Gut and Liver
https://doi.org/10.5009/gnl20306
pISSN 1976-2283 eISSN 2005-1212

Original Article

Development and External Validation of Survival Prediction Model for Pancreatic Cancer Using Two Nationwide Databases: Surveillance, Epidemiology and End Results (SEER) and Korea Tumor Registry System-Biliary Pancreas (KOTUS-BP)

Jae Seung Kang[1], Lydia Mok[2], Jin Seok Heo[3], In Woong Han[3], Sang Hyun Shin[3], Yoo-Seok Yoon[4], Ho-Seong Han[4], Dae Wook Hwang[5], Jae Hoon Lee[5], Woo Jung Lee[6], Sang Jae Park[7], Joon Seong Park[8], Yonghoon Kim[9], Huisong Lee[10], Young-Dong Yu[11], Jae Do Yang[12], Seung Eun Lee[13], Il Young Park[14], Chi-Young Jeong[15], Younghoon Roh[16], Seong-Ryong Kim[17], Ju Ik Moon[18], Sang Kuon Lee[19], Hee Joon Kim[20], Seungyeoun Lee[21], Hongbeom Kim[22], Wooil Kwon[22], Chang-Sup Lim[1], Jin-Young Jang[22], and Taesung Park[2]

[1]Department of Surgery, Seoul Metropolitan Government Seoul National University Boramae Medical Center, [2]Department of Statistics and Interdisciplinary Program in Bioinformatics, Seoul National University, [3]Department of Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, [4]Department of Surgery, Seoul National University Bundang Hospital, Seoul National University College of Medicine, [5]Department of Surgery, Asan Medical Center, University of Ulsan College of Medicine, [6]Division of Hepatobiliary and Pancreatic Surgery, Department of Surgery, Yonsei University College of Medicine, Seoul, [7]Center for Liver and Pancreatobiliary Cancer, National Cancer Center, Goyang, [8]Pancreatobiliary Cancer Clinic, Department of Surgery, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, [9]Department of Surgery, Keimyung University Dongsan Medical Center, Keimyung University School of Medicine, Daegu, [10]Department of Surgery, Ewha Womans University Mokdong Hospital, Ewha Womans University School of Medicine, [11]Division of HBP Surgery and Liver Transplantation, Department of Surgery, Korea University College of Medicine, Seoul, [12]Department of Surgery, Jeonbuk National University Medical School, Jeonju, [13]Department of Surgery, Chung-Ang University Hospital, Chung-Ang University College of Medicine, Seoul, [14]Department of General Surgery, Bucheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Bucheon, [15]Department of Surgery, Gyeongsang National University Hospital, Gyeongsang National University School of Medicine, Jinju, [16]Department of Surgery, Dong-A University College of Medicine, Busan, [17]Department of Surgery, Dongguk University Ilsan Hospital, Dongguk University College of Medicine, Goyang, [18]Department of Surgery, Konyang University Hospital, [19]Department of Surgery, Daejeon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Daejeon, [20]Department of Surgery, Chonnam National University Hospital, Gwangju, [21]Department of Mathematics and Statistics, Sejong University, and [22]Department of Surgery and Cancer Research Institute, Seoul National University College of Medicine, Seoul, Korea

See editorial on page 797.

Article Info

Received September 29, 2020
Revised December 31, 2020
Accepted January 15, 2021

Corresponding Author

Chang-Sup Lim
ORCID https://orcid.org/0000-0002-2349-9647
E-mail limcs7@gmail.com

Jin-Young Jang
ORCID https://orcid.org/0000-0003-3312-0503
E-mail jangjy4@snu.ac.kr

Taesung Park
ORCID https://orcid.org/0000-0002-8294-590X
E-mail tspark@stats.snu.ac.kr

Jae Seung Kang, Lydia Mok, and Jin Seok Heo contributed equally to this work as first authors.

Background/Aims: Several prediction models for evaluating the prognosis of nonmetastatic resected pancreatic ductal adenocarcinoma (PDAC) have been developed, and their performances were reported to be superior to that of the 8th edition of the American Joint Committee on Cancer (AJCC) staging system. We developed a prediction model to evaluate the prognosis of resected PDAC and externally validated it with data from a nationwide Korean database.

Methods: Data from the Surveillance, Epidemiology and End Results (SEER) database were utilized for model development, and data from the Korea Tumor Registry System-Biliary Pancreas (KOTUS-BP) database were used for external validation. Potential candidate variables for model development were age, sex, histologic differentiation, tumor location, adjuvant chemotherapy, and the AJCC 8th staging system T and N stages. For external validation, the concordance index (C-index) and time-dependent area under the receiver operating characteristic curve (AUC) were evaluated.

Results: Between 2004 and 2016, data from 9,624 patients were utilized for model development, and data from 3,282 patients were used for external validation. In the multivariate Cox proportional hazard model, age, sex, tumor location, T and N stages, histologic differentiation, and adjuvant chemotherapy were independent prognostic factors for resected PDAC. After an exhaustive search and 10-fold cross validation, the best model was finally developed, which included all prognostic variables. The C-index, 1-year, 2-year, 3-year, and 5-year time-dependent AUCs were 0.628, 0.650, 0.665, 0.675, and 0.686, respectively.

Conclusions: The survival prediction model for resected PDAC could provide quantitative survival probabilities with reliable performance. External validation studies with other nationwide databases are needed to evaluate the performance of this model. **(Gut Liver 2021;15:912-921)**

Key Words: Pancreatic neoplasms; Survival; Prognosis

INTRODUCTION

Despite the development of surgical technique and perioperative treatment, pancreatic cancer is still a lethal disease worldwide. Although 5-year overall survival (OS) of pancreatic cancer gradually increased in Korea, it was 11.4% in 2016, which was low compared with other gastrointestinal malignancies.[1] However, in patients who underwent pancreatectomy for nonmetastatic pancreatic adenocarcinoma (PDAC), 5-year OS rate was 20.2%, varying according to the 8th edition of the American Joint Committee on Cancer (AJCC) staging of pancreatic cancer, from 38.2% of stage IA to 0.0% of stage III.[2,3] Therefore, the prognosis of resected PDAC differs according to the disease severity, and the accurate prediction of survival probabilities is required to provide the information to the patients about their prognosis and the postoperative individualized treatment plan.

