**Name of Journal:** *World Journal of Gastrointestinal Oncology*

**Manuscript NO:** 81289

**Manuscript Type:** ORIGINAL ARTICLE

***Retrospective Study***

**Risk factors, prognostic predictors, and nomograms for pancreatic cancer patients with initially diagnosed synchronous liver metastasis**

Cao BY *et al*. Diagnostic and prognostic nomograms for PCLM

Bi-Yang Cao, Fang Tong, Le-Tian Zhang, Yi-Xin Kang, Chen-Chen Wu, Qian-Qian Wang, Wei Yang, Jing Wang

**Bi-Yang Cao, Fang Tong, Le-Tian Zhang, Yi-Xin Kang, Chen-Chen Wu, Qian-Qian Wang, Wei Yang, Jing Wang,** Department of Radiation Oncology, First Medical Center of Chinese PLA General Hospital, Beijing 100853, China

**Bi-Yang Cao, Le-Tian Zhang, Yi-Xin Kang, Chen-Chen Wu,** Medical School of Chinese PLA, Beijing 100853, China

**Author contributions:** Cao BY and Tong F contributed equally to this work; Cao BY and Tong F designed the research and wrote the first manuscript; Zhang LT, Kang YX, Wu CC, and Wang QQ contributed to conceiving the research and analyzing the data; Yang W and Wang J conducted the analysis and provided guidance for the research, and they were co-corresponding authors, and all authors reviewed and approved the final manuscript.

**Corresponding author: Wei Yang, PhD, Professor,** Department of Radiation Oncology, First Medical Center of Chinese PLA General Hospital, No. 28 Fuxing Road, Haidian District, Beijing 100853, China. yangwei1437@163.com

**Received:** November 2, 2022

**Revised:** November 17, 2022

**Accepted:** December 7, 2022

**Published online:**

**Abstract**

BACKGROUND

Liver metastasis (LM) remains a major cause of cancer-related death in patients with pancreatic cancer (PC) and is associated with a poor prognosis. Therefore, identifying the risk and prognostic factors in PC patients with LM (PCLM) is essential as it may aid in providing timely medical interventions to improve the prognosis of these patients. However, there are limited data on risk and prognostic factors in PCLM patients.

AIM

To investigate the risk and prognostic factors of PCLM and develop corresponding diagnostic and prognostic nomograms.

METHODS

Patients with primary PC diagnosed between 2010 and 2015 were reviewed from the Surveillance, Epidemiology, and Results Database. Risk factors were identified using multivariate logistic regression analysis to develop the diagnostic mode. The least absolute shrinkage and selection operator Cox regression model was used to determine the prognostic factors needed to develop the prognostic model. The performance of the two nomogram models was evaluated using receiver operating characteristic (ROC) curves, calibration plots, decision curve analysis (DCA), and risk subgroup classification. The Kaplan-Meier method with a log-rank test was used for survival analysis.

RESULTS

We enrolled 33459 patients with PC in this study. Of them, 11458 (34.2%) patients had LM at initial diagnosis. Age at diagnosis, primary site, lymph node metastasis, pathological type, tumor size, and pathological grade were identified as independent risk factors for LM in patients with PC. Age > 70 years, adenocarcinoma, poor or anaplastic differentiation, lung metastases, no surgery, and no chemotherapy were the independently associated risk factors for poor prognosis in patients with PCLM. The C- index of diagnostic and prognostic nomograms were 0.731 and 0.753, respectively. The two nomograms could accurately predict the occurrence and prognosis of patients with PCLM based on the observed analysis results of ROC curves, calibration plots, and DCA curves. The prognostic nomogram could stratify patients into prognostic groups and perform well in internal validation.

CONCLUSION

Our study identified the risk and prognostic factors in patients with PCLM and developed corresponding diagnostic and prognostic nomograms to help clinicians in subsequent clinical evaluation and intervention. External validation is required to confirm these results.

**Key Words:** Pancreatic neoplasms; Neoplasm metastasis; Liver; Prognosis; Nomograms; Surveillance, Epidemiology, and End Result program

Cao BY, Tong F, Zhang LT, Kang YX, Wu CC, Wang QQ, Yang W, Wang J. Risk factors, prognostic predictors, and nomograms for pancreatic cancer patients with initially diagnosed synchronous liver metastasis. *World J Gastrointest Oncol* 2022; In press

**Core Tip:** Pancreatic cancer (PC) has a poor prognosis owing to its risk of metastasis. The risk and prognostic factors for patients with distant metastasis at diagnosis have been studied; however, few studies have focused on liver metastasis (LM), the most common target organ. This study investigated the risk and prognostic factors for PC patients with LM at diagnosis using the Surveillance, Epidemiology, and End Results database on a population-based level and developed two nomograms for predicting the risk and prognosis for these patients. The nomograms developed in this study can be a convenient tool for facilitating clinical decision-making.

**INTRODUCTION**

Pancreatic cancer (PC) has now become a public health concern, with the National Institutes of Health reporting a 5-year survival rate of 10.2%[1]. It is the second most common cause of cancer-related deaths after lung cancer. In 2021, there were 48220 PC-related deaths in the United States, accounting for 7.9% of all cancer deaths[1]. PC is highly aggressive and prone to early metastasis. Approximately 50% of patients are diagnosed with metastatic disease at the onset[2]. PC commonly causes metastasis to the liver (90%), lymph nodes (LN) (25%), lung (25%), peritoneum (20%), and bones (10%-15%)[3]. The liver is the most common site of metastasis from PC, and it has a significantly poorer prognosis than lung or other distant [metastases](https://www.baidu.com/link?url=Drk7h5AROGD9oYNucF4AhLkoHKePNdFEQDjCNhDiUiwAFkl9XMW1Q81Z8_bCRfIYeqgvyVBHFpfSkIFdXV7AlbtwIfk6agmmDVh-RgAQ0Ae&wd=&eqid=bb2088b00000b10c00000003638ffbfd)[4].

Systemic chemotherapy is the standard treatment for patients with metastatic PC (MPC). Although fluorouracil, irinotecan, leucovorin, and oxaliplatin (FOLFIRINOX) and nab-paclitaxel plus gemcitabine regimens improve patient survival than gemcitabine alone, the median overall survival (OS) of patients with MPC is less than 12 mo[5]. PC patients with LM (PCLM) had a median survival period of less than 6 mo[6]. Patients with PC are generally treated differently depending on whether they have LM. Furthermore, if LM is diagnosed earlier, aggressive chemotherapeutic treatment will increase the survival rate than routine chemotherapeutic treatment alone. Essentially, all PC are at risk of developing LM. Most patients are asymptomatic and are detected using computed tomography (CT) or magnetic resonance imaging (MRI) at an outpatient follow-up visit. Therefore, early identification of patients at high risk for LM and intervening is critical for oncologists.

Because of the lack of large-scale population-based studies, the clinical characteristics of PCLM remain poorly investigated. The Surveillance, Epidemiology, and End Result (SEER) database is a population-based publicly available cancer reporting system that is available nationally. In epidemiological studies, the SEER database can be used to identify risk factors for LM from PC and adjust treatment regimens based on demographic factors[7]. Nomograms are predictive tools that are widely used in cancer prognosis research. It can simplify statistical predictive models to the point where they can be expressed mathematically as a single probability estimate of occurrence. Furthermore, nomograms are considered an alternative to the conventional TNM staging systems used for PC, breast cancer, and hepatocellular carcinoma[8-10]. Therefore, the aims of our study were: (1) To use the SEER database to analyze the risk and prognostic factors of PCLM patients; and (2) To develop nomograms for predicting the risk and prognosis of PCLM patients.

**MATERIALS AND METHODS**

***Patient selection***

Data from newly diagnosed patients with PC between 2010 and 2015 were obtained from the SEER database, the nation’s largest cancer database in the United States. The SEER database, established from the information provided by 26% of the population across 18 cancer registries in the United States, offers information on the survival characteristics and incidence of malignant tumors. Patients with PC were defined as those who were pathologically diagnosed with primary pancreatic malignant tumors using the International Classification of Diseases for Oncology (ICD-O) codes C25.0-C25.3 and C25.7-C25.9. The inclusion criteria used in this study were as follows: (1) Patients with PC diagnosed between 2010 and 2015; (2) Pathologic diagnosis of PC; (3) Clear LM status; and (4) Older than 18 years at the time of diagnosis. The exclusion criteria were as follows: (1) Diagnosis at autopsy or *via* death certificates; (2) Non-primary tumor or multiple primary tumors; (3) Survival time less than 1 mo; (4) Unknown surgical procedure for the primary and distant sites; (5) Incomplete or missing information concerning follow-up or other demographic and clinical characteristics; and (6) Unknown status of the lung, bone, or brain metastases. Supplementary Figure 1 shows the detailed selection process for the inclusion of patients. The data for this study were collected from the SEER database using SEER\*Stat v8.3.8 software (seer.cancer.gov/seerstat). The SEER Research Plus Data Agreement was signed, and the license for analyzing the study data was obtained in November 2021 (username: 15159-Nov2020).

