Tumour Marker Prognostic Study

INTERNATIONAL JOURNAL OF SURGERY

OPEN

Interpretable machine learning model for predicting early recurrence of pancreatic cancer: integrating intratumoral and peritumoral radiomics with body composition

Linxia Wu, MD, PhD[a,b,c], Chunyuan Cen, MD, PhD[d], Die Ouyang, MD[a,e], Licai Zhang, MD[a,e], Xin Li, MD, PhD[a,b,c], Heshui Wu, MD, PhD[f], Ming He, MD, PhD[d], Ping Han, MD, PhD[a,b,c], Wei Tan, MD, PhD[g,*], Lei Chen, MD, PhD[a,b,c,*], Chuansheng Zheng, MD, PhD[a,b,c,*]

Background: Pancreatic ductal adenocarcinoma (PDAC) is associated with a high rate of early recurrence (ER) after radical resection, which significantly affects long-term survival. Currently, no reliable system exists for predicting ER in these patients. This study aimed to develop a machine learning (ML) model combining intratumoral and peritumoral radiomic features with body composition to predict the ER risk in patients with PDAC following radical resection.

Materials and methods: This study included patients with PDAC who underwent upfront surgery at four hospitals between June 2014 and December 2023. Preoperative clinical information, computed tomography (CT) images, and postoperative pathological data were collected. CT-quantified body composition was measured; radiomic features were extracted from the intratumoral and peritumoral regions. Six ML algorithms were used to develop predictive models, including radiomics, clinical, clinical-radiomics, and clinicopathological-radiomics models. The SHapley Additive exPlanations (SHAP) method was applied for model interpretability.

Results: A total of 589 patients were evaluated, including 320 patients (mean age: 60.4 ± 8.3 years; 191 men) in the training cohort, 138 patients (mean age: 60.7 ± 8.9 years; 84 men) in the internal validation cohort, and 131 patients (mean age: 61.7 ± 10.9 years; 76 men) in the external validation cohort. The intra-peri-radiomics model, based on the random forest algorithm, achieved the best performance, with AUCs of 0.865, 0.849, and 0.839 in the training, internal validation, and external validation cohorts, respectively. Incorporating clinicopathological factors, the combined model showed superior performance, with AUCs of 0.936, 0.899, and 0.884 in the training, internal validation, and external validation cohorts, respectively. SHAP analysis revealed that radiomic features, adjuvant therapy, CA199, lymphovascular invasion, platelet-lymphocyte ratio, visceral fat index, CA125, visceral-subcutaneous fat tissue ratio, tumor size, and TNM stage significantly contributed to the prediction of ER.

Conclusion: The developed ML model, integrating radiomic features and clinicopathological factors, offered superior predictive accuracy for ER in patients with PDAC post-surgery. SHAP visualization enhanced the model's interpretability and facilitated clinical applications.

Keywords: body composition, early recurrence, machine learning, pancreatic ductal adenocarcinoma, radiomics, SHapley Additive exPlanations

Introduction

Pancreatic ductal adenocarcinoma (PDAC) is one of the most devastating malignancies, with a 5-year survival rate of only 12%[1]. Radical resection is the most effective curative treatment; however, only 20% of PDAC cases are resectable at the

[a]Department of Radiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China, [b]Hubei Provincial Clinical Research Center for Precision Radiology & Interventional Medicine, Wuhan, Hubei, China, [c]Hubei Key Laboratory of Molecular Imaging, Wuhan, Hubei, China, [d]Department of Radiology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Guangzhou, Guangdong, China, [e]Department of Radiology, Auto Valley Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China, [f]Department of Pancreatic Surgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China and [g]Department of Radiology, Geriatrics Hospital Affiliated to Wuhan University of Science and Technology, Wuhan, Hubei, China

Linxia Wu, Chunyuan Cen, and Die Ouyang contributed equally to this work.

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

*Corresponding authors. Address: Department of Radiology, Geriatrics Hospital Affiliated to Wuhan University of Science and Technology, Wuhan 430070, Hubei, China. E-mail: tanwei63317@163.com (W. Tan); Department of Radiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, Hubei, China. E-mail: chan0812@126.com (L. Chen); Tel.: +86 13329702158. E-mail: hqzcsxh@sina.com (C. Zheng).

International Journal of Surgery (2025) 111:8198–8211

Received 19 March 2025; Accepted 3 July 2025

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.lww.com/international-journal-of-surgery.