Although the AJCC staging system has been widely used, with many surgeons depending on it for predicting the prognosis, its prognostic accuracy was insufficient with concordance index (C-index) of 0.57 and 0.60 in two large-scaled external validation studies.[3,4] There were some models for predicting prognosis after surgery in patients with resected PDAC,[5-8] consisting of the AJCC T and N stage, along with other variables, such as age, sex, histologic differentiation, adjuvant treatment, or resection margin status. The formulated prediction models had higher performance power than that of the AJCC staging system, and provided the quantitative survival probability after the surgery.[6-8]

The aim of this study was to develop a prediction model for providing survival probability for resected PDAC with data from a Surveillance, Epidemiology and End Results (SEER) database in the United States and externally validate this model with nationwide Korean database.

MATERIALS AND METHODS

This was a retrospective cohort study with prospectively collected data in the United States and Republic of Korea. The institutional review board of each participating center in Korea approved this study (representative IRB number: 07-2020-058 in Seoul Metropolitan Government Seoul National University Boramae Medical Center, Seoul, Korea) and all patients provided consent to participate in this study.

1. Study population

Fig. 1 is the flowchart of study design for defining the final model development set and test set, including the inclusion and exclusion criteria. For unification of the study period, patients who underwent upfront curative-intent pancreatectomy between 2004 and 2016 were enrolled in this study. Data from a SEER database in National Cancer Institute were utilized for the development of survival prediction models of resected PDAC. The patient's inclusion criteria were: (1) a confirmed primary site in the pancreas (C250, C251, C252); and (2) ductal adenocarcinoma (SEER

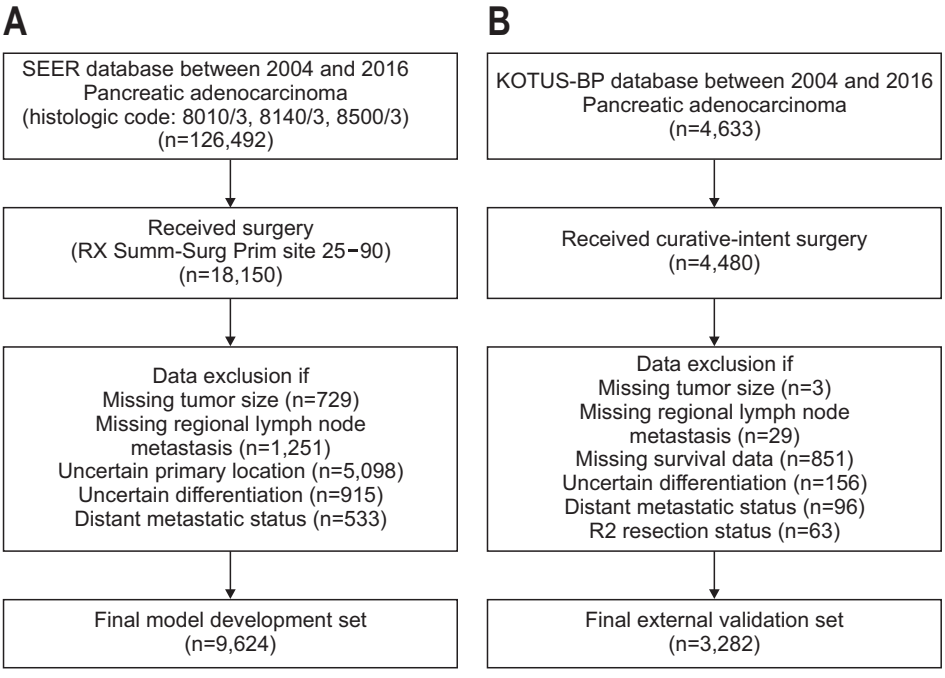

Fig. 1. The case selection criteria for defining a model development set and an external validation set.
SEER, Surveillance, Epidemiology, and End Results; KOTUS-BP, Korea Tumor Registry System-Biliary Pancreas.

histologic code: 8010/3, 8140/3, 8500/3). The detailed inclusion and exclusion codes of SEER are described in Supplementary Table 1. After exclusion, a total of 9,624 patients who underwent pancreatectomy due to PDAC were finally included in the model development set (Fig. 1A). Clinical variables derived from the SEER database were age, sex, histologic differentiation (well differentiated, moderately differentiated, and poorly differentiated), tumor location (head, body, or tail), adjuvant chemotherapy, and the 8th edition of AJCC T and N stage.[2]

A Korea Tumor Registry System-Biliary Pancreas (KOTUS-BP) database was established and launched by the Korean Association of Hepato-Biliary-Pancreatic Surgery in 2014. After exclusion, a total of 3,282 patients in 20 institutions in Korea were finally included in the external validation set (Fig. 1B). Clinical variables analyzed were same as the model development set variables.

2. Model development and external validation

Fig. 2 shows the overall scheme of model development and external validation. We exhaustively searched for the best models for each number of variables by comparing the Akaike information criterion (AIC), Harrell C-index, and the 2-year time-dependent area under the receiver operating characteristic curve (AUC) (Step 1). AIC estimates the prediction error and penalized for the number of variables in the model and lower value is better.[9] C-index calculates the concordance of the predicted and the observed survival time.[10] The closer the value of C-index is to 1, the higher the matching rate. Time-dependent AUC assesses the predictive accuracy of the survival model.[11] For survival data, cumulative sensitivity, and dynamic specificity are used to get time-dependent AUC. The closer the value is to 1, the better the predictive ability of the corresponding time cutoff value. For all possible combination of variables, we selected the models with the highest C-index, the highest AUC, or equivalently the lowest AIC for each number of variables. From the models selected from the previous step, we finally determined the single best model using 10-fold cross validation (CV) (Step 2). The best model had the highest 10-fold CV C-index or 10-fold CV 2-year time-dependent AUC.