***Data collection***

We collected information on the following baseline characteristics of patients with PC: age at diagnosis, race, sex, primary site, histological type, pathological grade, T stage, LN metastasis, tumor size, surgery (including surgery to the primary and distant sites), radiotherapy, chemotherapy, brain metastasis, bone metastasis, and lung metastasis. The following variables were used to determine risk factors for LM in patients with PC: Age at diagnosis, race, sex, primary site, histological type, pathological grade, T stage, node status, and tumor size. Survival analysis was conducted in PCLM patients to determine prognostic factors. Furthermore, three treatment options were investigated: Surgery, radiotherapy, and chemotherapy. Based on the cutoff value of the median age at diagnosis of PC, patients were classified into two groups, 70 years or younger and older than 70 years. In addition, histological codes were divided into three categories based on ICD-O-3 codes: Adenocarcinoma (histologic codes 8140, 8480, 8500), neuroendocrine carcinoma (histologic code 8246), and others (histologic codes 8010, 8012, 8013, 8020, 8021, 8041, 8046, 8070, 8150, 8240, 8244, 8249, 8481, 8490, and 8560). Tumor size was divided into three groups based on the American Joint Committee on Cancer 8th edition T classification: < 2 cm, 2-4 cm, and > 4 cm. The study’s primary endpoint was OS, defined as the time between diagnosis and death for any reason.

***Statistical analysis***

R software (version 4.0.1; R Core Team, 2022) was used for all statistical analysis, and a *P* value < 0.05 (two sides) was considered statistically significant. Using R software, we randomly assigned all patients with PC into training and validation groups and compared the distributions between the two groups using the chi-square test or Fisher’s exact test.

Univariate logistic regression analysis was used to select risk factors for LM, and those with *P* < 0.05 were subsequently evaluated using multivariate logistic regression analysis to identify independent risk factors. A novel diagnostic nomogram based on the identified independent risk factors was developed using the “rms” package. Receiver operating characteristic (ROC) curves for the nomogram and the corresponding area under the curve (AUC) were calculated to evaluate the discrimination. The nomogram's performance was assessed using decision curve analysis (DCA) and calibration curves.

To select nomogram characteristics in the prognostic cohort, a two-step analysis was performed. First, the least absolute shrinkage and selection operator (LASSO) method was used to select beneficial predictive factors, minimizing overfitting. The training set’s most significant LASSO regression attributes were further investigated using multivariate Cox proportional hazards analysis. Furthermore, the variables with *P* value < 0.05 in the multivariate analysis were incorporated into a nomogram aimed to predict the OS of PCLM patients, and the individual risk score was calculated using the nomogram formula. Additionally, the AUCs of the nomogram’s time-dependent ROC were estimated at 6, 12, and 18 mo to demonstrate prediction accuracy. Calibration and DCA curves were plotted at 6, 12, and 18 mo to analyze the nomograms. All patients with PC and LM were divided into high-and low-risk groups based on the median risk score. The difference in OS between the two groups was investigated using a Kaplan-Meier survival curve and the log-rank test.

**RESULTS**

***Patients baseline clinical characteristics***

33459 patients met the inclusion criteria with PC diagnosed between 2010 and 2015. 11458 patients (34.2%) had LM at initial diagnosis, and 22001 patients did not. The entire cohort’s median survival time was 9 mo, and the median age at diagnosis was 67 years. Table 1 shows the clinical characteristics of the patients with PC, with or without LM. PC patients with LM were much younger than PC patients without LM (≤ 70 years: 65.9% *vs* 59.7%, *P* < 0.001), more likely to be men (54.7% *vs* 49.1%, *P* < 0.001), and more commonly black (13.4% *vs* 11.8%, *P* < 0.001). The presence of LM was most common among patients with the body or tail involved by the tumor (41.5% *vs* 24.7%, *P* < 0.001). Furthermore, patients with LN metastasis (47.3% *vs* 43.5%, *P* < 0.001), tumors > 4 cm in diameter (56.3% *vs* 39.7%, *P* < 0.001), bone metastasis (5.6% *vs* 1.3%, *P* < 0.001), and lung metastasis (15.3% *vs* 5.3%, *P* < 0.001) had a higher risk of developing LM. Patients with LM were less likely to undergo surgery at the primary site (3.5% *vs* 40.7%, *P* < 0.001), radiation treatment (4.8% *vs* 24.7%, *P* < 0.001), and chemotherapy (62.5% *vs* 59.7%, *P* < 0.001).

***Independent risk factors for PCLM***

We enrolled and analyzed 33459 patients with PC for LM-related risk factors and stratified 22306 and 11153 patients into the training and validation sets at a 2:1 ratio. Supplementary Table 1 summarizes the baseline characteristics of patients in each set. Overall, baseline characteristics were balanced between the training and validation sets. Univariate logistic analysis revealed that all nine potential factors were significant (Table 2). Multivariate logistic regression analysis showed that six associated factors were independent risk factors for LM in PC patients: Age ≤ 70 years, primary site in the body or tail, LN metastasis, adenocarcinoma or neuroendocrine carcinoma, a larger tumor size, and a higher grade (Table 2).

***Diagnostic nomogram model establishment and validation***

Based on the independent risk factors identified by multivariable logistic regression, a nomogram model for predicting the risk of LM in patients with PC was developed (Figure 1). The C-index of the diagnostic nomogram was 0.731 [95% confidence interval (CI): 0.623-0.823]. According to ROC analysis, the AUC of the nomogram was 0.728, indicating excellent discriminatory ability (Figure 2A). The observed results were consistent with the predicted outcomes based on the calibration curve (Figure 2B). Additionally, DCA demonstrated the effectiveness of the nomogram model in clinical practice (Figure 2C). An internal validation cohort was created, and the corresponding validation curves were plotted to validate the model. ROC analysis revealed an AUC value of 0.737 for the nomogram, showing high discrimination in the validation population (Figure 2D). The calibration curve revealed excellent agreement between nomogram predictions and actual observations, and the internal verification cohort agreed with the training cohort (Figure 2E). DCA demonstrated in the validation cohort that the nomogram model performed well in clinical practice (Figure 2F).

***Independent prognostic factors for PCLM***

The 11458 eligible PC patients with LM were included in this study to explore the prognostic factors. Supplementary Table 2 shows no significant differences between the validation and training sets. Sixteen variables were included in this analysis. In the LASSO regression analysis, age at diagnosis, histology, pathological grade, surgery at the primary site, surgery at the distant site, chemotherapy, and lung metastasis were identified as OS risk factors (Figure 3). In the multivariate analysis, all seven factors were independently associated with OS and statistically significant (Table 3).

***Establishment and validation of prognostic nomogram model***

A nomogram was developed using these seven prognostic factors to predict the survival of patients with PCLM (Figure 4). The C-index of the prognostic nomogram was 0.753 (95%CI: 0.658-0.823). Based on the calibration curves of the nomograms for the probability of 6-, 12-, and 18-mo OS in training (Figures 5A-C) and validation (Figures 6A-C) sets, there was strong agreement between the nomogram-estimated OS and the actual outcome. Additionally, DCA curves indicated that the nomogram had acceptable performance in clinical practice (Figures 5D-F and 6D-F). Additionally, ROC analysis indicated that the AUCs of the nomogram for the training set at 6, 12, and 18 mo were 0.751, 0.749, and 0.775, respectively (Figure 7A). However, the corresponding values for the validation set were 0.751, 0.751, and 0.784, respectively (Figure 7B), suggesting that the nomogram also provided reasonable accuracy in predicting the survival of PC patients with LM. As evidenced by the Kaplan-Meier curves, a shorter OS was observed in the high-risk group than in the low-risk group (Figures 7C and 7D).

**DISCUSSION**

Our study evaluated the risk and prognostic factors for patients with PCLM using logistic regression and LASSO-Cox analyses. Furthermore, we incorporated the risk and prognostic factors identified to generate a diagnostic and prognostic nomogram. A score can be calculated based on information extracted from nomograms related to diagnosis and prognosis, thereby guiding subsequent clinical evaluation and intervention.

As revealed by the diagnostic nomogram, patients aged ≤ 70 years, with tumors of the body and/or tail, neuroendocrine cancer, poor or anaplastic differentiation, LN metastasis, and large tumors (> 4 cm) were more likely to develop LM. Younger patients have tumors with higher and more aggressive histopathology[11,12]. The presence of numerous genetic alterations in younger patients further supports the hypothesis that immature tumor cells are more prone to DNA damage. Therefore, younger patients may be at greater risk of developing metastasis. As explained in many previous reports[13,14], our results suggest that primary tumors located in the tail and body of the pancreas are more prone to metastasis to the liver than those located in the head of the pancreas. Most commonly, larger or more advanced tumors are found within the body or tail of the pancreas than within the head, perhaps because of the absence of obstructive jaundice, which may increase the risk of LM.Our findings indicated that pancreatic neuroendocrine tumors (PNETs) had a higher risk of LM than pancreatic adenocarcinomas. In a recent study, the DNA of neutrophil extracellular traps was found to play a crucial role in promoting LM in pancreatic neuroendocrine cells. This mechanism differed from previous studies on cytokines and chemokines, integrin complexes, metabolic programming, and proliferation signaling[15]. Poorly differentiated tumor cells are typically more aggressive[16], which may partially explain our finding that poor differentiation is significantly associated with a higher risk of LM. There is scientific evidence that LN metastasis is a common sign before distant metastasis[17,18], suggesting that patients with PC with LN metastasis should pay greater attention to distant metastasis. However, tumor size, which has proven to be a strong and consistent indicator of both distant metastasis and poor prognosis[19,20], was found to be directly correlated with the invasion of cancer cells into the liver. Larger tumors tend to be more aggressive and susceptible to the involvement of adjacent organs and blood vessels. This may indicate a greater tumor burden in patients with PC. Therefore, clinicians must maintain a keen awareness of these risk factors when treating patients with PC. MRI/positron emission tomography-CT is recommended for patients at an early stage with potential risks.