Published online 15 July 2025

http://dx.doi.org/10.1097/JS9.0000000000003078

HIGHLIGHTS

- The intra-peri-radiomics model, based on the random forest algorithm, demonstrated the best performance, with AUCs of 0.865, 0.849, and 0.839 in the training, internal validation, and external validation cohorts, respectively.
- The combined model, which incorporated clinicopathological factors and the intra-peri-radiomics model, showed superior performance, with areas under the receiver operating characteristic curves of 0.936, 0.899, and 0.884 in the training, internal validation, and external validation cohorts, respectively.
- SHapley Additive exPlanations analysis highlighted key predictors, including intra-peri-radiomics features, adjuvant therapy, CA199, and lymphovascular invasion.

time of diagnosis[2]. Despite radical resection, the 5-year survival rate remains suboptimal at approximately 20–30%[3,4]. The primary reason for this unfavorable postoperative outcome is the extremely high rate (50%) of early recurrence (ER), defined as recurrence within 1 year of curative resection, with a median recurrence-free survival of only 12 to 16 months[5-7]. Therefore, a reliable prognostic system is crucial for quantifying the risk of ER, as it can guide postoperative strategies and facilitate personalized management of high-risk patients with PDAC.

Computed tomography (CT) is the preferred imaging modality for diagnosing and evaluating PDAC[8]. However, traditional visual assessments often limit clinicians' ability to fully exploit the imaging features of PDAC. Recently, radiomics – a technique that analyzes imaging features extracted from regions of interest – has demonstrated significant effectiveness in predicting the biological complexity and intratumoral heterogeneity of tumors[9,10]. Through high-throughput analysis of intratumoral CT images, radiomics has been employed to predict the behavior of PDAC tumors[11–13]. Although most studies have focused on primary tumors, previous research has highlighted the significant role of the peritumoral region in clinically assessing cancerous heterogeneity and its importance in prognostic prediction[14–16]. In addition to tumor characteristics, individual differences among patients are critical factors that influence pancreatic cancer recurrence. Body composition is an important risk factor for recurrence and poor survival in PDAC, serving as an indicator of the patient's nutritional status and metabolism[17,18]. However, this easily accessible parameter is often overlooked during the development of prognostic models. To date, no study has combined intratumoral and peritumoral imaging features with CT-quantified body composition parameters to assess the prognosis and ER risk of pancreatic cancer.

Recently, machine learning (ML) algorithms, including random forests (RFs) and support vector machines (SVMs), have garnered widespread attention in the development of clinical predictive models. Compared with traditional linear models, ML models are capable of handling complex nonlinear relations and high-dimensional data, thereby improving predictive performance[19,20]. However, ML technology has some limitations. One notable limitation is the complexity and lack of interpretability of these models, which often leads to them being referred to as "black boxes"[21]. This study attempted to address this issue by employing the SHapley Additive exPlanations (SHAP) method to explain ML models and visualize the predictive contributions of each variable[22].

Therefore, this study aimed to develop and validate an ML model that integrated intratumoral and peritumoral CT radiomic features with body composition parameters to predict the risk of ER in patients with PDAC after radical resection. Furthermore, the inclusion of visualization tools enhanced the transparency of the ML model's decision-making process, increasing clinicians' trust in the predictions and aiding in the development of personalized treatment strategies.

Methods

Patients

Patients who underwent pancreatic surgery and were pathologically confirmed to have PDAC between June 2014 and December 2023 at four independent hospitals were enrolled (Fig. 1). The inclusion criteria were as follows:[1] curative surgery for resectable PDAC, as determined by a multidisciplinary team discussion among surgeons, oncologists, and radiologists, in line with the National Comprehensive Cancer Network guidelines for pancreatic cancer version 2.2024 (the resectable PDAC was defined as tumors based on preoperative CT with no major arterial contact and no contact or s180° contact with the portal vein or superior mesenteric vein without vein contour irregularity)[8]; and[2] pathologically confirmed PDAC. The exclusion criteria were as follows:[1] palliative or exploratory surgery;[2] preoperative neoadjuvant chemotherapy;[3] incomplete clinicopathological information;[4] other coexisting malignancies within the past 5 years;[5] poor quality or lack of preoperative CT examination; and[6] death within 90 days after surgery or loss to follow-up. The data from the derivation cohort comprising three separate hospitals (centers 1–3, June 2014–December 2023) were divided, with 70% utilized for training and 30% for internal validation, in order to avoid problems with overfitting. In addition, an external data set (center 4, June 2021–December 2023) was used for external validation. A flowchart of the patient recruitment process was shown in Supplementary Digital content, **Fig. S1**, Available at: http://links.lww.com/JS9/E783. This multicenter retrospective study was approved by the institutional ethics review board of each center; the requirement for written informed consent was waived. All procedures involving human participants were performed in accordance with the 1975 Declaration of Helsinki and its amendments. The work has been reported in line with the REMARK criteria[23]. Additionally, this study strictly adhered to the transparency principles of the TITAN 2025 Statement, ensuring the completeness and traceability of methodological descriptions[24].