After the determination of the best model, the external validation was performed with the external validation set

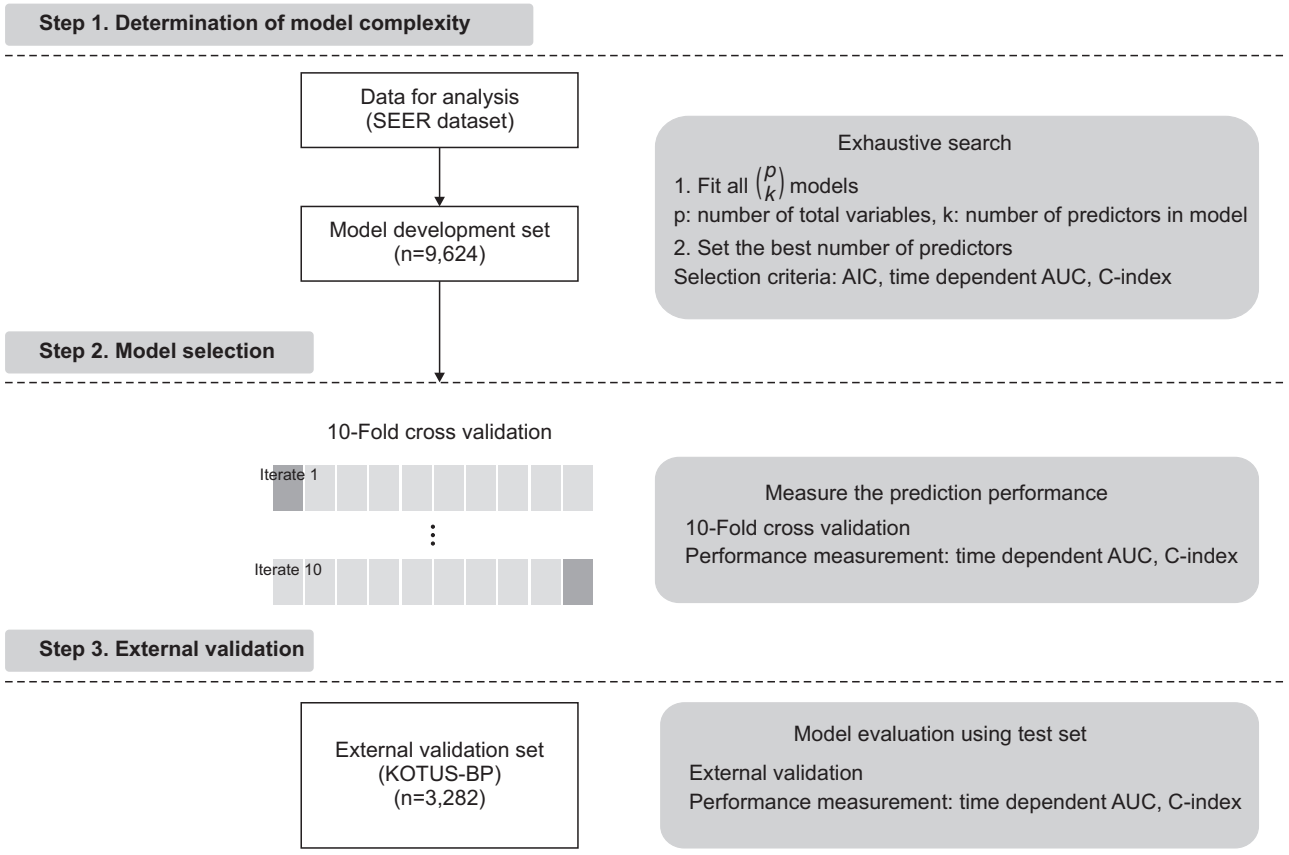

Fig. 2. Flowchart of model development and external validation process.
SEER, Surveillance, Epidemiology and End Results; KOTUS-BP, Korea Tumor Registry System-Biliary Pancreas; AIC, Akaike information criterion; AUC, area under the receiver operating characteristic curve; C-index, Harrell concordance index.

(KOTUS-BP). C-index along with 1-year, 2-year, 3-year, and 5-year time-dependent AUC were calculated.

3. Statistical analysis

All statistical analyses used R program version 3.6.2 (The R Foundation for Statistical Computing, Vienna, Austria). Survival outcomes were calculated using the Kaplan-Meier method and compared using the log-rank test. Variables for which p-value <0.05 in univariate analysis were entered into a multivariate Cox proportional hazard (PH) model to estimate the hazard ratios (HRs) for the corresponding predictors. Continuous variables were reported as the mean±standard deviation. A calibration plot was used to compare the predicted probability with the observed probability at a specific time point. If the model is ideal, pairs of observed and the predicted probabilities lie on the 45° angle line. To assess calibration for the prognostic model, the Greenwood-Nam-D'Agostino goodness of fit test was performed in each time cutoff value.[12] The nomogram plot was produced using "rms" packages in R program and the calibration plot was produced using "pec" packages.

RESULTS

1. Summary of 5-year OS rates according to the clinical variables in SEER and KOTUS-BP

Table 1 shows the 5-year OS rates according to the clinical variables in SEER and KOTUS-BP database, respectively. Overall, 5-year OS rate in SEER was 20.1% and median survival duration was 21 months. Five-year OS rate in KOTUS-BP was 32.1% and median survival duration was 24 months. Sex, tumor location, AJCC 8th T and N stage, histologic differentiation, and adjuvant chemotherapy could discriminate 5-year OS rates with statistical significance in the univariate analysis in both SEER and KOTUS-BP databases.