In addition, a survival analysis was performed. Our research suggests that being aged > 70 years, adenocarcinoma, poor or anaplastic differentiation, accompanying lung metastasis, and no previous surgery or chemotherapy might indicate a poor prognosis for patients with PCLM. The prognosis of older PC patients has generally been reported to be poorer than that of younger patients[21,22]. Our study indicated that older patients diagnosed with PCLM were more likely to have a shorter survival time, confirming previous studies’ results that demonstrated that older PCLM patients had a worse prognosis[23,24]. Despite our inability to obtain additional relevant data from the database, we hypothesized that this was likely related to impaired performance and reduced immunity in older individuals. Adenocarcinoma is the most common histological subtype of PC. In our study, adenocarcinomas constituted 76.3% of all tumors in patients with PCLM, which represents the most prevalent histological subtype. In our research, the adenocarcinoma subtype was a significant independent predictor of poorer OS in patients with PCLM, consistent with previous studies that showed that PC patients with adenocarcinoma had a lower OS rate and poor prognosis[25,26]. It has been suggested that non-adenocarcinoma tumors are distinct from pancreatic adenocarcinoma neoplasms in terms of morphology and biology[27]. Further studies are required to elucidate the specific mechanisms involved. Poor or anaplastic differentiation is associated with a poorer prognosis, as it has been observed to be strongly associated with a worse prognosis in multiple cancers[28,29], reflecting the nature of the tumor. We found that lung metastasis may indicate a poor prognosis in PC patients with LM. Lung metastasis worsens the prognosis of patients with PCLM, which is consistent with the results of previous PC studies [30,31].

Patients who received chemotherapy had superior survival based on prognostic analysis. Systemic chemotherapy is the primary treatment option, and recent studies have reported several palliative chemotherapy options for patients with MPC, including FOLFIRINOX or gemcitabine and nab-paclitaxel, which improved the median OS to 11 and 8.5 mo, respectively, compared with 6.7-7.0 mo for gemcitabine alone[5]. Patients with PCLM continue to have a poor prognosis owing to secondary chemoresistance[32], highlighting the urgent need for new treatment strategies. Our study also demonstrated that resecting the primary tumor or metastatic site could prolong survival in patients with PCLMs. Such treatment options should be considered for carefully selected patients. Cancer-directed surgery reliably increased the median OS from 5 to 10 mo in patients with pancreatic ductal adenocarcinoma with LM in a propensity-matched analysis of the SEER database[33]. Furthermore, several types of PNETs are treated with cancer-directed surgery, regardless of LM[34]. Up to 30% of patients with PNET develop LM at the time of diagnosis[35], and type I/II LM resection is recommended for G1/G2 PNETs[35]. Four decades ago, surgery was primarily used as palliative treatment for colorectal cancer LM (CRLM)[36]. Currently, resection is considered the standard of care for CRLM[37]. Patients with PC commonly develop subclinical metastases, which must be addressed macroscopically and microscopically. Therefore, surgical therapy involving the primary or distant site should be recommended when a multidisciplinary team at a high-volume PC center carefully selects patients with PCLM. Multi-center randomized controlled studies are required to confirm updated standards of care in this field. We provide a set of targets and hypotheses to guide the research agenda.

This study has several limitations. First, as part of this study, a retrospective analysis of the SEER database was conducted with potential selection bias and insufficient data granularity. Several potential prognostic variables, such as chemotherapy regimens, the extent of metastasis, and tumor markers such as carbohydrate antigen 19-9, as well as the patient’s physical condition, preoperative conditions, and postoperative complications, were not provided, which limited the estimation of their role in the nomogram. Second, our study was limited to patients who only had LM at diagnosis, and those with LM occurring at a later stage were excluded. Third, owing to data heterogeneity, we evaluated the nomogram only through our internal validation method; no publicly available data were enrolled for external validation to avoid selective bias. Finally, even though the nomogram has achieved acceptable predictive performance and a relatively complete evaluation to accurately estimate the risk and prognosis of PCLM, further investigation should be conducted.

**CONCLUSION**

Our study identified the risk and prognostic factors in patients with PCLM. The nomograms developed in this study can aid clinicians in providing better prognoses and prevention in high-risk patients.

**ARTICLE HIGHLIGHTS**

***Research background***

Pancreatic cancer (PC) with liver metastasis (LM) is a commonly fatal disease with an extremely poor prognosis. Identifying the risk and prognostic factors of PC patients with LM (PCLM) is essential and may aid in providing timely medical interventions to improve the prognosis of these patients.

***Research motivation***

Few studies have focused on investigating PCLM patients’ risk and prognostic factors, and there are no corresponding diagnostic and prognostic nomograms for these entities.

***Research objectives***

This study aimed to investigate the risk and prognostic factors of PCLM and establish corresponding diagnostic and prognostic nomograms.

***Research methods***

Patients from the Surveillance, Epidemiology, and End Results database with primary PC diagnosed between 2010 and 2015 were reviewed. A multivariate logistic regression analysis was used to identify risk factors to develop a diagnostic model. The least absolute shrinkage and selection operator Cox regression model was used to determine the prognostic factors used to establish a prognostic model. The performance of the two nomogram models was evaluated using receiver operating characteristic (ROC) curves, calibration plots, decision curve analysis (DCA), and risk subgroup classification. Survival analysis was performed using the Kaplan-Meier method with a log-rank test.

***Research results***

A total of 33459 PC patients were included in the study, with 11458 patients (34.2%) having LM at initial diagnosis. Age at diagnosis, primary site, lymph node metastasis, pathological type, tumor size, and pathological grade were identified as independent risk factors for LM in patients with PC. The independent factors associated with poor prognosis for patients with PCLM include age > 70 years, adenocarcinoma, poor or anaplastic differentiation, lung metastasis, no surgery, and no chemotherapy. The C-indices of the diagnostic and prognostic nomograms were 0.731 and 0.753, respectively. Based on the observed analysis results of ROC curves, calibration plots, and DCA curves, the two nomograms could accurately predict the occurrence and prognosis of patients with PCLM. The prognostic nomogram could stratify patients into prognostic groups and perform well in terms of internal validation.

***Research conclusions***

Our study identified the risk and prognostic factors in patients with PCLM and constructed corresponding diagnostic and prognostic nomograms to guide subsequent clinical evaluation and intervention for clinicians.

***Research perspectives***

The nomograms constructed in this study can help clinicians provide better prevention for high-risk subjects and monitor their prognosis. External validation is required to verify these results.

**ACKNOWLEDGEMENTS**

We would like to express our appreciation to the Surveillance, Epidemiology, and End Results Program tumor registries, which were instrumental in establishing the Surveillance, Epidemiology, and End Results Database.

**REFERENCES**

1 **National Cancer Institute**. Cancer Stat Facts: Pancreatic Cancer. [cited 5 October 2022]. Available from: https://seer.cancer.gov/statfacts/html/pancreas.html

2 **Mizrahi JD**, Surana R, Valle JW, Shroff RT. Pancreatic cancer. *Lancet* 2020; **395**: 2008-2020 [PMID: 32593337 DOI: 10.1016/S0140-6736(20)30974-0]

3 **He C**, Huang X, Zhang Y, Lin X, Li S. The impact of different metastatic patterns on survival in patients with pancreatic cancer. *Pancreatology* 2021; **21**: 556-563 [PMID: 33518454 DOI: 10.1016/j.pan.2021.01.014]

4 **Houg DS**, Bijlsma MF. The hepatic pre-metastatic niche in pancreatic ductal adenocarcinoma. *Mol Cancer* 2018; **17**: 95 [PMID: 29903049 DOI: 10.1186/s12943-018-0842-9]

5 **Sohal DPS**, Kennedy EB, Cinar P, Conroy T, Copur MS, Crane CH, Garrido-Laguna I, Lau MW, Johnson T, Krishnamurthi S, Moravek C, O'Reilly EM, Philip PA, Pant S, Shah MA, Sahai V, Uronis HE, Zaidi N, Laheru D. Metastatic Pancreatic Cancer: ASCO Guideline Update. *J Clin Oncol* 2020: JCO2001364 [PMID: 32755482 DOI: 10.1200/JCO.20.01364]

6 **De Dosso S**, Siebenhüner AR, Winder T, Meisel A, Fritsch R, Astaras C, Szturz P, Borner M. Treatment landscape of metastatic pancreatic cancer. *Cancer Treat Rev* 2021; **96**: 102180 [PMID: 33812339 DOI: 10.1016/j.ctrv.2021.102180]

7 **Doll KM**, Rademaker A, Sosa JA. Practical Guide to Surgical Data Sets: Surveillance, Epidemiology, and End Results (SEER) Database. *JAMA Surg* 2018; **153**: 588-589 [PMID: 29617544 DOI: 10.1001/jamasurg.2018.0501]

8 **Zhang W**, Ji L, Wang X, Zhu S, Luo J, Zhang Y, Tong Y, Feng F, Kang Y, Bi Q. Nomogram Predicts Risk and Prognostic Factors for Bone Metastasis of Pancreatic Cancer: A Population-Based Analysis. *Front Endocrinol (Lausanne)* 2021; **12**: 752176 [PMID: 35356148 DOI: 10.3389/fendo.2021.752176]

9 **Ji L**, Cheng L, Zhu X, Gao Y, Fan L, Wang Z. Risk and prognostic factors of breast cancer with liver metastases. *BMC Cancer* 2021; **21**: 238 [PMID: 33676449 DOI: 10.1186/s12885-021-07968-5]

