., et al., *Racial Disparities and Trends in Pancreatic Cancer Incidence and Mortality in the United States.* Clin Gastroenterol Hepatol, 2019.
5. Gordon-Dseagu, V.L., et al., *Pancreatic cancer incidence trends: evidence from the Surveillance, Epidemiology and End Results (SEER) population-based data.* Int J Epidemiol, 2018. **47**(2): p. 427-439.
6. Midha, S., S. Chawla, and P.K. Garg, *Modifiable and non-modifiable risk factors for pancreatic cancer: A review.* Cancer Lett, 2016. **381**(1): p. 269-77.
7. Bunduc, S., et al., *Very Early Onset Pancreatic Adenocarcinoma - Clinical Presentation, Risk Factors and Therapeutic Options.* Chirurgia (Bucur), 2018. **113**(3): p. 405-411.
8. Kang, J.S., et al., *Clinicopathologic and survival differences in younger patients with pancreatic ductal adenocarcinoma-A propensity score-matched comparative analysis.* Pancreatology, 2017. **17**(5): p. 827-832.
9. He, J., et al., *Young patients undergoing resection of pancreatic cancer fare better than their older counterparts.* J Gastrointest Surg, 2013. **17**(2): p. 339-44.
10. Ntala, C., et al., *Demographic, clinical, and pathological features of early onset pancreatic cancer patients.* BMC Gastroenterol, 2018. **18**(1): p. 139.
11. Ansari, D., et al., *Early-onset pancreatic cancer: a population-based study using the SEER registry.* Langenbecks Arch Surg, 2019.
12. Beeghly-Fadiel, A., et al., *Early onset pancreatic malignancies: Clinical characteristics and*

survival associations. Int J Cancer, 2016. **139**(10): p. 2169-77.
13. Adam, M.A., et al., *Rethinking the Current American Joint Committee on Cancer TNM Staging System for Medullary Thyroid Cancer.* JAMA Surg, 2017. **152**(9): p. 869-876.
14. Choi, K.H., et al., *Comparison the sixth and seventh editions of the AJCC staging system for T1 gastric cancer: a long-term follow-up study of 2124 patients.* Gastric Cancer, 2017. **20**(1): p. 43-48.
15. Shao, N., et al., *Comparison of the 7th and 8th edition of American Joint Committee on Cancer (AJCC) staging systems for breast cancer patients: a Surveillance, Epidemiology and End Results (SEER) Analysis.* Cancer Manag Res, 2019. **11**: p. 1433-1442.
16. Li, H.B., J. Zhou, and F.Q. Zhao, *A Prognostic Nomogram for Disease-Specific Survival in Patients with Pancreatic Ductal Adenocarcinoma of the Head of the Pancreas Following Pancreaticoduodenectomy.* Med Sci Monit, 2018. **24**: p. 6313-6321.
17. He, C., et al., *Overall survival and cancer-specific survival in patients with surgically resected pancreatic head adenocarcinoma: A competing risk nomogram analysis.* J Cancer, 2018. **9**(17): p. 3156-3167.
18. Wang, X.D., et al., *Marital status independently predicts pancreatic cancer survival in patients treated with surgical resection: an analysis of the SEER database.* Oncotarget, 2016. **7**(17): p. 24880-7.
19. Baine, M., et al., *Marital status and survival in pancreatic cancer patients: a SEER based analysis.* PLoS One, 2011. **6**(6): p. e21052.
20. Kattan, M.W., *Nomograms. Introduction.* Semin Urol Oncol, 2002. **20**(2): p. 79-81.
21. Kim, S.Y., et al., *Nomograms predicting survival of patients with unresectable or metastatic gastric cancer who receive combination cytotoxic chemotherapy as first-line treatment.* Gastric Cancer, 2018. **21**(3): p. 453-463.
22. Zheng, Z.F., et al., *Development and External Validation of a Simplified Nomogram Predicting Individual Survival After R0 Resection for Gastric Cancer: An International, Multicenter Study.* Ann Surg Oncol, 2018. **25**(8): p. 2383-2390.
23. Eguchi, H., et al., *Clinicopathological Characteristics of Young Patients With Pancreatic Cancer: An Analysis of Data From Pancreatic Cancer Registry of Japan Pancreas Society.* Pancreas, 2016. **45**(10): p. 1411-1417.
24. Song, W., D.L. Miao, and L. Chen, *Nomogram for predicting survival in patients with pancreatic cancer.* Onco Targets Ther, 2018. **11**: p. 539-545.
25. Li, J. and L. Liu, *Overall survival in patients over 40 years old with surgically resected pancreatic carcinoma: a SEER-based nomogram analysis.* BMC Cancer, 2019. **19**(1): p. 726.
26. Shi, S., et al., *Proposed Modification of the 8th Edition of the AJCC Staging System for Pancreatic Ductal Adenocarcinoma.* Ann Surg, 2019. **269**(5): p. 944-950.
27. Iasonos, A., et al., *How to build and interpret a nomogram for cancer prognosis.* J Clin Oncol, 2008. **26**(8): p. 1364-70.

