BMC Gastroenterology

RESEARCH

Open Access

Construction of a combined prognostic model for pancreatic ductal adenocarcinoma based on deep learning and digital pathology images

Kaixin Hu[1,2†], Chenyang Bian[1,2†], Jiayin Yu[2], Dawei Jiang[2], Zhangjun Chen[1,2], Fengqing Zhao[2] and Huangbao Li[1,2*]

Abstract

Background Deep learning has made significant advancements in the field of digital pathology, and the integration of multiple models has further improved accuracy. In this study, we aimed to construct a combined prognostic model using deep learning-extracted features from digital pathology images of pancreatic ductal adenocarcinoma (PDAC) alongside clinical predictive indicators and to explore its prognostic value.

Methods A retrospective analysis was conducted on 142 postoperative pathologically confirmed PDAC cases. These cases were divided into training ($n=114$) and testing sets ($n=28$) at an 8:2 ratio. Tumor whole-slide imaging features were extracted and screened to construct a pathological risk model based on a pre-trained deep learning model. Clinical and pathological data from the training set were used to select independent predictive factors for PDAC and establish a clinical risk model using LASSO, univariate, and multivariate Cox regression analyses. Based on the pathological and clinical risk models, a combined model was developed. The Harrell concordance index (C-index) was computed to assess the predictive performance of each model for PDAC survival prognosis.

Results For the training and testing sets, the C-index values for the clinical risk model were 0.76 and 0.75, respectively; for the pathological risk model, they were 0.82 and 0.73, respectively; and for the combined model, they were 0.86 and 0.77, respectively. The combined model exhibited appropriate calibration at 1-, 3-, and 5-year time points, as well as a superior area under the curve of the receiver operating characteristic curve and clinical net benefit compared to the single models.

Conclusions Integrating the pathological and clinical risk models may provide a higher predictive value for survival prognosis.

Keywords Pancreatic ductal adenocarcinoma, Deep learning, Combined model, Survival prognosis

†Kaixin Hu and Chenyang Bian made equal contributions and are co-first authors.

*Correspondence:
Huangbao Li
lhb641834@163.com

[1]Jiaxing University Master Degree Cultivation Base, Zhejiang Chinese Medical University, Jiaxing, Zhejiang, China
[2]Department of Hepatobiliary and Pancreatic Surgery, First Hospital of Jiaxing, Affiliated Hospital of Jiaxing University, Jiaxing, Zhejiang, China

Background

Pancreatic cancer (PC) is a common malignant tumor of the digestive system. According to the Global Cancer Statistics 2020 [1], the mortality rate is comparable to the incidence rate, with a poor prognosis and ranking as the seventh leading cause of cancer-related death in both sexes. Pancreatic ductal adenocarcinoma (PDAC) is the most prevalent form of PC, accounting for approximately 90% of PC cases, and is considered one of the most aggressive malignancies [2]. Despite advancements in detection and treatment modalities, the 5-year survival rate for PDAC remains notably low, hovering around 10% in the United States [3]. According to the estimates of the National Cancer Center (NCC) in China [4], the 5-year survival rate for PC is extremely low, at around 7.2%, which is lower than that in the United States [5]. Given the high lethality of PDAC, establishing an effective prognostic model for survival analysis is clinically significant. Such models will not only aid in predicting the prognosis of individual patients but also provide a basis for treatment decisions, thereby enhancing patient survival.

Currently, prognosis assessment in clinical practice relies heavily on the TNM staging system [6]. However, even among patients at the same disease stage, significant variability in outcomes exists, reflecting the inadequacy of the prognostic information provided by the existing TNM staging. In recent years, histopathological image analysis has emerged as a prominent field for directly predicting the prognosis and treatment response of patients with cancer [7]. Histopathological image analysis plays a crucial role in medical research, disease diagnosis, and treatment selection [8]. Pathological diagnosis widely uses whole-slide images (WSI) owing to their suitability for remote diagnosis, ease of storage, and absence of damage [9]. In recent years, some scholars have achieved promising results in PC diagnosis using WSI. However, there is a paucity of research on assessing the prognosis of PC using WSI. Pathologists' subjective assessments of WSI may overlook important features, necessitating more advanced research methods to extract rich information from WSI.

Deep learning (DL) is a machine learning method that uses the iterative abstraction of input images to perform specific tasks [10]. Convolutional neural networks (CNNs) are a specialized type of neural network structure in DL and are primarily used for processing data with spatial structures such as images and videos. CNNs consist of convolutional, pooling, and fully connected layers, with convolutional layers being their core components. The hierarchical stacking of CNNs enables the model to learn multi-level abstract features of the data, thereby enhancing the model's representational capacity and generalizability. The DL extraction process can objectively and automatically extract richer and more discriminative information features from datasets without requiring excessive manual intervention.

As DL and big data analytics technologies continue to advance, more studies are focusing on leveraging these advanced techniques to construct more precise prognostic models for PC. The computational analysis of medical images has yielded promising results in predicting overall survival (OS) in patients with PC [11]. Moreover, DL-based prognostic prediction based on pathological images has also been applied to diseases such as colorectal cancer [12], lung cancer [7, 13], liver cancer [14], gastric cancer [15], and breast cancer [16].

In this study, we developed a combined survival analysis prognostic model that integrates DL with traditional clinical features to more accurately predict the OS of patients with PDAC. Through a comprehensive analysis of patients' clinical and pathological features, we aim to provide a robust tool to enhance the prognosis assessment of PDAC and to support future personalized treatment strategies.

Methods

Data acquisition

The study data, including digitally scanned pathological slides and clinical information, were retrospectively collected from patients who underwent pancreaticoduodenectomy at the First Hospital of Jiaxing from March 1, 2012, to March 1, 2022, and were pathologically diagnosed with PDAC. The clinical data included preoperative laboratory tests, including laboratory tests for Carcinoembryonic Antigen (CEA), Cancer Antigen 19−9 (CA19-9), Cancer Antigen 242 (CA242), and bilirubin (supplementary tables S1, S2, S3), and postoperative clinical records, comprising sex (Sex), age (Age), tumor site (Site), vascular invasion (Vascular), tumor differentiation (Grade), and T, N, and M stage. Each case of PC was represented by a representative hematoxylin and eosin (H&E)-stained digital slide in mrxs format, magnified at 40x, totaling 142 slides from 142 patients with PC. None of the included patients with PDAC received treatment before specimen collection. The TNM staging and tumor differentiation grading were based on the AJCC Cancer Staging Manual (9th edition) [17].

Inclusion and Exclusion Criteria.

Inclusion criteria

(1) The patient was diagnosed postoperatively with PDAC.
(2) The study includes comprehensive clinical data, follow-up records, and corresponding high-resolution digital images of frozen pathological sections.

Exclusion criteria

(1) Patients with concomitant other neoplastic diseases.
(2) Clinical data are incomplete, patients lost to follow-up, poor quality digital pathological slides, and patients for whom digital pathological slides cannot be obtained.