10 **Hu C**, Yang J, Huang Z, Liu C, Lin Y, Tong Y, Fan Z, Chen B, Wang C, Zhao CL. Diagnostic and prognostic nomograms for bone metastasis in hepatocellular carcinoma. *BMC Cancer* 2020; **20**: 494 [PMID: 32487048 DOI: 10.1186/s12885-020-06995-y]

11 **Wang J**, Liu J, Pan H, Jiang C, Liu S, Zhu Z, Fang J, Zheng X, Hong S, Wang S. Young age increases the risk of lymph node positivity in papillary thyroid cancer patients: a SEER data-based study. *Cancer Manag Res* 2018; **10**: 3867-3873 [PMID: 30288118 DOI: 10.2147/CMAR.S167774]

12 **Mete O**, Asa SL. Pathological definition and clinical significance of vascular invasion in thyroid carcinomas of follicular epithelial derivation. *Mod Pathol* 2011; **24**: 1545-1552 [PMID: 21804527 DOI: 10.1038/modpathol.2011.119]

13 **Wagner M**, Redaelli C, Lietz M, Seiler CA, Friess H, Büchler MW. Curative resection is the single most important factor determining outcome in patients with pancreatic adenocarcinoma. *Br J Surg* 2004; **91**: 586-594 [PMID: 15122610 DOI: 10.1002/bjs.4484]

14 **Paye F**, Micelli Lupinacci R, Bachellier P, Boher JM, Delpero JR; French Surgical Association (AFC). Distal pancreatectomy for pancreatic carcinoma in the era of multimodal treatment. *Br J Surg* 2015; **102**: 229-236 [PMID: 25533421 DOI: 10.1002/bjs.9708]

15 **Grande E**, Rodriguez-Antona C, López C, Alonso-Gordoa T, Benavent M, Capdevila J, Teulé A, Custodio A, Sevilla I, Hernando J, Gajate P, Molina-Cerrillo J, Díez JJ, Santos M, Lanillos J, García-Carbonero R. Sunitinib and Evofosfamide (TH-302) in Systemic Treatment-Naïve Patients with Grade 1/2 Metastatic Pancreatic Neuroendocrine Tumors: The GETNE-1408 Trial. *Oncologist* 2021; **26**: 941-949 [PMID: 34190375 DOI: 10.1002/onco.13885]

16 **Matsuo K**, Mandelbaum RS, Machida H, Purushotham S, Grubbs BH, Roman LD, Wright JD. Association of tumor differentiation grade and survival of women with squamous cell carcinoma of the uterine cervix. *J Gynecol Oncol* 2018; **29**: e91 [PMID: 30207099 DOI: 10.3802/jgo.2018.29.e91]

17 **Wu SG**, Zhang WW, Sun JY, Li FY, Lin Q, He ZY. Patterns of Distant Metastasis Between Histological Types in Esophageal Cancer. *Front Oncol* 2018; **8**: 302 [PMID: 30135855 DOI: 10.3389/fonc.2018.00302]

18 **Gaitanidis A**, Alevizakos M, Tsaroucha A, Tsalikidis C, Pitiakoudis M. Predictive Nomograms for Synchronous Distant Metastasis in Rectal Cancer. *J Gastrointest Surg* 2018; **22**: 1268-1276 [PMID: 29663304 DOI: 10.1007/s11605-018-3767-0]

19 **He C**, Zhang Y, Cai Z, Lin X, Li S. Overall survival and cancer-specific survival in patients with surgically resected pancreatic head adenocarcinoma: A competing risk nomogram analysis. *J Cancer* 2018; **9**: 3156-3167 [PMID: 30210639 DOI: 10.7150/jca.25494]

20 **He C**, Mao Y, Wang J, Duan F, Lin X, Li S. Nomograms predict long-term survival for patients with periampullary adenocarcinoma after pancreatoduodenectomy. *BMC Cancer* 2018; **18**: 327 [PMID: 29580215 DOI: 10.1186/s12885-018-4240-x]

21 **Exarchakou A**, Papacleovoulou G, Rous B, Magadi W, Rachet B, Neoptolemos JP, Coleman MP. Pancreatic cancer incidence and survival and the role of specialist centres in resection rates in England, 2000 to 2014: A population-based study. *Pancreatology* 2020; **20**: 454-461 [PMID: 32014435 DOI: 10.1016/j.pan.2020.01.012]

22 **Loveday BPT**, Lipton L, Thomson BN. Pancreatic cancer: An update on diagnosis and management. *Aust J Gen Pract* 2019; **48**: 826-831 [PMID: 31774983 DOI: 10.31128/AJGP-06-19-4957]

23 **Su BB**, Bai DS, Yu JQ, Zhang C, Jin SJ, Zhou BH, Jiang GQ. Can Patients with Pancreatic Cancer and Liver Metastases Obtain Survival Benefit from Surgery? A Population-Based Study. *J Cancer* 2021; **12**: 539-552 [PMID: 33391450 DOI: 10.7150/jca.51218]

24 **Oweira H**, Petrausch U, Helbling D, Schmidt J, Mannhart M, Mehrabi A, Schöb O, Giryes A, Decker M, Abdel-Rahman O. Prognostic value of site-specific metastases in pancreatic adenocarcinoma: A Surveillance Epidemiology and End Results database analysis. *World J Gastroenterol* 2017; **23**: 1872-1880 [PMID: 28348494 DOI: 10.3748/wjg.v23.i10.1872]

25 **Li G**, Chen JZ, Chen S, Lin SZ, Pan W, Meng ZW, Cai XR, Chen YL. Development and validation of novel nomograms for predicting the survival of patients after surgical resection of pancreatic ductal adenocarcinoma. *Cancer Med* 2020; **9**: 3353-3370 [PMID: 32181599 DOI: 10.1002/cam4.2959]

26 **Zou Y**, Han H, Ruan S, Jian Z, Jin L, Zhang Y, Chen Z, Yin Z, Ma Z, Jin H, Dai M, Shi N. Development of a Nomogram to Predict Disease-Specific Survival for Patients After Resection of a Non-Metastatic Adenocarcinoma of the Pancreatic Body and Tail. *Front Oncol* 2020; **10**: 526602 [PMID: 33194585 DOI: 10.3389/fonc.2020.526602]

27 **Dhillon J**. Non-Ductal Tumors of the Pancreas. *Monogr Clin Cytol* 2020; **26**: 92-108 [PMID: 32987393 DOI: 10.1159/000455737]

28 **Chen J**, Cao N, Li S, Wang Y. Identification of a Risk Stratification Model to Predict Overall Survival and Surgical Benefit in Clear Cell Renal Cell Carcinoma With Distant Metastasis. *Front Oncol* 2021; **11**: 630842 [PMID: 33777784 DOI: 10.3389/fonc.2021.630842]

29 **Li Y**, Liu D, Zhao L, Güngör C, Song X, Wang D, Liu W, Tan F. Accurate nomograms with excellent clinical value for locally advanced rectal cancer. *Ann Transl Med* 2021; **9**: 296 [PMID: 33708923 DOI: 10.21037/atm-20-4144]

30 **Lianyuan T**, Deyu L, Haibo Y, Yadong D, Guanjing T. Clinical features and prognostic factors of elderly patients with metastatic pancreatic cancer: a population-based study. *Aging (Albany NY)* 2021; **13**: 7133-7146 [PMID: 33639615 DOI: 10.18632/aging.202570]

31 **Ni X**, Li D, Dai S, Pan H, Sun H, Ao J, Chen L, Kong H. Development and Evaluation of Nomograms to Predict the Cancer-Specific Mortality and Overall Mortality of Patients with Hepatocellular Carcinoma. *Biomed Res Int* 2021; **2021**: 1658403 [PMID: 33860031 DOI: 10.1155/2021/1658403]

32 **Zeng S**, Pöttler M, Lan B, Grützmann R, Pilarsky C, Yang H. Chemoresistance in Pancreatic Cancer. *Int J Mol Sci* 2019; **20** [PMID: 31514451 DOI: 10.3390/ijms20184504]

33 **Pausch TM**, Liu X, Cui J, Wei J, Miao Y, Heger U, Probst P, Heap S, Hackert T. Survival Benefit of Resection Surgery for Pancreatic Ductal Adenocarcinoma with Liver Metastases: A Propensity Score-Matched SEER Database Analysis. *Cancers (Basel)* 2021; **14** [PMID: 35008223 DOI: 10.3390/cancers14010057]

34 **Farley HA**, Pommier RF. Treatment of Neuroendocrine Liver Metastases. *Surg Oncol Clin N Am* 2016; **25**: 217-225 [PMID: 26610783 DOI: 10.1016/j.soc.2015.08.010]

35 **Modlin IM**, Lye KD, Kidd M. A 5-decade analysis of 13,715 carcinoid tumors. *Cancer* 2003; **97**: 934-959 [PMID: 12569593 DOI: 10.1002/cncr.11105]

36 **Bartolini I**, Bencini L, Risaliti M, Ringressi MN, Moraldi L, Taddei A. Current Management of Pancreatic Neuroendocrine Tumors: From Demolitive Surgery to Observation. *Gastroenterol Res Pract* 2018; **2018**: 9647247 [PMID: 30140282 DOI: 10.1155/2018/9647247]

37 **Pawlik TM**, Choti MA. Surgical therapy for colorectal metastases to the liver. *J Gastrointest Surg* 2007; **11**: 1057-1077 [PMID: 17530336 DOI: 10.1007/s11605-006-0061-3]

**Footnotes**

**Institutional review board statement:** Our research is based on the National Cancer Institute’s SEER program. For this study, we signed the SEER research data agreement to access SEER information, using reference number 15159-Nov2020. Data were obtained following the approved guidelines. The Office for Human Research Protection considered this research to be on nonhuman subjects because the subjects were patients who had been researched by the United States Department of Health and Human Services and were publicly accessible and de-dentified. Thus, no institutional review board approval was required.