28. Huitzil-Melendez, F.D., et al., *Advanced hepatocellular carcinoma: which staging systems best predict prognosis?* J Clin Oncol, 2010. **28**(17): p. 2889-95.
29. Pu, N., et al., *Survival prediction in pancreatic cancer patients with no distant metastasis: a large-scale population-based estimate.* Future Oncol, 2018. **14**(2): p. 165-175.
30. Mostafa, M.E., et al., *Pathologic classification of "pancreatic cancers": current concepts and challenges.* Chin Clin Oncol, 2017. **6**(6): p. 59.
31. Springfeld, C., et al., *Chemotherapy for pancreatic cancer.* Presse Med, 2019. **48**(3 Pt 2): p. e159-e174.
32. Stessin, A.M., J.E. Meyer, and D.L. Sherr, *Neoadjuvant radiation is associated with improved survival in patients with resectable pancreatic cancer: an analysis of data from the surveillance, epidemiology, and end results (SEER) registry.* Int J Radiat Oncol Biol Phys, 2008. **72**(4): p. 1128-33.
33. Schulte, A., et al., *Association between family cancer history and risk of pancreatic cancer.* Cancer Epidemiol, 2016. **45**: p. 145-150.
34. Korc, M., et al., *Tobacco and alcohol as risk factors for pancreatic cancer.* Best Pract Res Clin Gastroenterol, 2017. **31**(5): p. 529-536.

Table 1(on next page)

Clinical Characteristics of Training Set and Validation Set

1

Characteristics		Total 4,146(100%)	training set 2,904(70.0%)	validation set 1,242(30.0%)	p-value
Race					0.005
	White	3,073(74.1%)	2,192(75.5%)	881(70.9%)	
	Black	666(16.1%)	450(15.5%)	216(17.4%)	
	Others	407(9.8%)	262(9.0%)	145(11.7%)	
Gender					0.761
	Male	2,225(53.7%)	1,554(53.5%)	671(54.0%)	
	Female	1,921(46.3%)	1,350(46.5%)	571(46.0%)	
Primary site					0.205
	Head	1,911(46.1%)	1,355(46.7%)	556(44.8%)	
	Body	473(11.4%)	327(11.3%)	146(11.8%)	
	Tail	791(19.1%)	566(19.5%)	225(18.1%)	
	Others	971(23.4%)	656(22.6%)	315(25.4%)	
Grade					0.393
	High differentiated	1,272(30.7%)	909(31.3%)	363(29.2%)	
	Low differentiated	744(17.9%)	520(17.9%)	224(18.0%)	
	Unknown	2,130(51.4%)	1,475(50.8%)	655(52.7%)	
Pathological types					0.251
	Ductal adenocarcinoma	2,820(68.0%)	1,991(68.6%)	829(66.7%)	
	Non-ductal adenocarcinoma	1,326(32.0%)	913(31.4%)	413(33.3%)	
AJCC					0.833
	I	418(10.1%)	297(10.2%)	121(9.7%)	
	II	1,098(26.5%)	775(26.7%)	323(26.0%)	
	III	321(7.7%)	228(7.9%)	93(7.5%)	
	IV	2,309(55.7%)	1,604(55.2%)	705(56.8%)	
Surgery					0.256
	No	2,593(62.5%)	1,800(62.0%)	793(63.8%)	
	Yes	1,553(37.5%)	1,104(38.0%)	449(36.2%)	