2. Potential variables for model development in SEER database

In the multivariate Cox-PH model, age (HR, 1.007; 95% confidence interval [CI], 1.004 to 1.009; p<0.001), male sex (HR, 1.070; 95% CI, 1.019 to 1.123; p=0.006), head cancer (HR, 1.12; 95% CI, 1.041 to 1.198; p=0.002), AJCC 8th T stage (T2: HR, 1.394; 95% CI, 1.297 to 1.498; p<0.001 and T3: HR, 1.723; 95% CI, 1.586 to 1.872; p<0.001), N stage (N1: HR, 1.566; 95% CI, 1.476 to 1.663; p<0.001 and N2:

Table 1. Five-Year Overall Survival Rates According to the Variables in the SEER Database and KOTUS-BP

Variable	SEER database (n=9,624)			KOTUS-BP (n=3,282)		
	Patients	5-Year OS, %	p-value*	Patients	5-Year OS, %	p-value*
Age, yr	65.6±10.5	20.1†		63.9±10.1	32.1†	
Sex			<0.001			0.007
Female	4,755 (49.4)	21.3		1,381 (42.1)	36.1	
Male	4,869 (50.6)	18.9		1,901 (57.9)	29.2	
Tumor location			0.002			<0.001
Head	8,079 (83.9)	19.2		2,046 (62.3)	28.4	
Body/tail	1,545 (16.1)	25.0		1,236 (37.7)	37.7	
AJCC 8th T stage						
T1	1,603 (16.7)	32.7		671 (20.5)	45.1	
T2	5,830 (60.6)	18.8	<0.001	2,009 (61.2)	29.6	<0.001
T3	2,191 (22.7)	14.3	<0.001	602 (18.3)	24.5	<0.001
AJCC 8th N stage						
N0	3,155 (32.8)	32.4		1,312 (40.0)	42.5	
N1	4,030 (41.9)	16.8	<0.001	2,363 (72.0)	28.4	<0.001
N2	2,439 (25.3)	9.6	<0.001	543 (16.5)	16.4	<0.001
Differentiation						
Well	1,013 (10.5)	37.4		376 (11.5)	44.9	
Moderately	5,055 (52.5)	20.5	<0.001	2,363 (72.0)	32.9	<0.001
Poorly	3,556 (37.0)	14.6	<0.001	543 (16.5)	20.5	<0.001
Adjuvant chemotherapy			<0.001			0.004
Yes	2,948 (30.6)	21.3		2,008 (61.2)	36.0	
No	6,676 (69.4)	17.3		1,274 (38.8)	29.4	

Data are presented as mean±SD or number (%).
OS, overall survival; SEER, Surveillance, Epidemiology and End Results; KOTUS-BP, Korea Tumor Registry System-Biliary Pancreas; AJCC, American Joint Committee on Cancer.
*Log-rank test.

HR, 1.980; 95% CI, 1.852 to 2.116; p<0.001), histologic differentiation (moderately differentiated: HR, 1.565; 95% CI, 1.428 to 1.715; p<0.001 and poorly differentiated: HR, 2.069; 95% CI, 1.884 to 2.272; p<0.001), and no adjuvant chemotherapy (HR, 1.789; 95% CI, 1.696 to 1.887; p<0.001) were independent prognostic factors for worse outcome in patients with resected PDAC (Table 2).

3. Model development with SEER database and external validation with KOTUS-BP database

After the exhaustive search, the best combination of variables with the lowest AIC, the highest C-index and the highest 2-year time-dependent AUC were models including all potential variables (Fig. 2, Step 1). For the models with all variables, the single best model was fitted and determined using 10-fold CV. The C-index, 1-year, 2-year, and 3-year time-dependent AUCs of the best model after the 10-fold CV were 0.654, 0.712, 0.689, and 0.694, respectively (Fig. 2, Step 2). This model was visualized to nomogram form (Fig. 3).

Then, the external validation was performed with KOTUS-BP database (Fig. 2, Step 3). The C-index, 1-year, 2-year, 3-year, and 5-year time-dependent AUCs were 0.628, 0.650, 0.665, 0.675, and 0.686, respectively (see Supplementary Fig. 1). The p-value of Greenwood-Nam-D'Agostino test for each time cutoff value (1-, 2-, 3-, and 5-year) was 0.78, 0.18, 0.39, and 0.17, respectively, indicating the developed model to be well calibrated (Fig. 4).

DISCUSSION

This study conducted the model development with SEER database by the rigorous statistical techniques and demonstrated the reliable performance (C-index of 0.628) in the external validation with nationwide database in Korea (KOTUS-BP). Age, sex, histologic differentiation, AJCC 8th T and N stage, tumor location, and adjuvant chemotherapy were included in this prediction model, with a few of them already included in other prognostic models.[6-8] The visualized nomogram was established, and time-dependent survival probability could be easily calculated (Fig. 3). Calibration demonstrated good consistency between the predictive survival and actual observation of 1-year, 2-year, 3-year, and 5-year survival rates (Fig. 4).