**Informed consent statement:** Patients were not required to provide informed consent to the study because this study used a public database with anonymous clinical data, and the patients’ personal privacy information was not available.

**Conflict-of-interest statement:** All authors report no relevant conflicts of interest for this article.

**Data sharing statement:** The dataset from SEER database analyzed during the current study are available in the SEER dataset repository (<https://seer.cancer.gov/>).

**Open-Access:** This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: https://creativecommons.org/Licenses/by-nc/4.0/

**Provenance and peer review:** Unsolicited article; Externally peer reviewed

**Peer-review model:** Single blind

**Peer-review started:** November 2, 2022

**First decision:** November 14, 2022

**Article in press:**

**Specialty type:** Oncology

**Country/Territory of origin:** China

**Peer-review report’s scientific quality classification**

Grade A (Excellent): A

Grade B (Very good): B, B

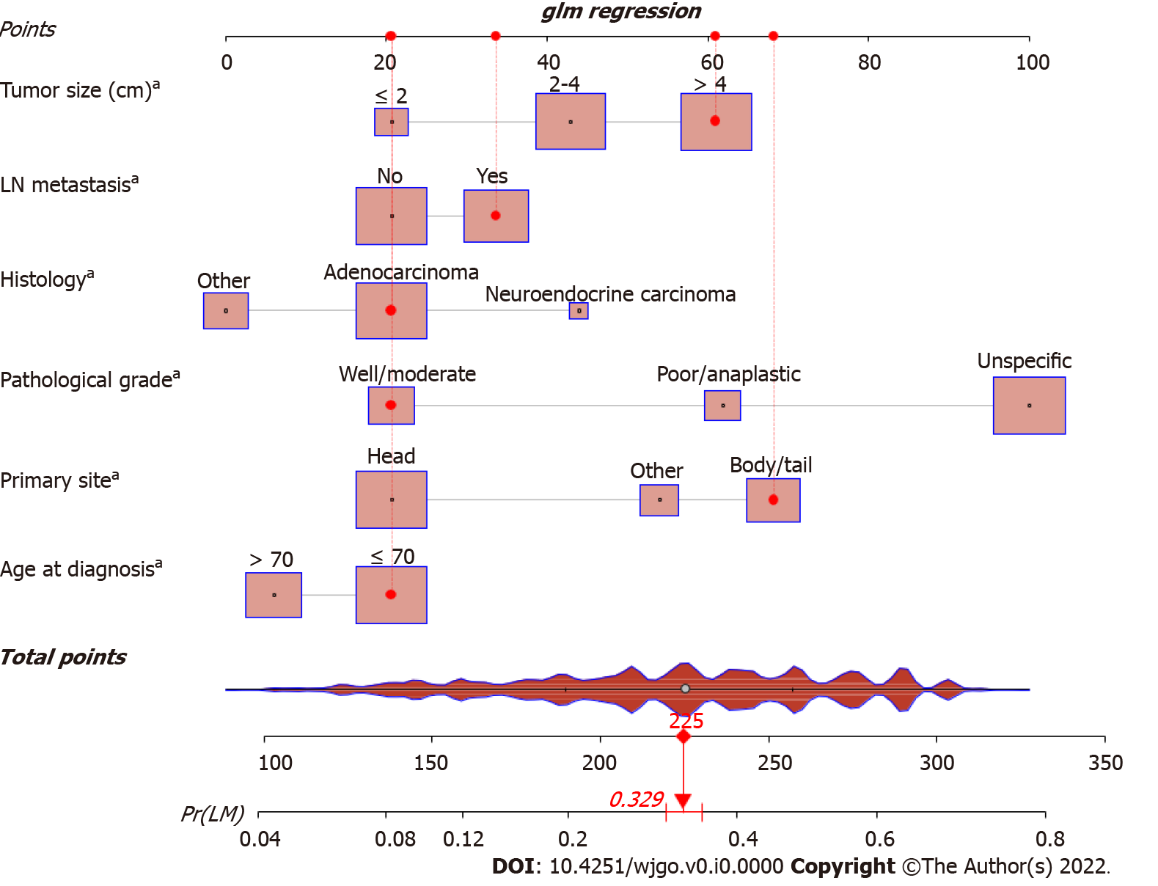
Grade C (Good): 0

Grade D (Fair): 0

Grade E (Poor): 0

**P-Reviewer:** Cascinu S, Italy; Katayama Y, Japan; Yasukawa K, Japan **S-Editor:** Wang JJ **L-Editor:** A **P-Editor:**