Image preprocessing

Histopathological sections were post-stained with H&E, fixed in formalin, and embedded in paraffin for each WSI sample. All H&E-stained slides were scanned using a digital slide scanner. The WSI was sampled into small image patches sized 256×256 pixels, with a magnification of 20×(0.5 μm/pixel). Empty image patches were inspected and removed after output. Subsequently, the Vahadane algorithm was employed in Python to perform color normalization on all images, with the aim to mitigate the impact of color variations between different samples on computer vision tasks.

Tumor feature extraction and construction of a pathological risk signature

Deep neural networks were used to extract features from image patches using transfer learning techniques. Three pre-trained models, namely MobileNet_v3_small, ResNet18, and DenseNet121, were pre-trained on ImageNet. The images were inputted into each network, and the input features of the last fully connected layer of each network were selected as output features to obtain the final multi-dimensional image features from the three models. Subsequently, the mean value of each dimension of the deep features was calculated to obtain patient-level features. The average feature of multiple image patch levels represents the feature set for each patient. (Fig. 1)

The dataset was randomly partitioned into 80% training and 20% testing sets. Patient-level features from the three neural networks were used to construct a multivariate Cox proportional hazards model for each feature set in the training set for predicting patient OS. The features derived from the Cox model were defined as the "pathological risk signature (PRS)." Patient-level features were constructed through linear combinations by weighting their coefficients in the training and testing datasets. The prognostic performance of the three neural network

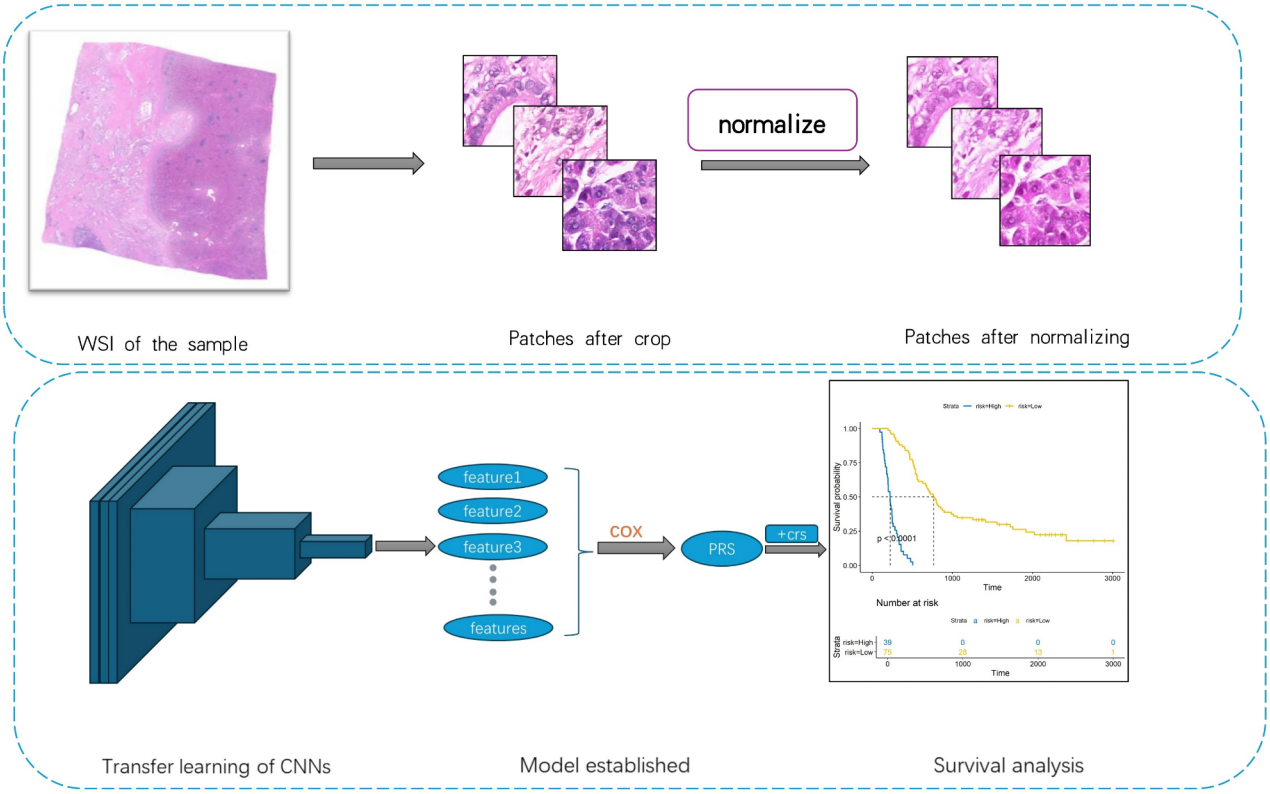

Fig. 1 Flowchart of the research methodology. (1) Employing deep learning networks for transfer learning techniques to extract features from whole-slide images to facilitate the construction of a Cox model for the computation of PRS. (2) Concomitantly, integrating the CRS in the Cox prognostic model construction, followed by Kaplan–Meier survival analysis to assess the discriminatory efficacy of the combined model in stratifying high and low-risk subgroups. WSI: Whole-slide images, PRS: Pathological risk signature, CRS: Clinical risk signature, CNNs: Convolutional neural networks

feature sets was assessed by computing the C-index values on the training and testing datasets. Features with a C-index ≥ 0.60 [18] were selected as better prognostic variables for further analysis.

Survival analysis of clinical and clinicopathological features

The patients' preoperative laboratory test results served as continuous variables, and the training set was dimensionally reduced using the least absolute shrinkage and selection operator (LASSO) algorithm. Through 10-fold cross-validation, features with non-zero coefficients were selected, and the optimal variables were combined with the clinical pathological data. The final inclusion of variables was determined using univariate and multivariate Cox regression analyses and clinical expertise. A clinical risk model was established, and its performance was evaluated by calculating the C-index in the training and testing datasets. Subsequently, the Cox model was used to calculate the risk scores of the samples, and the resulting features were defined as "clinical risk signatures (CRS)." This model linearly combines patient-level features through weighted use of their coefficients in the training and testing datasets.

Construction of a survival prognosis model combining the PRS and CRS

The obtained CRS was combined with the PRS. Univariate and multivariate Cox regression analyses were conducted on the training set to analyze their correlation with survival and respective independence. A Cox proportional hazards model was fitted, and 10-fold cross-validation was performed to calculate the average C-index. Subsequently, the C-index was calculated for the training and testing datasets to assess the performance of the prognostic model. To further validate the model's performance, patients were stratified into high- or low-risk groups based on the calculated final risk scores using thresholds determined in the training set using X-tile software. The same thresholds were applied for stratification in the testing set. Kaplan–Meier survival analysis was employed to analyze survival curves, with differences assessed using the log-rank test.