The development of nomogram is already popular among clinicians for a variety of diseases as they provide quantitative information by simple calculation, helping clinicians decide the customized treatment strategies.[13-17] The current AJCC tumor-node-metastasis staging system discriminate patients with limited variables, and the C-index of 8th edition of AJCC pancreatic cancer stage was 0.588, which was lower than that of the present model (0.628). However, variables other than tumor size or lymph nodes,

Table 2. Variables for Model Development in the Multivariate Cox Proportional Hazard Model

Variable	SEER database (n=9,624)		
	Hazard ratio	95% CI	p-value
Age	1.007	1.004–1.009	<0.001
Sex			
Female	Reference	-	-
Male	1.070	1.019–1.123	0.006
Tumor location			
Body/tail	Reference	-	-
Head	1.12	1.041–1.198	0.002
AJCC 8th T stage			
T1	Reference	-	-
T2	1.394	1.297–1.498	<0.001
T3	1.723	1.586–1.872	<0.001
AJCC 8th N stage			
N0	Reference	-	-
N1	1.566	1.476–1.663	<0.001
N2	1.980	1.852–2.116	<0.001
Differentiation			
Well	Reference	-	-
Moderately	1.565	1.428–1.715	<0.001
Poorly	2.069	1.884–2.272	<0.001
Adjuvant chemotherapy			
Yes	Reference	-	-
No	1.789	1.696–1.887	<0.001

SEER, Surveillance, Epidemiology and End Results; CI, confidence interval; AJCC, American Joint Committee on Cancer.

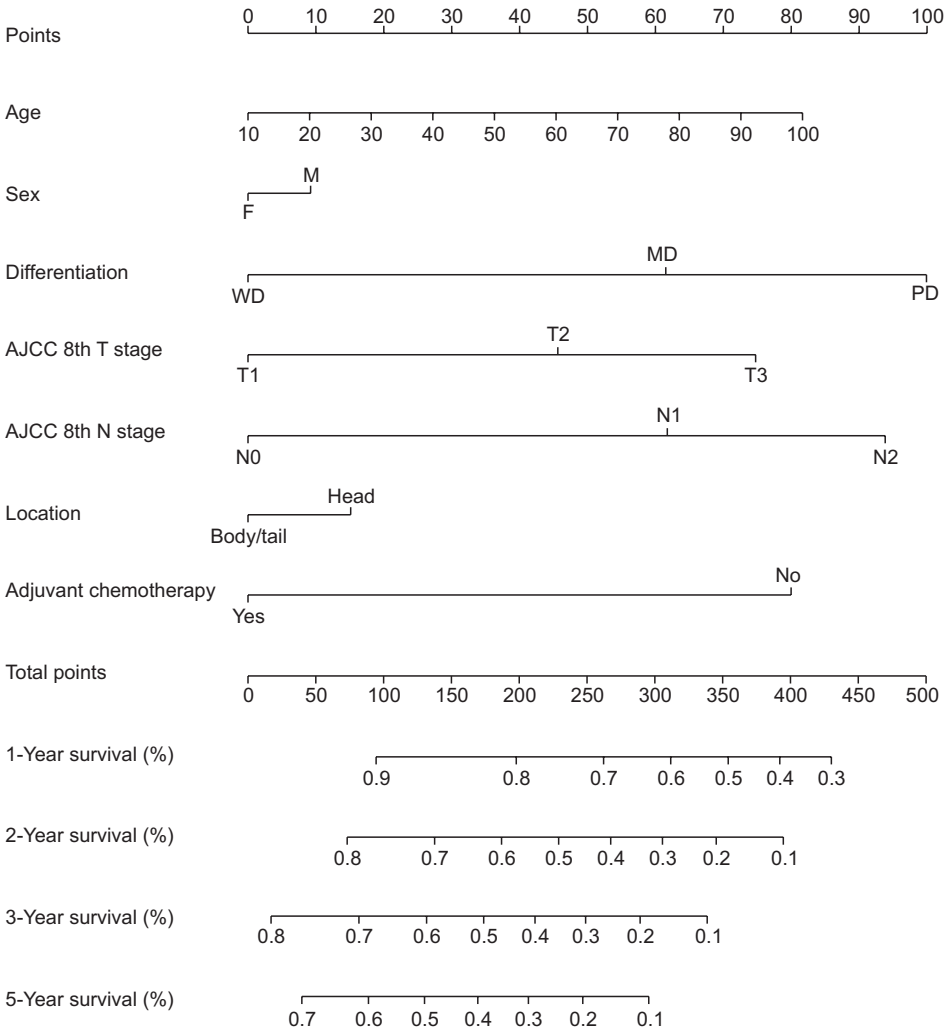

Fig. 3. Nomogram for survival in patients with resected pancreatic ductal adenocarcinoma. Prediction of survival can be made by drawing a vertical line from the total points scale to the survival probabilities scale.
M, male; F, female; WD, well differentiated; MD, moderately differentiated; PD, poorly differentiated; AJCC, American Joint Committee on Cancer.

such as perineural invasion,[18] resection margin status,[19] or adjuvant chemotherapy,[20] are also associated with survival outcomes in patients with pancreatic cancer. Altogether, three studies developed the nomograms for predicting the survival of resected nonmetastatic PDAC and revealed higher performances than that of the 8th edition of AJCC staging system.[6-8] For the customized treatment, the current tumor-node-metastasis staging system should be revised or replaced with more delicate and accurate prediction model.

Different from the other prediction models,[7,8] this model included the tumor location. Traditionally, the OS rate of pancreatic body or tail cancer was lower than that of the pancreatic head cancer, because of its discovery at more advanced or distant metastatic status, and the lower R0 resection rate.[21-23] However, in this prediction model, tumor location was associated with survival outcome in resected PDAC with the head cancer showing worse survival outcome than body or tail cancer (HR, 1.12; 95% CI, 1.041 to 1.198; p=0.002). It was probably because this prediction model was developed only with the patients who underwent curative-intent pancreatectomy. The patients with body or tail lesion showed better survival outcome than those with head lesion in resected PDAC in two studies. In the Nationwide Inpatient Sample database in the United States, patients who underwent distal pancreatectomy showed lower in-hospital mortality rate than those who underwent pancreatoduodenectomy.[24] The nationwide database in Germany showed similar results with the lower mortality rate of distal pancreatectomy than that of pancreatoduodenectomy in patients with pancreatic neoplasms.[25] It might be due to pancreatoduodenectomy being more invasive than distal pancreatectomy, with pancreatoduodenectomy showing higher mortality rate in patients requiring intervention for the pancreatoduodenectomy-related complications.[25]