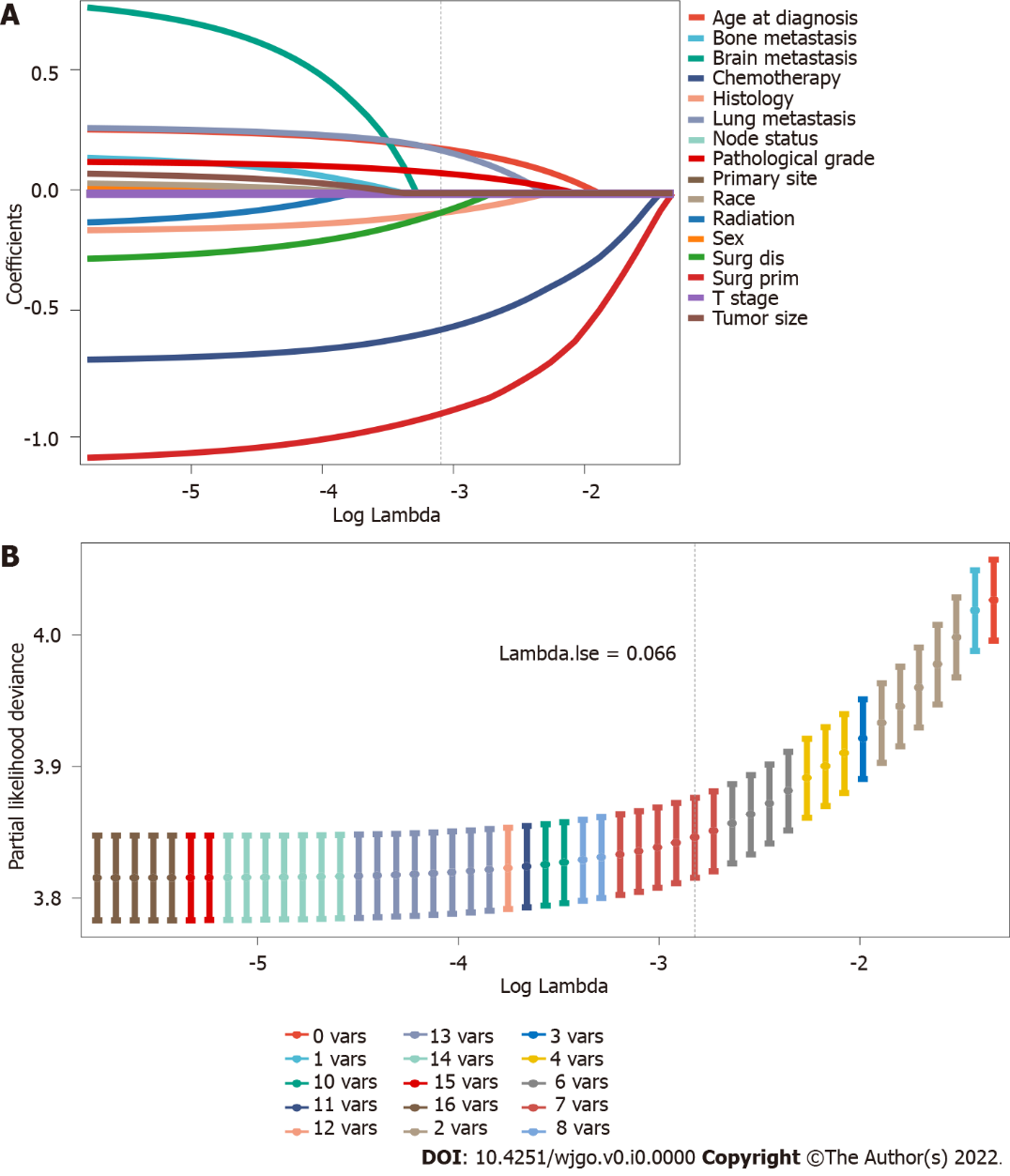
**Figure Legends**



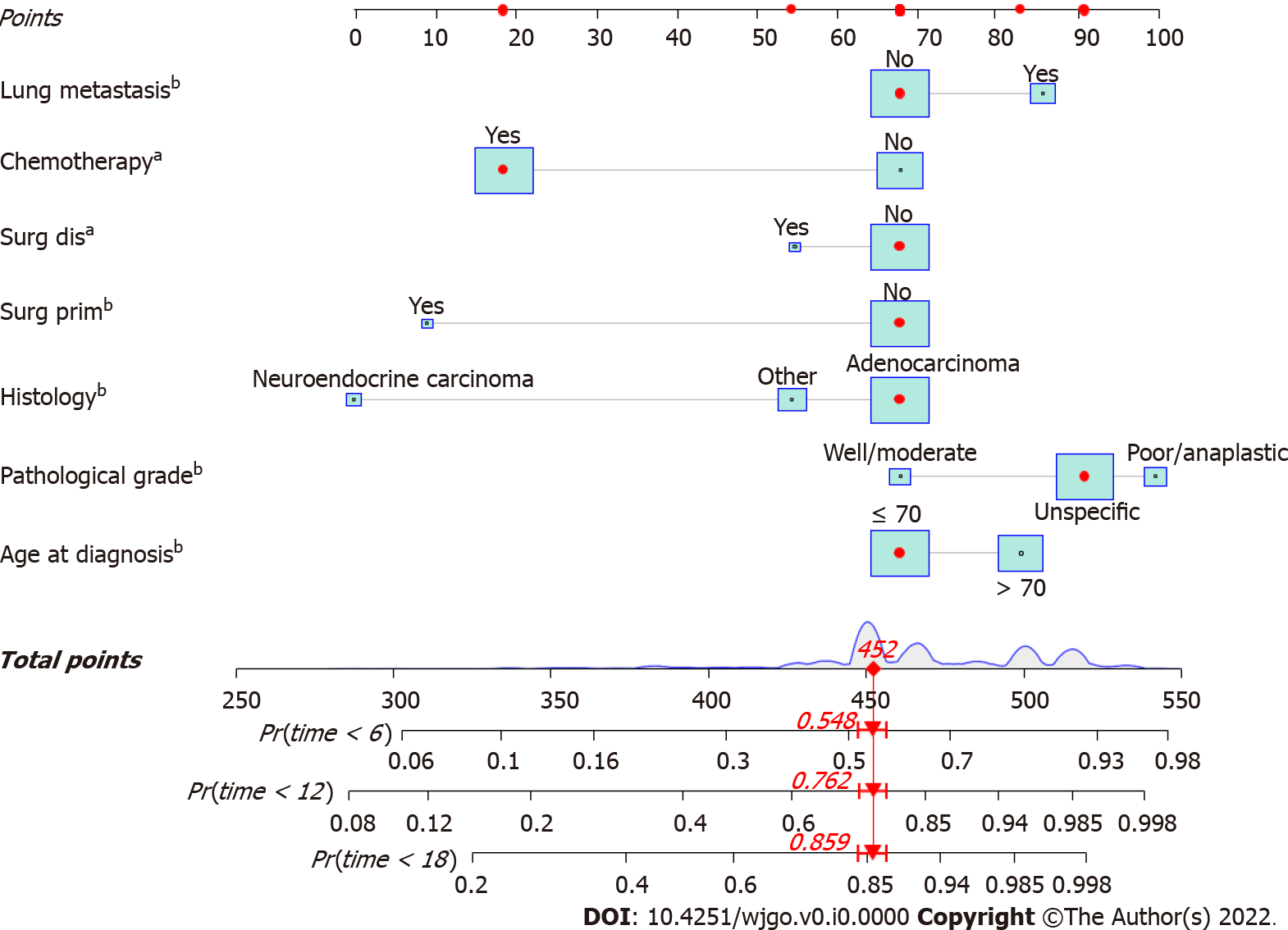
**Figure 1 Nomogram for predicting liver metastasis from pancreatic cancer patients.** a*P* < 0.001.LN: Lymph node; LM: Liver metastasis.



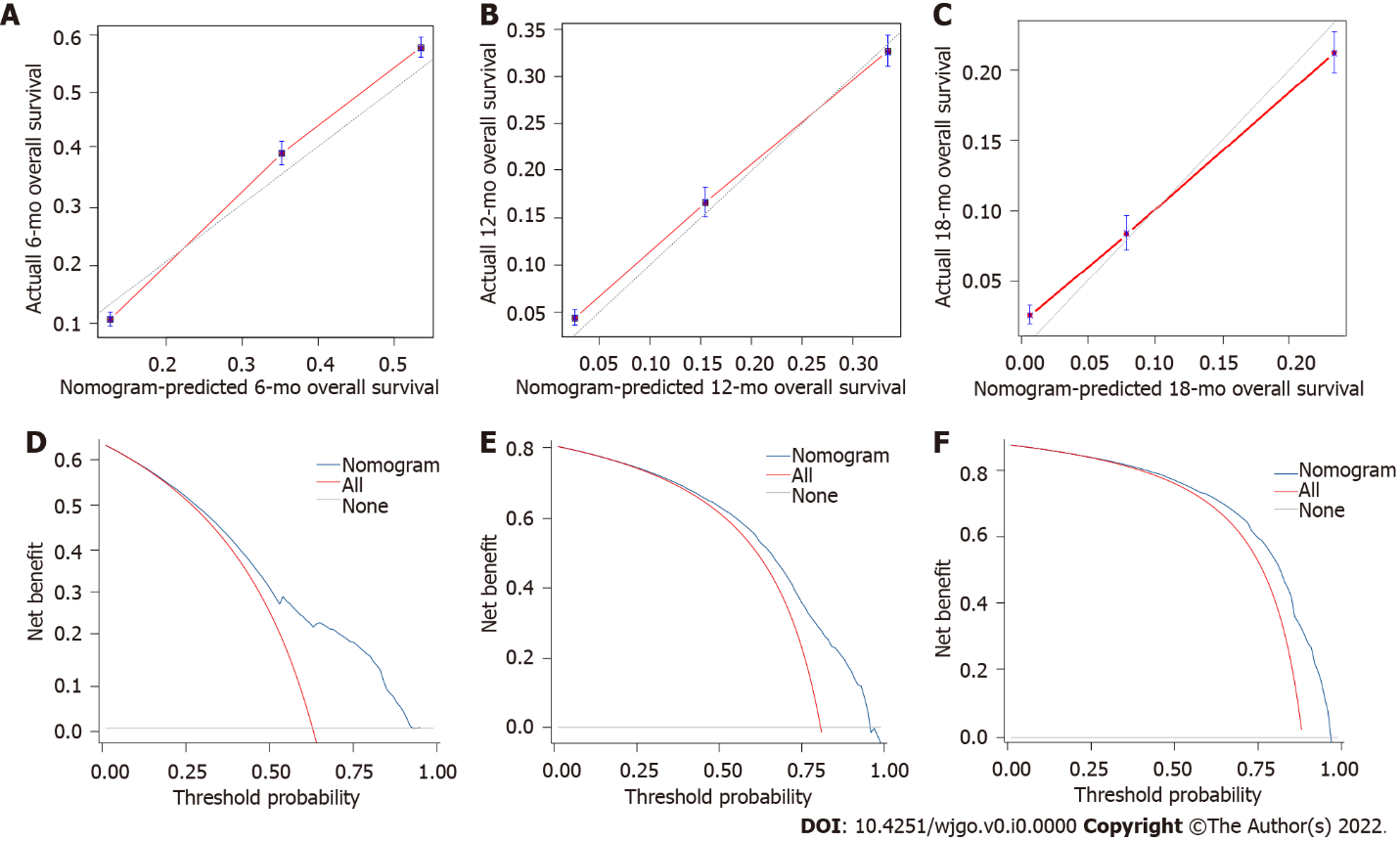
**Figure 2 Validation of the diagnostic nomogram in the training and validation sets.** A: The receiver operating characteristic curve of the training set; B: The calibration curve of the training set; C: The decision curve analysis of the training set; D: The receiver operating characteristic curve of the validation set; E: The calibration curve of the validation set; F: The decision curve analysis of the validation set. AUC: Area under curve.