Statistical analysis

Statistical analysis was performed using R software version 4.4.0, using packages such as survival, survminer, ggplot2, glmnet, dplyr, mice, caret, sampling, riskRegression, and dcurves. The C-index was computed using the R package survival to measure discrimination. Ten-fold cross-validation was used for model validation, where the data were split into ten subsets. Nine subsets were used for training and one subset for testing, and this process was repeated ten times to obtain the average model performance. Subgroup division was performed using X-tile software. Survival curves were generated based on Kaplan–Meier analysis. In Kaplan–Meier survival analysis, a weighted log-rank test was applied to obtain the p-value to test the significance of survival differences between high-risk and low-risk subgroups. DeLong's test was used to assess the difference between two AUCs. Furthermore, p-values < 0.05 were considered statistically significant, whereas p-values < 0.1 were considered indicative of potential relevance. We performed image preprocessing and feature extraction using Python (version 3.11.8) and TensorFlow (version 2.9.0).

Results

Patient characteristics

Patients with tumors classified as highly and moderately differentiated were classified into the low-risk differentiation group, whereas those classified as poorly and undifferentiated were classified into the high-risk differentiation group. The ages of all patients ranged from 39 to 82 years, with a mean age of 65 years. The median follow-up time for OS was 483 days (Table 1).

CRS model

For preoperative clinical laboratory test indicators, we employed LASSO regression to screen parameters and provided the coefficient change characteristics of these variables (Supplementary Fig. 1). To obtain a high-performance model with minimal variables, we conducted iterative analysis using a 10-fold cross-validation method. The preoperative laboratory tests finally included the lymphocyte ratio (LR, %), platelet count (PLC, 10^9/L), platelet distribution width (PDW, %), globulin concentration (GLC, g/L), potassium ion concentration (PIC, mmol/L), chloride concentration (CHC, mmol/L), CA242 (U/ml), uric acid concentration (UAC, μmol/L), and mean hemoglobin concentration (MHC, g/L), totaling 9 clinical indicators. The predicted model's λ value is 0.1254058. Clinical pathological data combined with clinical experience included sex, age, tumor site, vascular invasion, tumor differentiation, and T, N, and M stage. In the univariate analysis, variables with significant significance ($p < 0.05$) were included in the multivariate analysis. In the multivariate analysis, variables with significant significance ($p < 0.05$) were included in the final clinical risk model (Table 2). The C-index values of the final clinical risk model for the training and test sets were 0.76 (95% CI: 0.72–0.81) and 0.75 (95% CI: 0.64–0.86), respectively. The average C-index obtained by 10-fold cross-validation in the training set was 0.761.

PRS model construction

We extracted features of 576 dimensions from the MobileNet_v3_small model, 512 dimensions from

Table 1 Laboratory test indicators and clinical-pathological characteristics of the training and testing sets

Variables	Train	Test	*P*
	(*n* = 114)	(*n* = 28)	
LR, mean ± SD (%)	20.44 ± 8.94	20.20 ± 11.28	0.905
PDW, mean ± SD (%)	15.79 ± 2.11	15.21 ± 1.44	0.168
MHC, mean ± SD (g/L)	332.69 ± 12.57	328.79 ± 11.75	0.138
GLC, mean ± SD (g/L)	24.84 ± 4.46	25.70 ± 3.70	0.348
PIC, mean ± SD (mmol/L)	3.93 ± 0.43	3.85 ± 0.44	0.349
CA242, mean ± SD (U/ml)	79.02 ± 83.67	96.06 ± 89.20	0.342
Age, mean ± SD (years)	64.61 ± 8.65	67.50 ± 7.52	0.107
PLC, M (Q_1, Q_3) (10^9/L)	178.50 (142.50, 218.00)	162.00 (119.50, 200.00)	0.174
CHC, M (Q_1, Q_3) (mmol/L)	103.15 (101.00, 105.68)	104.20 (101.33, 106.08)	0.454
UAC, M (Q_1, Q_3) (μmol/L)	225.70 (170.67, 269.42)	283.20 (230.52, 310.85)	0.002
T			0.572
T1	14 (12.28)	1 (3.57)	
T2	59 (51.75)	18 (64.29)	
T3	27 (23.68)	6 (21.43)	
T4	14 (12.28)	3 (10.71)	
N			0.767
N0	61 (53.51)	13 (46.43)	
N1	44 (38.60)	12 (42.86)	
N2	9 (7.89)	3 (10.71)	
M			1
M0	109 (95.61)	27 (96.43)	
M1	5 (4.39)	1 (3.57)	
Vascular			0.258
No	78 (68.42)	16 (57.14)	
Yes	36 (31.58)	12 (42.86)	
Grade			0.602
High	51 (44.74)	11 (39.29)	
Low	63 (55.26)	17 (60.71)	
Site			0.139
Head	78 (68.42)	15 (53.57)	
Tail	36 (31.58)	13 (46.43)	
Sex			0.723
Male	65 (57.02)	17 (60.71)	
Female	49 (42.98)	11 (39.29)	

Continuous normally distributed variables are described as the mean ± standard deviation and were compared between groups using the t-test. Continuous skewed variables are described as the median and interquartile range, and between-group differences were compared using the non-parametric Wilcoxon rank-sum test. Differences between qualitative variables were assessed using the *p*-value of the chi-square test

LR: Lymphocyte ratio (%), PLC: Platelet count (10^9/L), PDW: Platelet distribution width (%), GLC: Globulin concentration (g/L), PIC: Potassium ion concentration (mmol/L), CHC: Chloride concentration (mmol/L), CA 242: Carbohydrate antigen 242 (U/ml), UAC: Uric acid concentration (μmol/L), MHC: Mean hemoglobin concentration (g/L)

Table 2 Univariate and multivariate analyses for overall survival

Variables	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*P*	HR (95% CI)	*P*
T				
T1	1		1	
T2	1.20 (0.63–2.26)	0.582	1.39 (0.69–2.76)	0.355
T3	1.42 (0.70–2.89)	0.326	2.30 (1.07–4.95)	0.033
T4	2.80 (1.28–6.12)	0.01	1.66 (0.65–4.20)	0.286
N				
N0	1		1	
N1	1.58 (1.03–2.41)	0.034	1.61 (1.03–2.52)	0.038
N2	2.85 (1.36–5.97)	0.005	2.23 (0.91–5.49)	0.081
M				
M0	1		1	
M1	6.04 (2.38–15.34)	< 0.001	12.16 (3.53–41.92)	< 0.001
Vascular				
No	1			
Yes	1.43 (0.93–2.20)	0.1		
Grade				
High	1		1	
Low	2.31 (1.52–3.52)	< 0.001	1.73 (1.08–2.79)	0.024
LR	0.97 (0.95–0.99)	0.021	0.97 (0.95–0.99)	0.01
PLC	1.00 (1.00–1.00)	0.099		
PDW	1.22 (1.11–1.34)	< 0.001	1.25 (1.13–1.38)	< 0.001
GLC	0.95 (0.91–0.99)	0.022	0.93 (0.89–0.98)	0.011
PIC	1.36 (0.85–2.19)	0.204		
CHC	0.92 (0.87–0.97)	0.002	0.92 (0.87–0.99)	0.018
CA242	1.01 (1.01–1.01)	0.042	1.01 (1.01–1.01)	0.026
UAC	0.99 (0.99–0.99)	0.009	1.00 (0.99–1.00)	0.359
MHC	1.02 (1.01–1.04)	0.024	1.02 (1.00–1.04)	0.094