Adjuvant chemotherapy for resected PDAC was considered to be the standard treatment because survival outcome was better in patients with adjuvant chemotherapy than in patients with observation after pancreatectomy.[26]

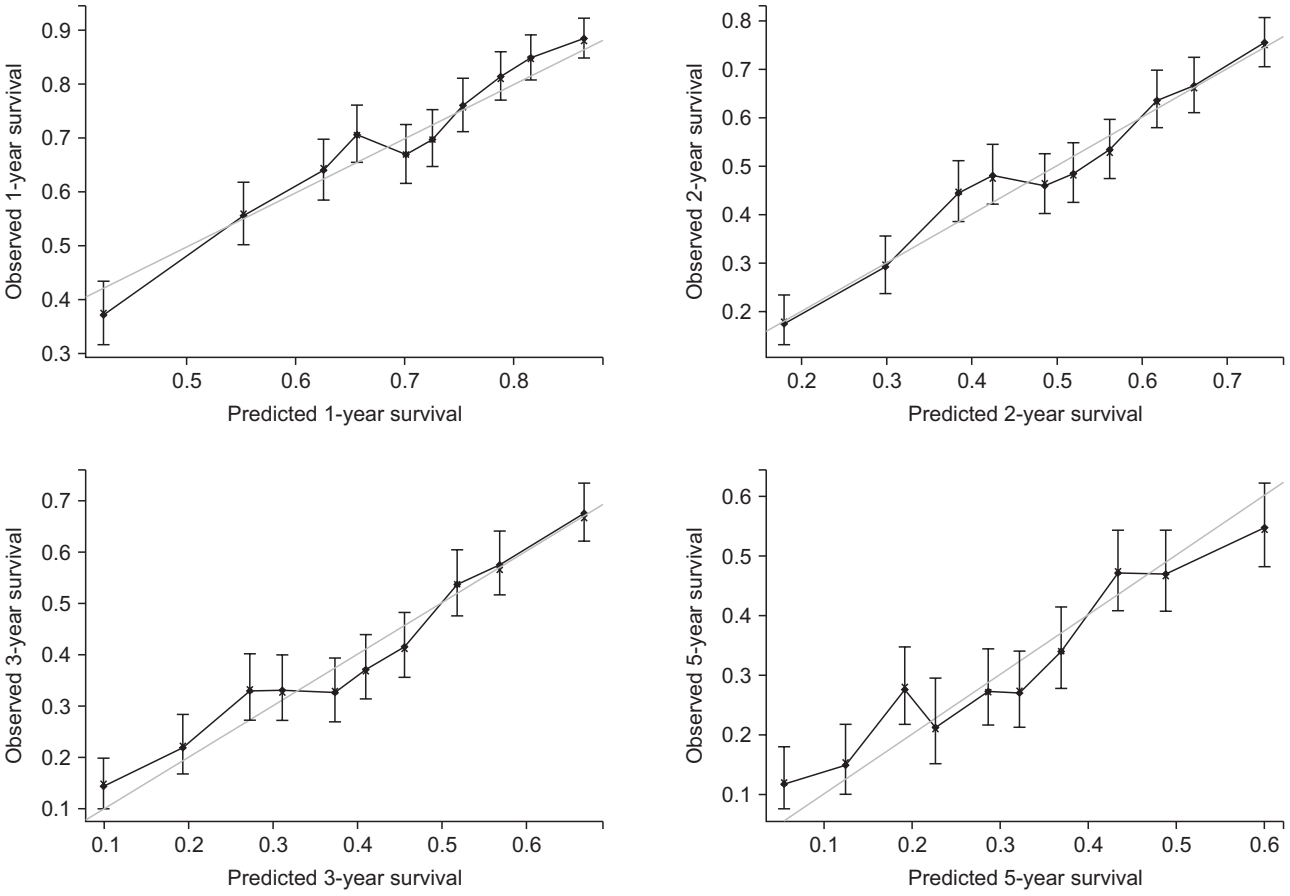

Fig. 4. Calibration curves for 1-year, 2-year, 3-year and 5-year overall survival.

In this study, the patients who received adjuvant chemotherapy had better 5-year OS rate than those who did not, in SEER database (21.3% vs 17.3%, p<0.001) (Table 1) and KOTUS-BP (36.0% vs 29.4%, p=0.004) (Table 1). In the multivariate Cox-PH model, no adjuvant chemotherapy was associated with worse survival outcome (HR, 1.789; 95% CI, 1.696 to 1.887; p<0.001) (Table 2). Although the detailed chemotherapy regimens were not investigated in this study, and the chemotherapy regimens might be heterogeneous in both databases, this model suggested that with or without adjuvant chemotherapy had more statistical power than that of AJCC T stage because of higher HR (1.789) of no chemotherapy than that of AJCC T stage (HR T2 vs T1, 1.394; HR T3 vs T1, 1.723).

Histologic differentiation was one of the prognostic factors associated with poor survival outcome in other malignancies.[27,28] A study revealed the association of histologic differentiation with survival outcome of PDAC in the univariate analysis.[29] In this study, the histologic differentiation discriminated the survival outcomes with statistical significance in both databases, with poor survival outcome in the multivariate Cox-PH model (Table 2). In addition, the statistical power of histologic differentiation was similar to that of AJCC N stage with comparable HRs comparable between two variables.