2

Table 2(on next page)

Univariate and multivariate COX analysis of OS in training set

1

Characteristics	Univariate analysis	Multivariate analysis		
	p-value	HR	95%CI	p-value
Race	0.028			0.041
White		Reference		
Black		1.15	1.03-1.28	0.016
Others		0.97	0.84-1.12	0.670
Gender	<0.001			0.019
Male		Reference		
Female		0.91	0.83-0.98	0.019
Primary site	<0.001	NA		
Head				
Body				
Tail				
Others				
Grade	<0.001			＜0.001
High differentiated		Reference		
Low differentiated		1.59	1.40-1.80	＜0.001
Unknown		1.18	1.05-1.31	0.004
Pathological types	<0.001			＜0.001
Ductal adenocarcinoma		Reference		
Non-ductal adenocarcinoma		0.40	0.36-0.45	＜0.001
AJCC	<0.001			＜0.001
I		Reference		
II		3.05	2.23-3.96	＜0.001
III		3.10	2.30-4.18	＜0.001
IV		4.63	3.53-6.07	＜0.001
Surgery	<0.001			＜0.001
No		Reference		
Yes		0.43	0.38-0.50	＜0.001

2

Table 3(on next page)

Univariate and multivariate competing analysis of CSS in training set

Table 3. Univariate and multivariate competing analysis of CSS in training set

	Univariate analysis	Multivariate analysis		
	p-value	SHR	95%CI	p-value
Race		NA		
White		Reference		
Black	0.210	1.07	0.94-1.21	0.30
Others	0.260	1.06	0.90-1.24	0.51
Gender		NA		
Male		Reference		
Female	0.78	1.02	0.94-1.12	0.63
Primary site				
Head		Reference		
Body	0.046	1.01	0.88-1.17	0.867
Tail	0.091	1.09	0.9686-1.24	0.1821
Others	0.750	0.87	0.77-0.99	0.037
Grade				
High differentiated		Reference		
Low differentiated	0.024	1.09	0.96-1.23	0.178
Unknown	<0.001	1.098	0.97-1.21	0.148
Pathological types				
Ductal adenocarcinoma		Reference		
Non-ductal adenocarcinoma	<0.001	0.57	0.50-0.65	<0.001
AJCC				
I		Reference		
II	0.063	1.22	0.998-1.510	0.072
III	<0.001	1.26	0.989-1.61	0.0692
IV	<0.001	1.53	1.22-1.932	<0.001
Surgery				
No		Reference		
Yes	<0.001	0.77	0.68-0.876	<0.001

Figure 1

Flowchart of patient selection

Detailed selection of PC patients in 2004-2015 from SEER database.