HR: Hazard ratio, CI: Confidence interval

LR: Lymphocyte ratio (%), PLC: Platelet count (10^9/L), PDW: Platelet distribution width (%), GLC: Globulin concentration (g/L), PIC: Potassium ion concentration (mmol/L), CHC: Chloride concentration (mmol/L), CA 242: Carbohydrate antigen 242 (U/ml), UAC: Uric acid concentration (μmol/L), MHC: Mean hemoglobin concentration (g/L)

the ResNet18 model, and 1000 dimensions from the DenseNet121 model. Then, based on the patient-level feature sets obtained from these three network models, three Cox models were constructed. We randomly divided the dataset into a training set (114 patients) and a test set (28 patients). Owing to the large number of features extracted by each model, we conducted correlation analysis to identify the 110 least correlated features from each model, before constructing a Cox proportional hazards model using these selected features. To optimize model performance, we used a grid search method to systematically explore the hyperparameter space, specifically, by searching Penalizer within the range of 0.001 to 0.01. The grid search evaluated the performance of each parameter set using 10-fold cross validation, where the C-index was the performance metric. When the Penalizer was set to 0.004, the C-index values reached 0.823 on the training set and 0.727 on the test set. The average C-index value from the 10-fold cross-validation on the training set was 0. 832. The results demonstrate that the model using the DenseNet121 network feature set as input performed the best, with C-index values of 0.823 and 0.727 for the training and test sets, respectively; the model that used the ResNet18 network feature

set as input achieved C-index values of 0.826 and 0.456 for the training and test sets, respectively; and the model that used the MobileNet_v3_small network feature set as input achieved C-index values of 0.835 and 0.525 for the training and test sets, respectively. Therefore, based on the DenseNet121 feature set, the model calculated the final PRS values of the training and test sets. The training and testing set C-indices of the final pathological risk signature model were 0.82 (95% CI: 0.78–0.87) and 0.73 (95% CI: 0.63–0.83), respectively.

Construction of the combined model

The risk scores, CRS and PRS were computed by integrating the clinical and pathological risk models of the training set. The analysis reveals a lack of statistical differences between the two variables across the training and testing sets (PRS: $P=0.114$, CRS: $P=0.428$). For univariate variable analysis, both variables were significantly correlated with survival prognosis ($P<0.001$). Upon inclusion in the multifactorial variable analysis, both variables continued to exhibit a significant correlation with survival prognosis ($P<0.001$) (Table 3). The variance inflation factor (VIF) between the two variables was 1.003217, indicating almost no multicollinearity between the variables and a strong level of independence. The combined model based on the PRS and CRS yielded an average C-index of 0.826 through 10-fold cross-validation on the training set. The calculated C-index values for the training and testing sets were 0.86 (95% CI: 0.83–0.89) and 0.77 (95% CI: 0.68–0.85), respectively. The combined model was calculated as follows: 0.423683 PRS model+0.263226 CRS model. The combined model significantly outperformed the single PRS and single CRS models in fitting survival data, with a likelihood ratio test result of $P<0.001$. The ROC curve AUCs for the combined model are shown in Supplementary Fig. 4, indicating that the AUC of the combined model exceeded those of the PRS and CRS models. DeLong's test revealed that the differences between the combined model and the CRS model ($P=0.1055$) and the PRS model ($P=0.1658$) were not statistically significant. Nevertheless, the combined model demonstrated the highest AUC and net benefit. Using the threshold determined by the x-tile software (0.96), the training set was divided into high-risk subgroup (>0.96) and low-risk subgroup (≤0.96). The same threshold was applied to the testing set. In the Kaplan–Meier survival analysis, the high- and low-risk subgroups exhibited good discriminative ability, with the training set showing statistical significance ($P<0.001$) and the testing set showing significance at $P=0.0076$ (Fig. 2).

Table 3 Univariate and multivariate analyses for overall survival

Variables	Univariate analysis		Multivariate analysis	
	HR (95% CI)	P	HR (95% CI)	P
PRS	1.54 (1.40–1.69)	< 0.001	1.52 (1.38–1.68)	< 0.001
CRS	1.35 (1.25–1.47)	< 0.001	1.30 (1.19–1.42)	< 0.001

HR: Hazard ratio, CI: Confidence interval

PRS: Pathological risk signature, CRS: Clinical risk signature

Evaluation of model performance at different time points

Based on clinical experience and previous research, the time points for evaluation were 1, 3, and 5 years. The calibration curves and calibration plots for our joint model at different time points on the training and testing sets are shown in Fig. 3. The calibration of the training and testing sets appeared to be reasonable at various time points. The ROC curve AUC comparison between the integrated model and the pathological risk feature model and clinical risk model is illustrated in Supplementary Fig. 2, whereas the decision curve is depicted in Supplementary Fig. 3.

Discussion

In this retrospective study, we developed a combined model to predict survival outcomes for patients with PDAC, achieving reasonable discrimination and calibration in the derivation and validation cohorts. Additionally, our model demonstrated superior predictive ability compared to commonly used staging systems. Ultimately, the combined model classified patients into two risk groups, highlighting a significant difference in overall survival between the groups.

PDAC is the most common type of PC and is characterized by high lethality and poor prognosis. The traditional anatomically based TNM staging system has facilitated the objective description and classification of the anatomical extent of malignant diseases through simple, reproducible, and site-specific algorithms [19]. However, the traditional TNM staging system has an accuracy rate of 0.60 in predicting the prognosis of patients with PC. Efforts to gradually optimize the TNM staging system have yielded minimal improvements in accuracy [20]. Simplistic classification alone cannot meet the demand for precise prediction. In our study of clinical risk, we augmented the TNM staging system using pathological feature descriptions and preoperative examination indicators. We conducted univariate and multivariate variable analyses on clinical variables of PDAC and demonstrated that T, N, and M stage, tumor differentiation degree, as well as preoperative examination indicators, including lymphocyte proportion, PDW, globulin concentration, CHC, and CA242, were significant factors influencing patient prognosis. The TNM staging system and tumor differentiation degree are widely used indicators of prognostic prediction. David et al. [21] found that the preoperative neutrophil-to-lymphocyte ratio (NLR) was significantly associated with OS and recurrence-free survival in patients undergoing