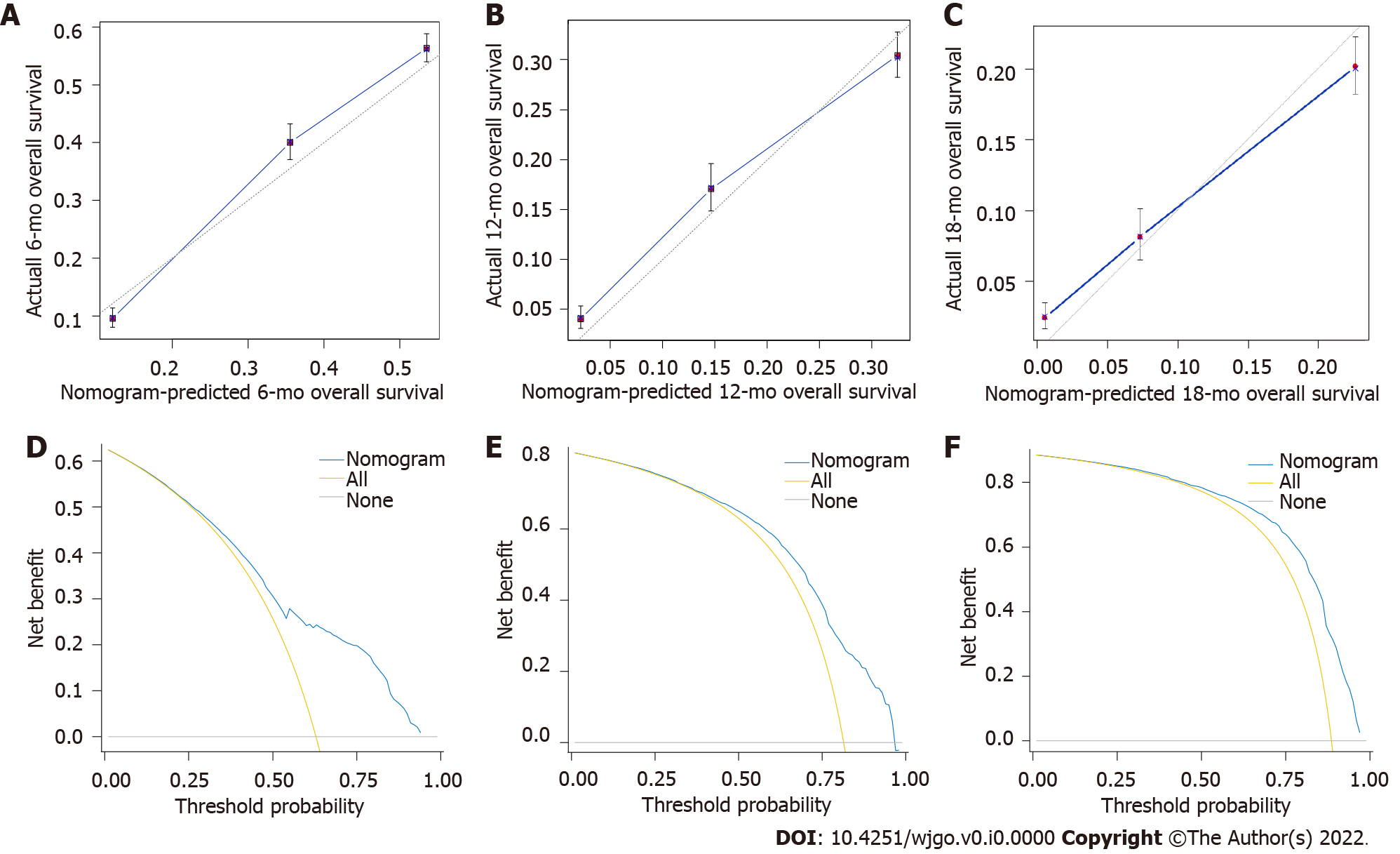
**Figure 3 The** **least absolute shrinkage and selection operator regression used to select prognostic factors for** **overall survival.** A: Least absolute shrinkage and selection operator (LASSO) coefficient profiles of 16 variables for overall survival (OS); B: LASSO Cox analysis identified 7 variables for OS. The LASSO regression analysis run in R runs 10 times K cross-validation for centralization and normalization of included variables and then selects the most appropriate lambda value depending on the type measure of -2 Log-likelihood and binomial family. “Lambda.lse” gives a model with good performance but the least number of independent variables.



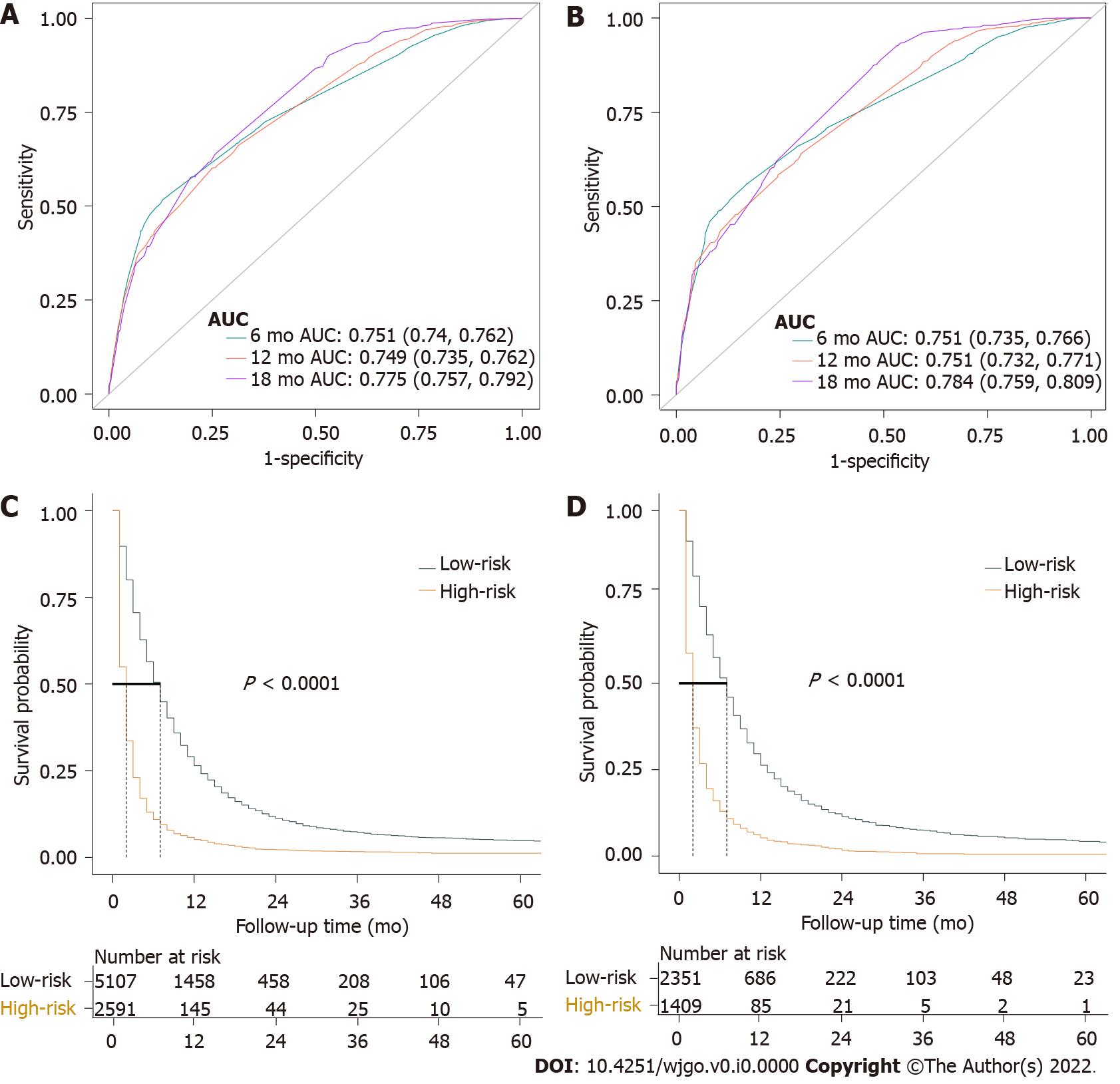
**Figure 4 A prognostic nomogram for pancreatic cancer patients with liver metastasis.** a*P* < 0.01; b*P* < 0.001. Surg prim: Surgical treatments of the primary site; Surg dis: Surgical treatments of the distant site.



**Figure 5 The calibration curves and decision curve analysis of the prognostic nomogram in the training set.** A: The calibration curve of the nomogram for 6 mo in the training set; B: The calibration curve of the nomogram for 12 mo in the training set; C: The calibration curve of the nomogram for 18 mo in the training set; D: The decision curve analysis of the nomogram for 6 mo in the training set; E: The decision curve analysis of the nomogram for 12 mo in the training set; F: The decision curve analysis of the nomogram for 18 mo in the training set.



**Figure 6 The calibration curves and decision curve analysis of the prognostic nomogram in the validation set.** A: The calibration curve of the nomogram for 6 mo in the validation set; B: The calibration curve of the nomogram for 12 mo in the validation set; C: The calibration curve of the nomogram for 18 mo in the validation set; D: The decision curve analysis of the nomogram for 6 mo in the validation set; E: The decision curve analysis of the nomogram for 12 mo in the validation set; F: The decision curve analysis of the nomogram for 18 mo in the validation set.



**Figure 7 Time-dependent receiver operating characteristic curve analysis and Kaplan-Meier survival curves of prognostic nomogram.** A: Time-dependent receiver operating characteristic curve of the prognostic nomogram for 6, 12, and 18 mo in the training set; B: Time-dependent receiver operating characteristic curve of the prognostic nomogram for 6, 12, and 18 mo in the validation set; C: The Kaplan-Meier survival curves of the patients in the training set; D: The Kaplan-Meier survival curves of the patients in the validation set. AUC: Area under curve.