Because the nomograms were established based on the prognostic factors related with the disease, similar variables would be selected and included among the models of the same disease. Although previous models of PDAC were developed with different cohorts, the variables included in these models were quite similar (e.g., AJCC T stage, N stage or lymph node ratio, histologic differentiation, resection margin status, etc.).[5-8] In previous survival prediction models, their performances were also comparable with the C-indices of these models in the external validations as 0.58 to 0.65 (Table 3).[6-8] Consequently, the predictive performance of models consisting of clinical variables only using conventional multivariate Cox-PH analysis seems difficult to exceed 0.7. Recently, machine learning techniques have been utilized to develop prediction models for increasing the prediction power better than the conventional multivariate logistic regression model.[30] In addition, if high-dimensional variables, such as genomics or transcriptomics data, were available, the performance might be improved.[31] Therefore, a better prediction model might be established when the clinical information data along with genomic

Table 3. Summary of Previous Studies That Had Independent Model Development and External Validation

Study	Model development cohort	External validation cohort	C-index
Huang *et al.*[6]	SEER database (n=9,519)	European database (4 countries, n=2,318)	0.58–0.63
Pu *et al.*[7]	SEER database (n=12,343)	Zhongshan Hospital (n=127)	0.63
van Roessel *et al.*[8]	International database (8 countries, n=3,081)	Academic Medical Center, Amsterdam (n=350)	0.65

SEER, Surveillance, Epidemiology and End Results; C-index, Harrell concordance index.

data, and other statistical methods are utilized.

This study had some limitations being a retrospective study. In addition, because the two databases were national cohort-based, specific chemotherapy regimens were not investigated, thereby not reflecting the effect of survival benefit of the recent chemotherapy protocols, such as FOLFIRINOX or gemcitabine plus nab-paclitaxel, etc. However, the prediction model in this study was created with the SEER database, one of the largest, qualified database, and was externally validated with the KOTUS-BP data, prospectively registered and regularly managed by the pancreatobiliary surgeons at the specialized centers in Korea. The lack of detailed regimen of adjuvant chemotherapy without information about resection margin status and information on whether neoadjuvant chemotherapy was performed or not in the SEER database were other limitations. The difference of race distribution between two databases was also one of the limitations that 81.6% of patients were White, 10.4% were Black, and less than 5% were Eastern-Northern Asian (1.6% of Chinese, 1.2% of Japanese, 0.96% of Korean) in the SEER database. The international collaborate prospective studies should be performed in future, to develop and validate the global prediction model of resected PDAC with higher performance power. Moreover, a new prediction model with preoperative variables would be helpful for clinicians to decide tailored treatment strategy for treating pancreatic cancer.

In conclusion, the survival prediction model of resected PDAC could predict the 1-, 2-, 3-, and 5-year survival with the reliable performance (C-indices, 0.650, 0.665, 0.675, and 0.686, respectively) when applied to the Korean patients. The external validation studies with other nationwide databases are needed to evaluate the performance power of this model.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGEMENTS

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute KHIDI, funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI16C2307) and the Collaborative Genome Program for Fostering New Post-Genome Industry of the National Research Foundation funded by the Ministry of Science and ICT (NRF-2017M3C9A5031591).

AUTHOR CONTRIBUTIONS

Study concept and design: J.S.K., L.M., J.S.H., C.S.L., J.Y.J., T.P. Data acquisition, analysis and interpretation: J.S.K., I.W.H., S.H.S., Y.S.Y., H.S.H., D.W.H., J.H.L., W.J.L., S.J.P., J.S.P., Y.K., H.L., Y.D.Y., J.D.Y., S.E.L., I.Y.P., C.Y.J., Y.R, S.R.K., J.I.M., S.K.L., H.J.K., H.K., W.K. Drafting of the manuscript, critical revision: J.S.K., L.M., C.S.L., J.Y.J. Statistical analysis: J.S.K., L.M., S.L., T.P. Obtained funding: J.Y.J., T.P. Study supervision: C.S.L., J.Y.J., T.P.

ORCID

Jae Seung Kang https://orcid.org/0000-0001-6587-9579
Lydia Mok https://orcid.org/0000-0002-4029-5793
Jin Seok Heo https://orcid.org/0000-0001-6767-2790
In Woong Han https://orcid.org/0000-0001-7093-2469
Sang Hyun Shin https://orcid.org/0000-0002-2533-4491
Yoo-Seok Yoon https://orcid.org/0000-0001-7621-8557
Ho-Seong Han https://orcid.org/0000-0001-9659-1260
Dae Wook Hwang https://orcid.org/0000-0002-1749-038X
Jae Hoon Lee https://orcid.org/0000-0002-6170-8729
Woo Jung Lee https://orcid.org/0000-0001-9273-261X
Sang Jae Park https://orcid.org/0000-0001-5582-9420
Joon Seong Park https://orcid.org/0000-0001-8048-9990
Yonghoon Kim https://orcid.org/0000-0001-9968-645X
Huisong Lee https://orcid.org/0000-0002-3565-6064

Young-Dong Yu https://orcid.org/0000-0003-3452-338X
Jae Do Yang https://orcid.org/0000-0001-9701-7666
Seung Eun Lee https://orcid.org/0000-0003-1830-9666
Il Young Park https://orcid.org/0000-0003-4590-2297
Chi-Young Jeong https://orcid.org/0000-0003-4121-6695
Younghoon Roh https://orcid.org/0000-0002-0165-0318
Seong-Ryong Kim https://orcid.org/0000-0002-7903-9797
Ju Ik Moon https://orcid.org/0000-0002-8120-5854
Sang Kuon Lee https://orcid.org/0000-0002-3720-2461
Hee Joon Kim https://orcid.org/0000-0002-8636-5726
Seungyeoun Lee https://orcid.org/0000-0001-7941-8933
Hongbeom Kim https://orcid.org/0000-0002-1595-0135
Wooil Kwon https://orcid.org/0000-0002-4827-7805
Chang-Sup Lim https://orcid.org/0000-0002-2349-9647
Jin-Young Jang https://orcid.org/0000-0003-3312-0503
Taesung Park https://orcid.org/0000-0002-8294-590X