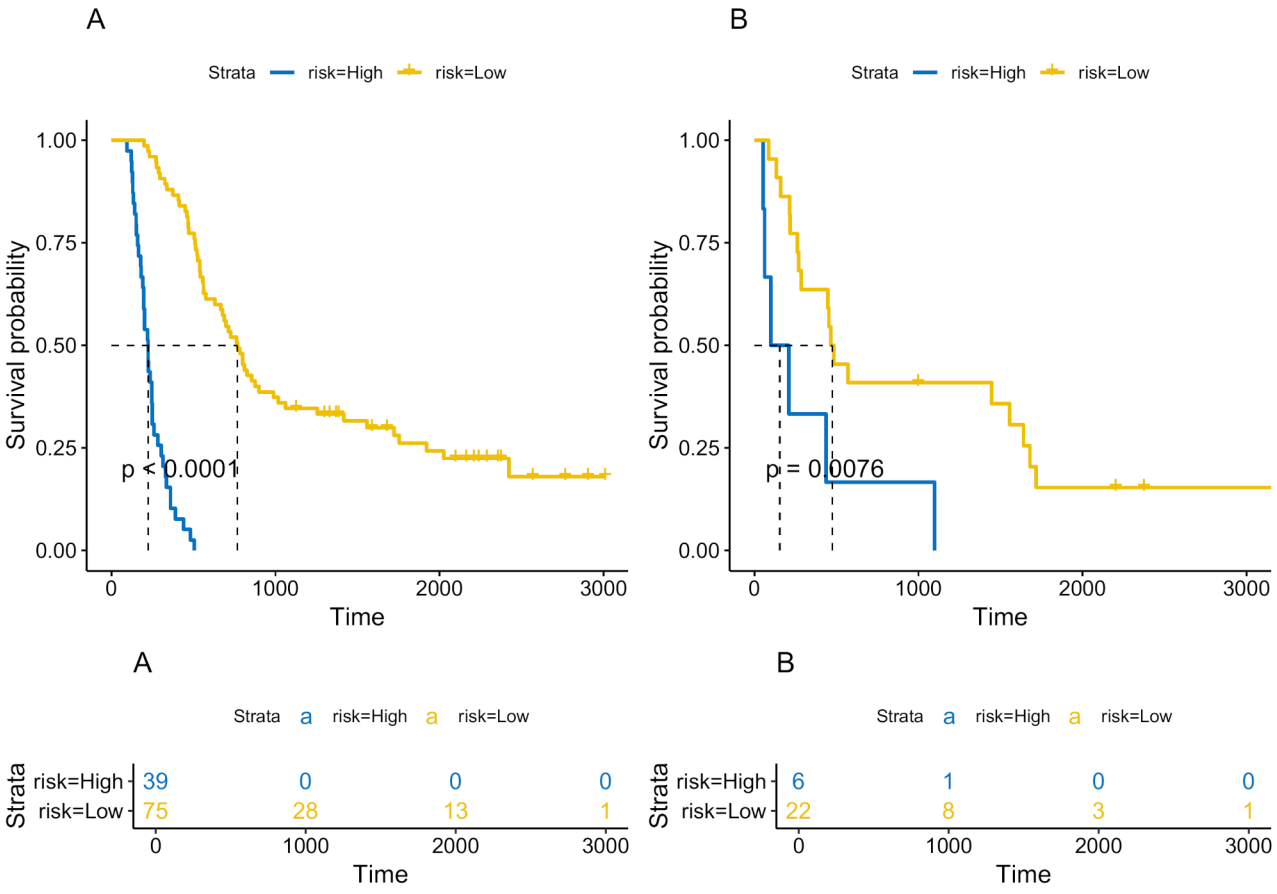

Fig. 2 Kaplan–Meier survival curves for high- and low-risk subgroups based on the combined model. **A**. Kaplan–Meier survival curves of the high- and low-risk subgroups in the training set. **B** Kaplan–Meier survival curves of the high- and low-risk subgroups in the test set

early PDAC resection in a retrospective study. A higher NLR is associated with a poorer prognosis, suggesting that lymphocyte counts alone may serve as a significant driver of survival. Our study revealed that elevated platelet width is associated with an increased risk of adverse prognosis. Previous investigations have revealed that a heightened preoperative platelet distribution width is predictive of adverse outcomes in various malignancies, including hepatocellular carcinoma [22], lung cancer [23], and breast cancer [24]. Cancer-related inflammation may be a common mechanism underlying PDW. Before treatment, the albumin-to-globulin ratio (AGR) is a commonly used inflammation-related factor. A retrospective analysis [25] revealed that a high AGR before treatment may serve as a favorable prognostic factor in patients with PC undergoing curative resection after receiving preoperative multimodal therapy, with the same study demonstrating similar findings for gastric cancer [26] and esophageal cancer [27]. Single studies focusing solely on globulin concentration are uncommon. Our research revealed that higher levels of globulin suggest higher survival rates, which may be attributable to its involvement in immune response mechanisms. Ni et al. [28] found that CA19-9 had a higher diagnostic value in patients with PC, whereas CA242 exhibited a higher prognostic value in a retrospective study. Preoperatively, elevated CA242 levels indicate a poorer survival prognosis, consistent with our study findings. CA19-9 and CEA are prominent serum biomarkers for PC [29, 30] and are commonly used in clinical practice for its diagnosis and prognosis. However, recent studies have shown a strong correlation between CA19-9 and CA242, with CA242 potentially offering greater specificity for PC. Additionally, CEA has a higher value in monitoring postoperative prognosis, assessing treatment response, and providing early warning of disease progression in locally advanced PC [31, 32]. The results of our study using DL in patients with PDAC indicate that PRS is a significant prognostic factor, independent of CRS. Patients with high PRS have a poorer postoperative survival prognosis. Zhou et al. [33] evaluated three CNN models for extracting tumor component features from whole-slide images of decomposed colorectal cancer (CRC) tissues and established prognostic models; these included VGG19, Densenet121,