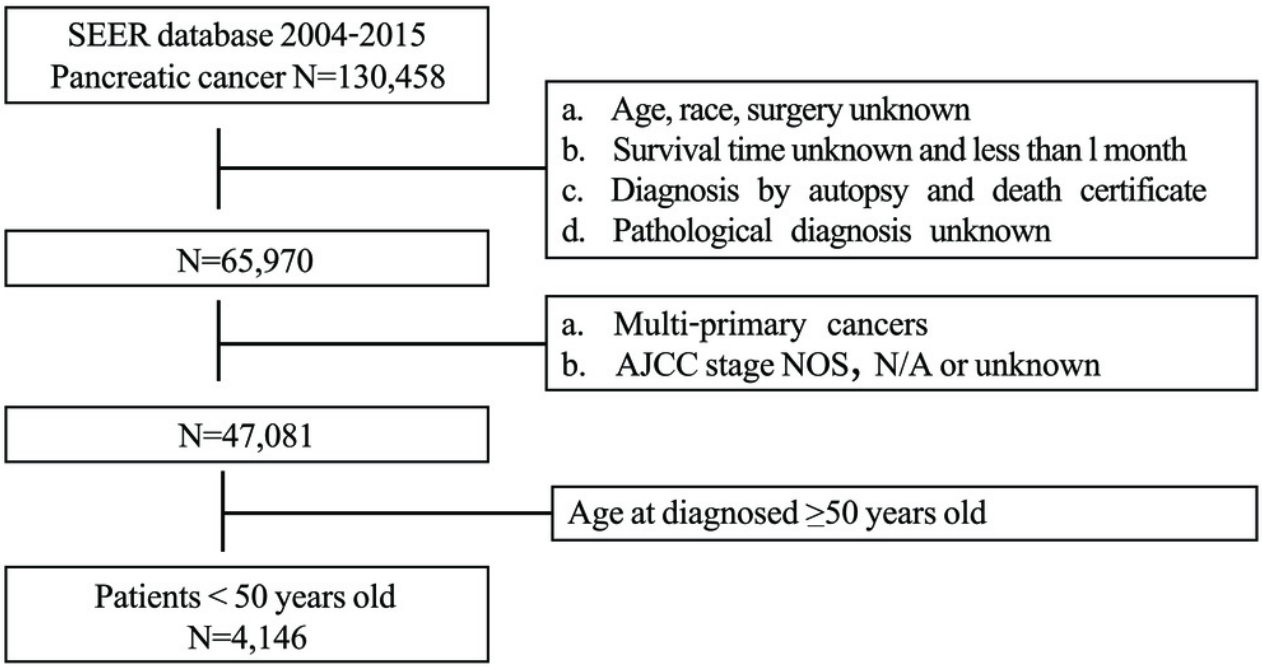


Figure 2

Nomogram for predicting OS and CSS of young patients with PC.

A. Nomogram for predicting 1-, 3- and 5-year OS for young patients with PC; B. Nomogram for predicting 1-, 3- and 5-year CSS for young patients with PC