REFERENCES

1. Jung KW, Won YJ, Kong HJ, Lee ES. Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2016. Cancer Res Treat 2019;51:417-430.
2. Chun YS, Pawlik TM, Vauthey JN. 8th Edition of the AJCC cancer staging manual: pancreas and hepatobiliary cancers. Ann Surg Oncol 2018;25:845-847.
3. 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:e183617.
4. Kamarajah SK, Burns WR, Frankel TL, Cho CS, Nathan H. Validation of the American Joint Commission on Cancer (AJCC) 8th edition staging system for patients with pancreatic adenocarcinoma: a Surveillance, Epidemiology and End Results (SEER) analysis. Ann Surg Oncol 2017;24:2023-2030.
5. Deng QL, Dong S, Wang L, et al. Development and validation of a nomogram for predicting survival in patients with advanced pancreatic ductal adenocarcinoma. Sci Rep 2017;7:11524.
6. Huang L, Balavarca Y, van der Geest L, et al. Development and validation of a prognostic model to predict the prognosis of patients who underwent chemotherapy and resection of pancreatic adenocarcinoma: a large international population-based cohort study. BMC Med 2019;17:66.
7. Pu N, Lv Y, Zhao G, et al. Survival prediction in pancreatic cancer patients with no distant metastasis: a large-scale population-based estimate. Future Oncol 2018;14:165-175.
8. van Roessel S, Strijker M, Steyerberg EW, et al. International validation and update of the Amsterdam model for prediction of survival after pancreatoduodenectomy for pancreatic cancer. Eur J Surg Oncol 2020;46:796-803.
9. Forster MR. Key concepts in model selection: performance and generalizability. J Math Psychol 2000;44:205-231.
10. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361-387.
11. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics 2000;56:337-344.
12. Demler OV, Paynter NP, Cook NR. Tests of calibration and goodness-of-fit in the survival setting. Stat Med 2015;34:1659-1680.
13. Freedman AN, Seminara D, Gail MH, et al. Cancer risk prediction models: a workshop on development, evaluation, and application. J Natl Cancer Inst 2005;97:715-723.
14. Sun F, Ma K, Yang X, et al. A nomogram to predict prognosis after surgery in early stage non-small cell lung cancer in elderly patients. Int J Surg 2017;42:11-16.
15. Wang C, Yang C, Wang W, et al. A prognostic nomogram for cervical cancer after surgery from SEER database. J Cancer 2018;9:3923-3928.
16. Wang SJ, Fuller CD, Kim JS, Sittig DF, Thomas CR Jr, Ravdin PM. Prediction model for estimating the survival benefit of adjuvant radiotherapy for gallbladder cancer. J Clin Oncol 2008;26:2112-2117.
17. Xie X, Zhou Z, Song Y, et al. The management and prognostic prediction of adenocarcinoma of appendix. Sci Rep 2016;6:39027.
18. Chen JW, Bhandari M, Astill DS, et al. Predicting patient survival after pancreaticoduodenectomy for malignancy: histopathological criteria based on perineural infiltration and lymphovascular invasion. HPB (Oxford) 2010;12:101-108.
19. Kim KS, Kwon J, Kim K, Chie EK. Impact of resection margin distance on survival of pancreatic cancer: a systematic review and meta-analysis. Cancer Res Treat 2017;49:824-833.
20. Jang JY, Kang JS, Han Y, et al. Long-term outcomes and recurrence patterns of standard versus extended pancreatectomy for pancreatic head cancer: a multicenter prospective randomized controlled study. J Hepatobiliary Pancreat Sci 2017;24:426-433.
21. Artinyan A, Soriano PA, Prendergast C, Low T, Ellenhorn JD, Kim J. The anatomic location of pancreatic cancer is a prognostic factor for survival. HPB (Oxford) 2008;10:371-376.
22. Lau MK, Davila JA, Shaib YH. Incidence and survival of pancreatic head and body and tail cancers: a population-based study in the United States. Pancreas 2010;39:458-462.

23. van Erning FN, Mackay TM, van der Geest LGM, et al. Association of the location of pancreatic ductal adenocarcinoma (head, body, tail) with tumor stage, treatment, and survival: a population-based analysis. Acta Oncol 2018;57:1655-1662.
24. McPhee JT, Hill JS, Whalen GF, et al. Perioperative mortality for pancreatectomy: a national perspective. Ann Surg 2007;246:246-253.
25. Nimptsch U, Krautz C, Weber GF, Mansky T, Grützmann R. Nationwide in-hospital mortality following pancreatic surgery in Germany is higher than anticipated. Ann Surg 2016;264:1082-1090.
26. Neoptolemos JP, Moore MJ, Cox TF, et al. Effect of adjuvant chemotherapy with fluorouracil plus folinic acid or gemcitabine vs observation on survival in patients with resected periampullary adenocarcinoma: the ESPAC-3 periampullary cancer randomized trial. JAMA 2012;308:147-156.
27. Alexander D, Jhala N, Chatla C, et al. High-grade tumor differentiation is an indicator of poor prognosis in African Americans with colonic adenocarcinomas. Cancer 2005;103:2163-2170.
28. Sun Z, Aubry MC, Deschamps C, et al. Histologic grade is an independent prognostic factor for survival in non-small cell lung cancer: an analysis of 5018 hospital- and 712 population-based cases. J Thorac Cardiovasc Surg 2006;131:1014-1020.
29. Ahn SJ, Park MS, Lee JD, Kang WJ. Correlation between 18F-fluorodeoxyglucose positron emission tomography and pathologic differentiation in pancreatic cancer. Ann Nucl Med 2014;28:430-435.
30. Lee HC, Yoon SB, Yang SM, et al. Prediction of acute kidney injury after liver transplantation: machine learning approaches vs. logistic regression model. J Clin Med 2018;7:428.
31. Grapov D, Fahrmann J, Wanichthanarak K, Khoomrung S. Rise of deep learning for genomic, proteomic, and metabolomic data integration in precision medicine. OMICS 2018;22:630-636.