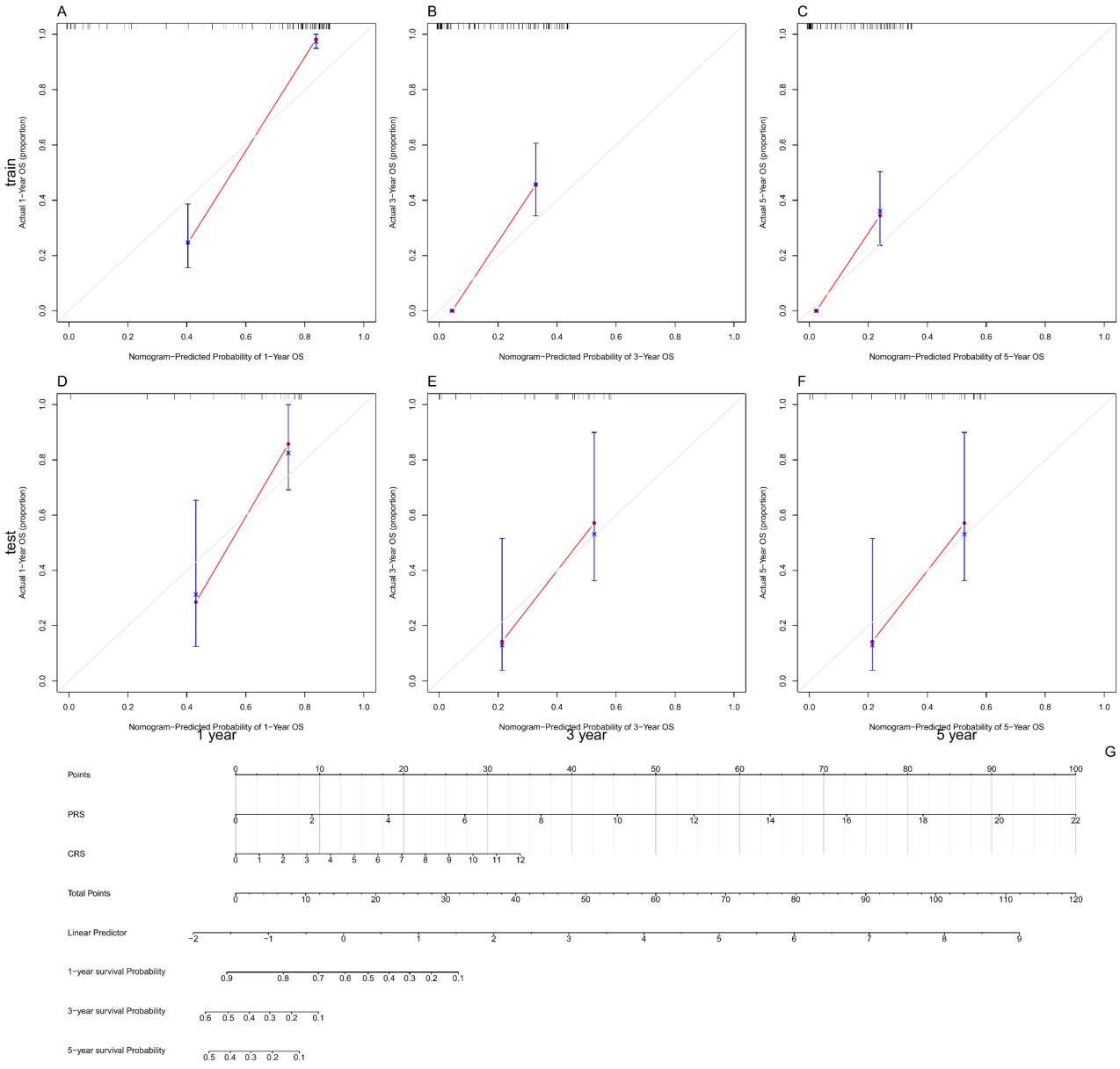

Fig. 3 Calibration curves and nomograms at different time points for the combined model. **A**, **B**, and **C**. Calibration curves of the combined model's training set at the 1-, 3-, and 5-year time points, respectively. **D**, **E**, and **F**. Calibration curves of the combined model's test set at the 1-, 3-, and 5-year time points, respectively. G. Nomogram of the combined model. PRS: Pathological risk signature, CRS: Clinical risk signature

and Resnet50. They found that the prognostic model based on features extracted by Densenet121 demonstrated the best performance. Subsequently, the deep neural network was used to extract patch-level features from whole-slide images of tumor tissues. They obtained patient-level features by averaging the patch-level features and inputting them into a Cox model to derive new prognostic biomarkers. In this study, we drew inspiration from the work of Zhou et al., by employing pre-trained DL networks to extract features from whole-slide images. In contrast to their study, we employed lighter DL models, namely MobileNet_v3_small, in conjunction with ResNet18 and DenseNet121, for performance comparison. Furthermore, to streamline the process, we opted not to perform the histological decomposition of whole-slide images. Instead, after straightforward image preprocessing, global features were extracted directly from the entire whole-slide image. We found that DenseNet121 continued to exhibit outstanding performance. Subsequently, the global features extracted using DenseNet121

were inputted into a Cox model to derive the prognostic biomarker PRS. The survival analysis results demonstrated that the PRS exhibited significant discriminative ability, as patients with higher PRS experienced a poorer prognosis. Further research findings indicated that PRS serves as a potential prognostic marker independent of CRS. Therefore, when clinical risk factors are insufficient for predicting patient prognosis, the PRS can serve as a novel factor to aid treatment decision-making. Subsequently, we combined the PRS with the CRS to predict patient prognosis, yielding superior results compared to predictions made solely based on the PRS or CRS alone. The combined model may compensate for the individual shortcomings of the PRS and CRS.

Currently, DL has shown promising performance in the diagnosis and differential diagnosis of PC [34] across various domains, including traditional omics [35], radiomics [36], and pathology [37]. As a branch of machine learning, DL has been extensively researched in prognostic prediction for various cancers. However, its application to PC prognosis remains relatively limited. Research has indicated that machine learning might be superior to DL models in pancreatic cancer prognosis studies, and the interpretability of these models has been significantly enhanced by the application of SHAP [38]. It must be acknowledged that the long follow-up periods required for PC prognosis, low incidence rates, difficulties in early diagnosis, interpretability of DL models, and complexity of PC tissue may limit the application of DL models in PC prognosis research. We conducted long-term follow-ups on patients with PC to ensure the accuracy of prognostic information. In addition, in our DL models for WSI, we moved away from the previous approach of annotating regions of interest, and instead, used pre-trained models to extract global features directly from the entire WSI. Through the development and comparison of DL models, we found that DenseNet121 performed exceptionally well. Its unique feature reuse capability helps to capture multi-level feature information, making it potentially more suitable for image feature research. This may provide valuable insights into selecting DL models in future studies.

Research has been conducted on several biochemical and genetic biomarkers regarding survival in PDAC [39, 40]. However, owing to technological limitations for gene expression testing and high patient costs, the widespread application of genetic prognostic biomarkers for clinical practice remains challenging. In this study, we leveraged multiscale information from WSI and used DL methods to extract abstract and comprehensive high-dimensional features. These features enabled the automatic prediction of prognosis and prediction, whereas the computational methods of DL may also help reduce the costs associated with prognosis prediction. Integrating our clinical risk features improved the accuracy of prognosis prediction. Our findings demonstrate the significant potential of DL methods and WSI in assessing the prognosis of patients with cancer. Multi-omics research is becoming a growing trend [41], and the integration of multiple models can reduce the risk of localized prediction failures [42]. However, as we strive to improve accuracy, we must also consider the efficiency and complexity of integrated models.

This study has several limitations. First, this is a retrospective analysis with a small sample size from a single center and lacks external validation. Therefore, further multicenter or prospective studies with larger sample sizes are warranted to validate and extend the findings. Second, our clinical model incorporates preoperative CHC as an indicator, with its role identified as a protective factor. However, no evidence currently supports its value as a prognostic predictor. Interestingly, chlorine radicals can effectively induce oxidative stress in normoxic and hypoxic tumor cells without relying on O2/H2O2, thereby overcoming the challenge of hypoxia insensitivity in clinical therapy [43]. PC typically exists in severely hypoxic environments. The introduction of Cl^- as a chlorine source can support the efficient generation of chlorine radicals, and preoperative CHC may be linked to these radicals. Future studies should explore the potential value of preoperative CHC in predicting PC prognosis. Third, because of the heterogeneity of pathological images, it is necessary to further expand the sample size and acquire higher-quality whole-slide images to improve the predictive accuracy of the prognostic model. Finally, the ambiguity surrounding the features extracted by DL and their represented meanings constitutes a primary obstacle in clinical applications. Enhancing the interpretability of DL has been a longstanding goal in clinical translation.