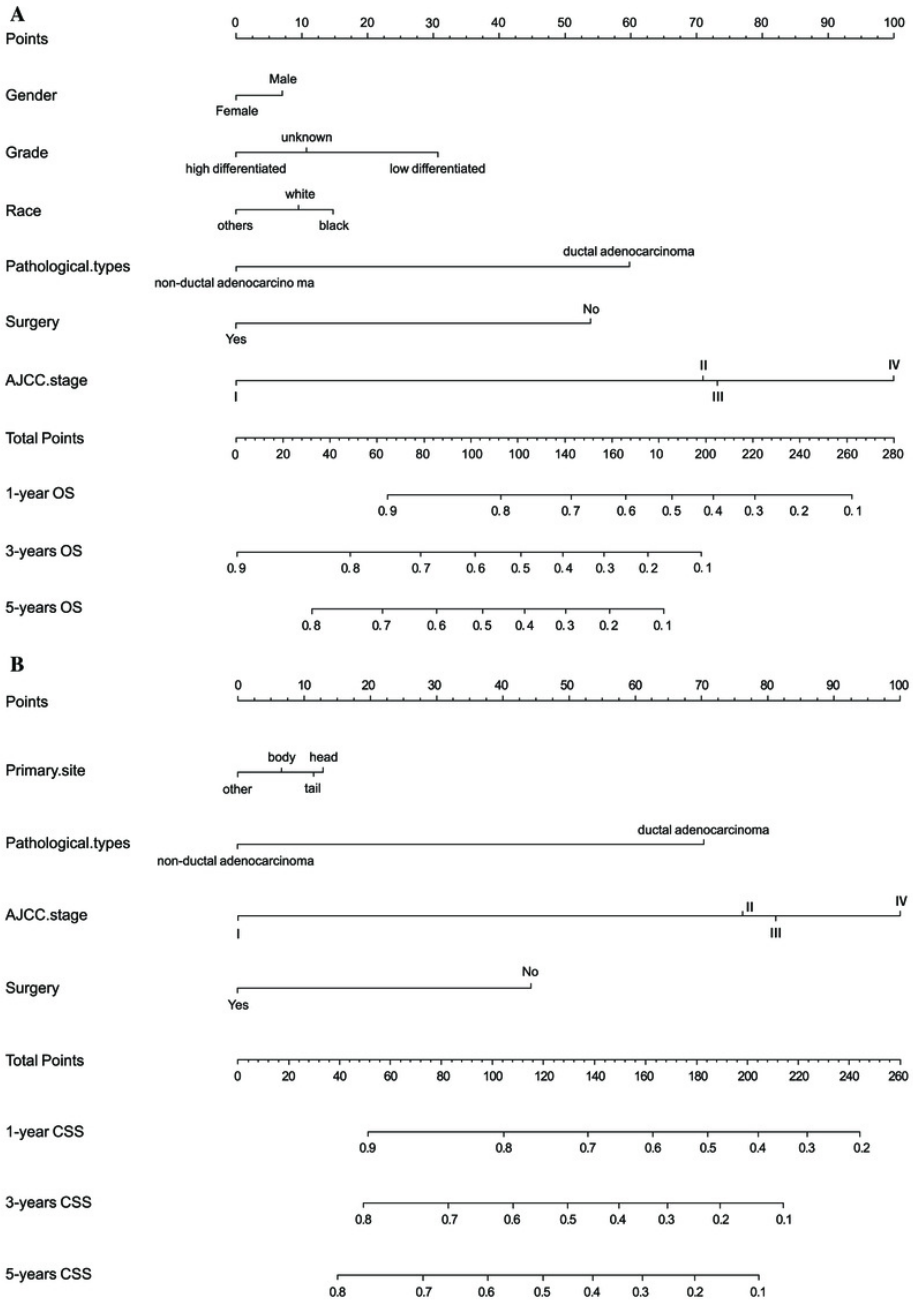

A
Points
0 10 20 30 40 50 60 70 80 90 100
Gender
Male
Female
Grade
unknown
high differentiated
low differentiated
Race
white
others
black
Pathological.types
ductal adenocarcinoma
non-ductal adenocarcino ma
Surgery
No
Yes
AJCC.stage
II
IV
I
III
Total Points
0 20 40 60 80 100 100 140 160 10 200 220 240 260 280
1-year OS
0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1
3-years OS
0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1
5-years OS
0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1
B
Points
0 10 20 30 40 50 60 70 80 90 100
Primary.site
body
head
other
tail
Pathological.types
ductal adenocarcinoma
non-ductal adenocarcinoma
AJCC.stage
II
IV
I
III
Surgery
No
Yes
Total Points
0 20 40 60 80 100 120 140 160 180 200 220 240 260
1-year CSS
0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.2
3-years CSS
0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1
5-years CSS
0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1

Figure 3

Calibration curves for 1-, 3- and 5-year OS in training set and validation set.

A. Calibration curves for 1-year OS in training set; B. Calibration curves for 3-year OS in training set; C. Calibration curves for 5-year OS in training set; D. Calibration curves for 1-year OS in validation set; E. Calibration curves for 3-year OS in validation set; F. Calibration curves for 5-year OS in validation set.