**Table 1 Clinical characteristics for patients diagnosed with pancreatic cancer with and without liver metastasis in Surveillance, Epidemiology, and End Result database (2010-2015), *n* (%)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **Levels** | **Total (*n* = 33459)** | **With LM (*n* = 11458** | **Without LM (*n* = 22001)** | ***P* value1** |
| Age at diagnosis (yr) | ≤ 70 | 20680 (61.8) | 7552 (65.9) | 13128 (59.7) | < 0.001 |
| > 70 | 12779 (38.2) | 3906 (34.1) | 8873 (40.3) |
| Sex | Female | 16390 (49.0) | 5195 (45.3) | 11195 (50.9) | < 0.001 |
| Male | 17069 (51.0) | 6263 (54.7) | 10806 (49.1) |
| Race | White | 26459 (79.1) | 9048 (79) | 17411 (79.1) | < 0.001 |
| Black | 4139 (12.4) | 1533 (13.4) | 2606 (11.8) |
| Other | 2861 (8.6) | 877 (7.7) | 1984 (9) |
| Primary site | Head | 17950 (53.6) | 4441 (38.8) | 13509 (61.4) | < 0.001 |
| Body/tail | 10200 (30.5) | 4762 (41.5) | 5438 (24.7) |
| Other | 5309 (15.9) | 2255 (19.7) | 3054 (13.9) |
| Histology | Adenocarcinoma | 22948 (68.6) | 8746 (76.3) | 14202 (64.6) | < 0.001 |
| Neuroendocrine carcinoma | 1585 (4.7) | 615 (5.4) | 970 (4.4) |
| Other | 8926 (26.7) | 2097 (18.3) | 6829 (31) |
| Pathological grade | Well/moderate | 8293 (24.8) | 1185 (10.3) | 7108 (32.3) | < 0.001 |
| Poor/anaplastic | 5105 (15.3) | 1367 (11.9) | 3738 (17) |
| Unspecific | 20061 (60.0) | 8906 (77.7) | 11155 (50.7) |
| T stage | T1 | 2063 (6.2) | 362 (3.2) | 1701 (7.7) | < 0.001 |
| T2 | 8383 (25.1) | 3934 (34.3) | 4449 (20.2) |
| T3 | 14526 (43.4) | 3541 (30.9) | 10985 (49.9) |
| T4 | 6635 (19.8) | 2116 (18.5) | 4519 (20.5) |
| TX | 1852 (5.5) | 1505 (13.1) | 347 (1.6) |
| LN metastasis | No | 18469 (55.2) | 6044 (52.7) | 12425 (56.5) | < 0.001 |
| Yes | 14990 (44.8) | 5414 (47.3) | 9576 (43.5) |
| Tumor size (cm) | < 2 | 3617 (10.8) | 664 (5.8) | 2953 (13.4) | < 0.001 |
| 2-4 | 14644 (43.8) | 4338 (37.9) | 10306 (46.8) |
| > 4 | 15198 (45.4) | 6456 (56.3) | 8742 (39.7) |
| Surg prim | No | 24100 (72.0) | 11054 (96.5) | 13046 (59.3) | < 0.001 |
| Yes | 9359 (28.0) | 404 (3.5) | 8955 (40.7) |
| Surg dis | No | 32408 (96.9) | 11102 (96.9) | 21306 (96.8) | 0.822 |
| Yes | 1051 (3.1) | 356 (3.1) | 695 (3.2) |
| Radiation | No | 27483 (82.1) | 10909 (95.2) | 16574 (75.3) | < 0.001 |
| Yes | 5976 (17.9) | 549 (4.8) | 5427 (24.7) |
| Chemotherapy | No | 13156 (39.3) | 4298 (37.5) | 8858 (40.3) | < 0.001 |
| Yes | 20303 (60.7) | 7160 (62.5) | 13143 (59.7) |
| Bone metastasis | No | 32524 (97.2) | 10812 (94.4) | 21712 (98.7) | < 0.001 |
| Yes | 935 (2.8) | 646 (5.6) | 289 (1.3) |
| Brain metastasis | No | 33365 (99.7) | 11405 (99.5) | 21960 (99.8) | < 0.001 |
| Yes | 94 (0.3) | 53 (0.5) | 41 (0.2) |
| Lung metastasis | No | 30546 (91.3) | 9701 (84.7) | 20845 (94.7) | < 0.001 |
| Yes | 2913 (8.7) | 1757 (15.3) | 1156 (5.3) |
| Vital status | Alive | 6025 (18.0) | 804 (7) | 5221 (23.7) | < 0.001 |
| Dead | 27434 (82.0) | 10654 (93) | 16780 (76.3) |

1 Statistical differences between pancreatic cancer with liver metastasis (LM) and without LM groups.

LN: Lymph node; LM: Liver metastasis; Surg prim: Surgical treatments of the primary site; Surg dis: Surgical treatments of the distant site.

**Table 2 Univariate and multivariate logistic analyses of liver metastasis in pancreatic cancer patients, *n* (%)**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **Levels** | **Without LM (*n* = 14608)** | **With LM (*n* = 7698)** | **Univariate analysis (OR, 95%CI)** | **Multivariate analysis (OR, 95%CI)** |
| Age at diagnosis (yr) | ≤ 70 | 8783 (60.1) | 5032 (65.4) |  |  |
| > 70 | 5825 (39.9) | 2666 (34.6) | 0.80 (0.75-0.85, *P* < 0.001) | 0.72 (0.68-0.77, *P* < 0.001) |
| Sex | Female | 7393 (50.6) | 3485 (45.3) |  |  |
| Male | 7215 (49.4) | 4213 (54.7) | 1.24 (1.17-1.31, *P* < 0.001) | 1.16 (0.89-1.23, *P* = 0.056) |
| Race | White | 11590 (79.3) | 6048 (78.6) |  |  |
| Black | 1730 (11.8) | 1047 (13.6) | 1.16 (1.07-1.26, *P* < 0.001) | 1.07 (0.97-1.17, *P* = 0.186) |
| Other | 1288 (8.8) | 603 (7.8) | 0.90 (0.81-0.99, *P* = 0.036) | 0.93 (0.83-1.04, *P* = 0.228) |
| Primary site | Head | 8939 (61.2) | 3003 (39) |  |  |
| Body/tail | 3604 (24.7) | 3212 (41.7) | 2.65 (2.49-2.83, *P* < 0.001) | 2.30 (2.14-2.47, *P* < 0.001) |
| Other | 2065 (14.1) | 1483 (19.3) | 2.14 (1.98-2.31, *P* < 0.001) | 1.76 (1.61-1.92, *P* < 0.001) |
| Histology | Adenocarcinoma | 9394 (64.3) | 5867 (76.2) |  |  |
| Neuroendocrine carcinoma | 640 (4.4) | 411 (5.3) | 1.03 (0.90-1.17, *P* = 0.670) | 1.28 (1.10-1.49, *P* = 0.001) |
| Other | 4574 (31.3) | 1420 (18.4) | 0.50 (0.46-0.53, *P* < 0.001) | 0.61 (0.56-0.66, *P* < 0.001) |
| Pathological grade | Well/moderate | 4727 (32.4) | 799 (10.4) |  |  |
| Poor/anaplastic | 2476 (16.9) | 919 (11.9) | 2.20 (1.97-2.44, *P* < 0.001) | 2.26 (2.01-2.53, *P* < 0.001) |
| Unspecific | 7405 (50.7) | 5980 (77.7) | 4.78 (4.40-5.19, *P* < 0.001) | 4.23 (3.86-4.64, *P* < 0.001) |
| T stage | T1 | 1131 (7.7) | 245 (3.2) |  |  |
| T2 | 2901 (19.9) | 2653 (34.5) | 4.22 (3.65-4.90, *P* < 0.001) | 1.53 (0.89-1.86, *P* = 0.080) |
| T3 | 7325 (50.1) | 2402 (31.2) | 1.51 (1.31-1.75, *P* < 0.001) | 0.81 (0.64-1.02, *P* = 0.054) |
| T4 | 3019 (20.7) | 1411 (18.3) | 2.16 (1.86-2.52, *P* < 0.001) | 0.72 (0.56-1.01, *P* = 0.051) |
| TX | 232 (1.6) | 987 (12.8) | 9.4 (6.4-9.7, *P* < 0.001) | 7.18 (5.52-9.35, *P* < 0.001) |
| LN metastasis | No | 8213 (56.2) | 4046 (52.6) |  |  |
| Yes | 6395 (43.8) | 3652 (47.4) | 1.16 (1.10-1.23, *P* < 0.001) | 1.38 (1.30-1.47, *P* < 0.001) |
| Tumor size (cm) | < 2 | 1966 (13.5) | 460 (6) |  |  |
| 2-4 | 6800 (46.5) | 2939 (38.2) | 1.85 (1.66-2.06, *P* < 0.001) | 1.29 (1.08-1.55, *P* = 0.006) |
| > 4 | 5842 (40) | 4299 (55.8) | 3.15 (2.82-3.51, *P* < 0.001) | 2.05 (1.71-2.46, *P* < 0.001) |

OR: Odds ratio; CI: Confidence interval; LN: Lymph node; LM: Liver metastasis.

**Table 3 Multivariate Cox analysis of the training cohort based on the results of least absolute shrinkage and selection operator regression, *n* (%)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Levels** | **All** | **Multivariate analysis (HR, 95%CI)** |
| Age at diagnosis (yr) | ≤ 70 | 4812 (62.5) |  |
| > 70 | 2886 (37.5) | 1.29 (1.23-1.36, *P* < 0.001) |
| Histology | Adenocarcinoma | 5867 (76.2) |  |
| Neuroendocrine carcinoma | 411 (5.3) | 0.31 (0.28-0.35, *P* < 0.001) |
| Other | 1420 (18.4) | 0.79 (0.75-0.85, *P* < 0.001) |
| Pathological grade | Well/moderate | 799 (10.4) |  |
| Poor/anaplastic | 919 (11.9) | 1.72 (1.55-1.91, *P* < 0.001) |
| Unspecific | 5980 (77.7) | 1.48 (1.36-1.61, *P* < 0.001) |
| Surg prim | No | 7429 (96.5) |  |
| Yes | 269 (3.5) | 0.37 (0.31-0.43, *P* < 0.001) |
| Surg dis | No | 7450 (96.8) |  |
| Yes | 248 (3.2) | 0.80 (0.69-0.93, *P* = 0.004) |
| Chemotherapy | No | 2934 (38.1) |  |
| Yes | 4764 (61.9) | 0.43 (0.41-0.45, *P* < 0.001) |
| Lung metastasis | No | 6510 (84.6) |  |
| Yes | 1188 (15.4) | 1.35 (1.27-1.44, *P* < 0.001) |

HR: Hazard ratio; CI: Confidence interval; Surg prim: Surgical treatments of the primary site; Surg dis: Surgical treatments of the distant site.