In summary, our study investigated preoperative serum biomarkers associated with PC prognosis and, in combination with the traditional TNM staging system, developed the CRS as a prognostic biomarker. Additionally, we used DL methods to develop the PRS as another prognostic biomarker. The combination of these markers may improve the prognosis of PDAC.

Abbreviations

AGR	Albumin-to-globulin ratio
AUC	Area under the curve
CA242	Carbohydrate antigen 242
CHC	Chloride concentration
CRS	Clinical Risk Signatures
CRC	Colorectal cancer
DL	Deep learning
PDW	Platelet distribution width
H&E	Hematoxylin and eosin
LASSO	Least absolute shrinkage and selection operator
LR	Lymphocyte ratio
MHC	Mean hemoglobin concentration
CNN	Convolutional Neural Networks
NLR	Neutrophil-to-lymphocyte ratio

OS	Overall survival
PC	Pancreatic cancer
PDAC	Pancreatic ductal adenocarcinoma
PRS	Pathological Risk Signature
PLC	Platelet count
ROC	Receiver operating characteristic
RFS	Recurrence-free survival
UAC	Uric acid concentration
VIF	Variance inflation factor
WSI	Whole-slide imaging

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12876-024-03469-4.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Acknowledgements
Not applicable.

Author contributions
All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Jiayin Yu, Chenyang Bian, Zhangjun Chen, Fengqing Zhao, and Kaixin Hu. The first draft of the manuscript was written by Kaixin Hu. Huangbao Li and Dawei Jiang commented on subsequent versions of the manuscript. All authors read and approved the final manuscript.

Funding
The work was funded by grants from the Science and Technology Project of Zhejiang Medical and Health [2021KY1103], Key Discipline of Hepatobiliary and Pancreatic Surgery of Jiaxing City [2023-zc-005], Translational Therapy Center for Hepatobiliary Pancreatic Cancer [2021-YJZX-04], Science and Technology Project of Jiaxing [2022AD30066, 2022AD30060,2023AD31059], Jointly Cultivate Disciplines of Jiaxing City and Zhejiang province [2023-PYXK-001], and National Clinical Key Specialty Construction Project [2023-GJZK-001].

Data availability
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The study protocol was approved by the Institutional Review Board of the First Hospital of Jiaxing (No. 2023-LY-522). All procedures followed the ethical standards of the responsible committee on human experimentation (institutional and national) and the Helsinki Declaration of 1964 and later versions. A waiver of the informed consent was approved by The Institutional Review Board of the First Hospital of Jiaxing owing to the retrospective nature of the study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 28 June 2024 / Accepted: 16 October 2024
Published online: 31 October 2024

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.
2. Halbrook CJ, Lyssiotis CA, Pasca di Magliano M, Maitra A. Pancreatic cancer: advances and challenges. Cell. 2023;186(8):1729–54.
3. Mizrahi JD, Surana R, Valle JW, Shroff RT. Pancreatic cancer. Lancet. 2020;395(10242):2008–20.
4. Zheng R, Zhang S, Zeng H, Wang S, Sun K, Chen R, Li L, Wei W, He J. Cancer incidence and mortality in China, 2016. J Natl Cancer Cent. 2022;2(1):1–9.
5. Maomao C, He L, Dianqin S, Siyi H, Xinxin Y, Fan Y, Shaoli Z, Changfa X, Lin L, Ji P, et al. Current cancer burden in China: epidemiology, etiology, and prevention. Cancer Biology Med. 2022;19(8):1121–38.
6. Birnbaum DJ, Bertucci F, Finetti P, Birnbaum D, Mamessier E. Head and body/tail pancreatic carcinomas are not the same tumors. Cancers (Basel). 2019;11(4).
7. Wang S, Yang DM, Rong R, Zhan X, Fujimoto J, Liu H, et al. Artificial intelligence in lung cancer pathology image analysis. Cancers (Basel). 2019;11(11):1673.
8. Xie X, Wang X, Liang Y, Yang J, Wu Y, Li L, et al. Tian GJFiO: evaluating cancer-related biomarkers based on pathological images: a systematic review. Front Oncol. 2021;11:763527.
9. Kumar N, Gupta R, Gupta SJJ. Whole slide imaging (WSI) in pathology: current perspectives and future directions. J Digit Imaging. 2020;33(4):1034–40.
10. LeCun Y, Bengio Y, Hinton GJ. Deep learn. 2015;521(7553):436–44.
11. Yao J, Cao K, Hou Y, Zhou J, Xia Y, Nogues I, Song Q, Jiang H, Ye X, Lu J, et al. Deep learning for fully automated prediction of overall survival in patients undergoing resection for pancreatic cancer: a retrospective multicenter study. Ann Surg. 2023;278(1):e68–79.
12. Wang R, Dai W, Gong J, Huang M, Hu T, Li H, Lin K, Tan C, Hu H, Tong T, et al. Development of a novel combined nomogram model integrating deep learning-pathomics, radiomics and immunoscore to predict postoperative outcome of colorectal cancer lung metastasis patients. J Hematol Oncol. 2022;15(1):11.
13. She Y, Jin Z, Wu J, Deng J, Zhang L, Su H, Jiang G, Liu H, Xie D, Cao N, et al. Development and validation of a deep learning model for non-small cell lung cancer survival. JAMA Netw Open. 2020;3(6):e205842.
14. Shi J-Y, Wang X, Ding G-Y, Dong Z, Han J, Guan Z, et al. Exploring prognostic indicators in the pathological images of hepatocellular carcinoma based on deep learning. Gut. 2021;70(5):951–61.
15. Huang B, Tian S, Zhan N, Ma J, Huang Z, Zhang C, Zhang H, Ming F, Liao F, Ji M, et al. Accurate diagnosis and prognosis prediction of gastric cancer using deep learning on digital pathological images: a retrospective multicentre study. EBioMedicine. 2021;73:103631.
16. Wang Y, Ali MA, Vallon-Christersson J, Humphreys K, Hartman J, Rantalainen M. Transcriptional intra-tumour heterogeneity predicted by deep learning in routine breast histopathology slides provides independent prognostic information. Eur J Cancer. 2023;191:112953.
17. Cancer, AJCo. AJCC cancer staging manual. 9th ed. edn: Springer; 2021.
18. Yan J, Zhang S, Li KK-W, Wang W, Li K, Duan W, Yuan B, Wang L, Liu L, Zhan Y, et al. Incremental prognostic value and underlying biological pathways of radiomics patterns in medulloblastoma. EBioMedicine. 2020;61:103093.
19. Katz MHG, Hwang R, Fleming JB, Evans DB. Tumor-node-metastasis staging of pancreatic adenocarcinoma. CA Cancer J Clin. 2008;58(2):111–25.
20. Schouten TJ, Daamen LA, Dorland G, van Roessel SR, Groot VP, Besselink MG, Bonsing BA, Bosscha K, Brosens LAA, Busch OR, et al. Nationwide validation of the 8th American Joint Committee on Cancer TNM Staging System and five proposed modifications for resected pancreatic cancer. Ann Surg Oncol. 2022;29(9):5988–99.
21. Pointer DT, Roife D, Powers BD, Murimwa G, Elessawy S, Thompson ZJ, Schell MJ, Hodul PJ, Pimiento JM, Fleming JB, et al. Neutrophil to lymphocyte ratio, not platelet to lymphocyte or lymphocyte to monocyte ratio, is predictive of patient survival after resection of early-stage pancreatic ductal adenocarcinoma. BMC Cancer. 2020;20(1):750.
22. Zuo X, Kong W, Feng L, Zhang H, Meng X, Chen W. Elevated platelet distribution width predicts poor prognosis in hepatocellular carcinoma. Cancer Biomark. 2019;24(3):307–13.
23. Cui M-M, Li N, Liu X, Yun Z-Y, Niu Y, Zhang Y, Gao B, Liu T, Wang R-T. Platelet distribution width correlates with prognosis of non-small cell lung cancer. Sci Rep. 2017;7(1):3456.