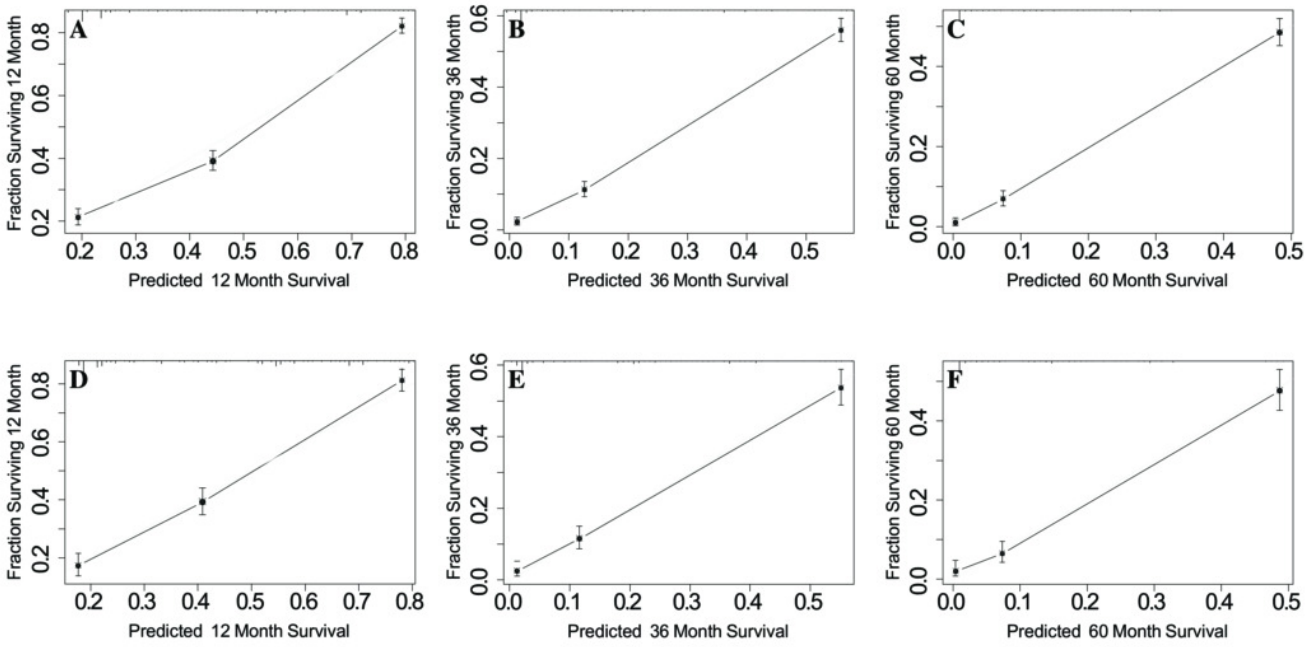
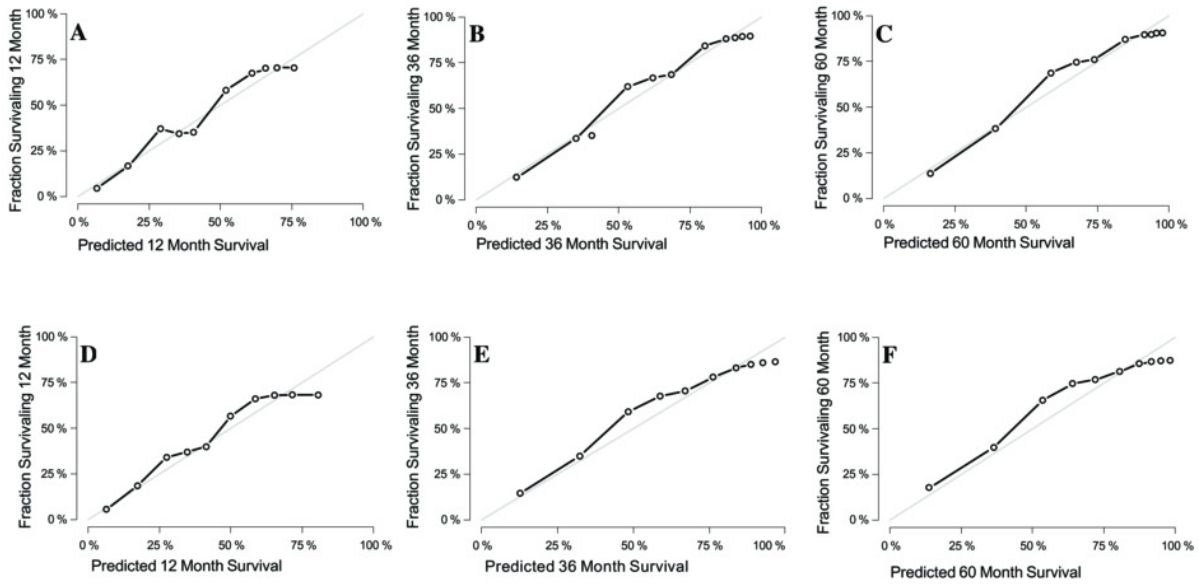


Figure 4

Calibration curves for 1-, 3- and 5-year CSS in training set and validation set.

A. Calibration curves for 1-year CSS in training set; B. Calibration curves for 3-year CSS in training set; C. Calibration curves for 5-year CSS in training set; D. Calibration curves for 1-year CSS in validation set; E. Calibration curves for 3-year CSS in validation set; F. Calibration curves for 5-year CSS in validation set.