24. Huang Y, Cui M-M, Huang Y-X, Fu S, Zhang X, Guo H, Wang R-T. Preoperative platelet distribution width predicts breast cancer survival. Cancer Biomark. 2018;23(2):205–11.
25. Hayashi M, Kobayashi D, Takami H, Inokawa Y, Tanaka N, Kurimoto K, Nakanishi K, Umeda S, Shimizu D, Hattori N, et al. Albumin-globulin ratio indicates the survival outcome of pancreatic cancer cases who underwent preoperative treatment and curative surgical resection. Nutr Cancer. 2023;75(5):1330–9.
26. Zhang Y, Zhu J-Y, Zhou L-N, Tang M, Chen M-B, Tao M. Predicting the prognosis of gastric cancer by albumin/globulin ratio and the prognostic nutritional index. Nutr Cancer. 2020;72(4):635–44.
27. Atsumi Y, Kawahara S, Kakuta S, Onodera A, Hara K, Kazama K, Numata M, Aoyama T, Tamagawa A, Tamagawa H, et al. Low preoperative albumin-to-globulin ratio is a marker of poor prognosis in patients with esophageal cancer. Vivo. 2021;35(6):3555–61.
28. Ni XG, Bai XF, Mao YL, Shao YF, Wu JX, Shan Y, Wang CF, Wang J, Tian YT, Liu Q, et al. The clinical value of serum CEA, CA19-9, and CA242 in the diagnosis and prognosis of pancreatic cancer. Eur J Surg Oncol. 2005;31(2):164–9.
29. Ermiah E, Eddfair M, Abdulrahman O, Elfagieh M, Jebriel A, Al-Sharif M, Assidi M, Buhmeida A. Prognostic value of serum CEA and CA19-9 levels in pancreatic ductal adenocarcinoma. Mol Clin Oncol. 2022;17(2):126.
30. Ahmadipour M, Bhattacharya A, Sarafbidabad M, Syuhada Sazali E, Krishna Ghoshal S, Satgunam M, Singh R, Rezaei Ardani M, Missaoui N, Kahri H, et al. CA19-9 and CEA biosensors in pancreatic cancer. Clin Chim Acta. 2024;554:117788.
31. O'Neill RS, Stoita A. Biomarkers in the diagnosis of pancreatic cancer: are we closer to finding the golden ticket? World J Gastroenterol. 2021;27(26):4045–87.
32. Kato H, Kishiwada M, Hayasaki A, Chipaila J, Maeda K, Noguchi D, Gyoten K, Fujii T, Iizawa Y, Tanemura A, et al. Role of serum Carcinoma Embryonic Antigen (CEA) level in localized pancreatic adenocarcinoma: CEA level before operation is a significant prognostic indicator in patients with locally advanced pancreatic cancer treated with neoadjuvant therapy followed by surgical resection: a retrospective analysis. Ann Surg. 2022;275(5):e698–707.
33. Zhou X, Lu Y, Wu Y, Yu Y, Liu Y, Wang C, Zhao Z, Wang C, Gao Z, Li Z, et al. Construction and validation of a deep learning prognostic model based on digital pathology images of stage III colorectal cancer. Eur J Surg Oncol. 2024;50(7):108369.
34. Huang B, Huang H, Zhang S, Zhang D, Shi Q, Liu J, Guo J. Artificial intelligence in pancreatic cancer. Theranostics. 2022;12(16):6931–54.
35. Karar ME, El-Fishawy N, Radad M. Automated classification of urine biomarkers to diagnose pancreatic cancer using 1-D convolutional neural networks. J Biol Eng. 2023;17(1):28.
36. Fu N, Fu W, Chen H, Chai W, Qian X, Wang W, Jiang Y, Shen B. A deep-learning radiomics-based lymph node metastasis predictive model for pancreatic cancer: a diagnostic study. Int J Surg (London England). 2023;109(8):2196–203.
37. Fu H, Mi W, Pan B, Guo Y, Li J, Xu R, Zheng J, Zou C, Zhang T, Liang Z, et al. Automatic pancreatic ductal adenocarcinoma detection in whole slide images using deep convolutional neural networks. Front Oncol. 2021;11:665929.
38. Lin J, Yin M, Liu L, Gao J, Yu C, Liu X, et al. The development of a prediction model based on random survival forest for the postoperative prognosis of pancreatic cancer: a SEER-based study. Cancers (Basel). 2022;14(19):4667.
39. Gheysarzadeh A, Ansari A, Emami MH, Razavi AE, Mofid MR. Over-expression of low-density lipoprotein receptor-related protein-1 is associated with poor prognosis and invasion in pancreatic ductal adenocarcinoma. Pancreatology. 2019;19(3):429–35.
40. Chen S, Hu S, Zhou B, Cheng B, Tong H, Su D, Li X, Chen Y, Zhang G. Telomere-related prognostic biomarkers for survival assessments in pancreatic cancer. Sci Rep. 2023;13(1):10586.
41. Fiocchi C. Omics and multi-omics in IBD: no integration, no breakthroughs. Int J Mol Sci. 2023;24(19).
42. Yang Z, Wang C, Wang Y, Lafata KJ, Zhang H, Ackerson BG, Kelsey C, Tong B, Yin FF. Development of a multi-feature-combined model: proof-of-concept with application to local failure prediction of post-SBRT or surgery early-stage NSCLC patients. Front Oncol. 2023;13:1185771.
43. Song R, Wang H, Zhang M, Liu Y, Meng X, Zhai S, Wang CC, Gong T, Wu Y, Jiang X, et al. Near-infrared light-triggered chlorine radical ((.) Cl) stress for cancer therapy. Angewandte Chemie (International ed English). 2020;59(47):21032–40.

Publisher's note