Clinicopathological data

Demographic characteristics and medical history, including age, sex, BMI, smoking status, and history of diabetes, were extracted from electronic medical records. Tumor markers, including carbohydrate antigen (CA) 199, CA 125, carcinoembryonic antigen (CEA), white blood cell count, and platelet count, were measured preoperatively at least 2 weeks before surgery. Inflammation-based prognostic indices, including the

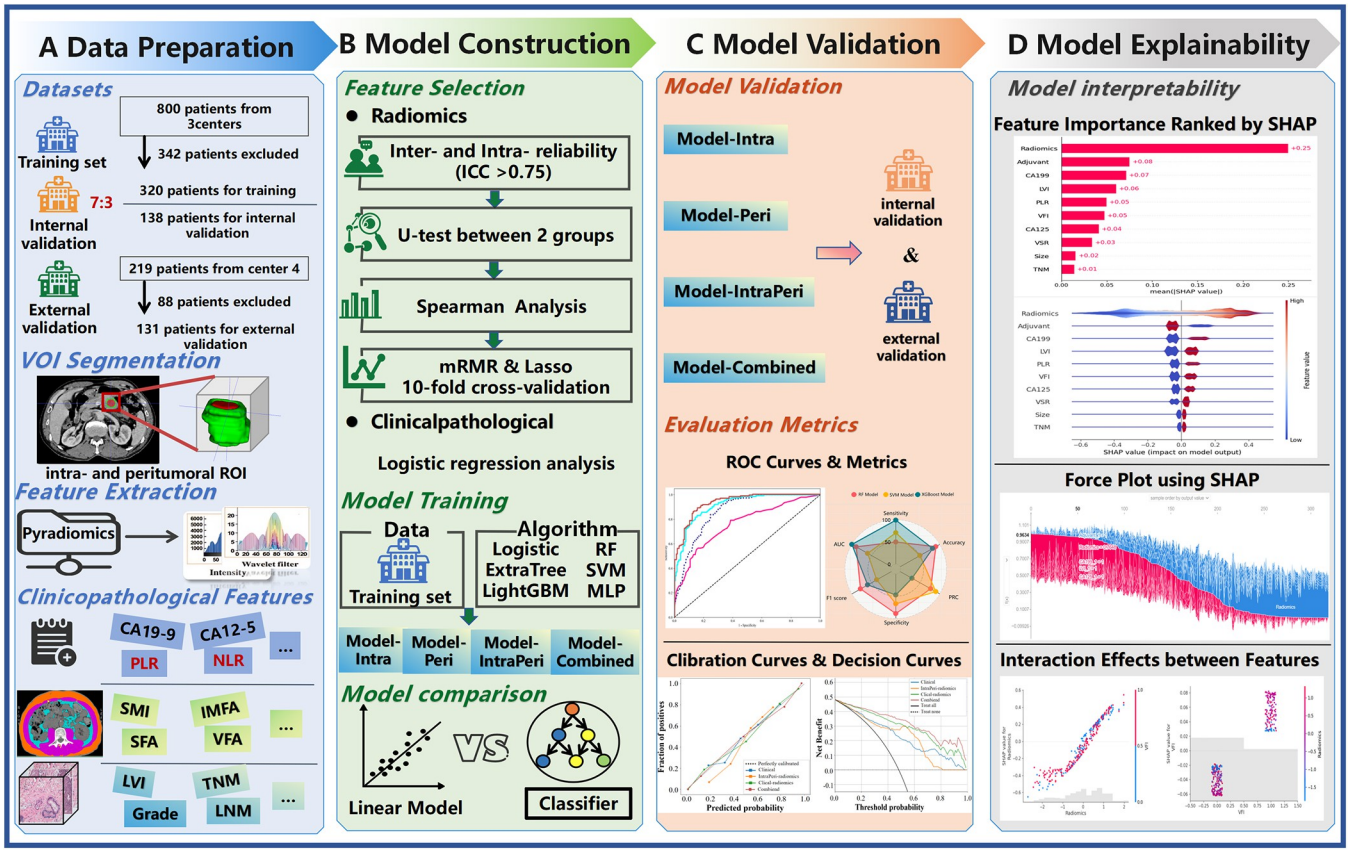

Figure 1. Study workflow overview. (A) Data Preparation: Incorporation of four datasets, delineation of intra- and peritumoral regions to define regions of interest (ROIs), extraction of intra- and peritumoral radiomic features, and collection of clinicopathological features. (B) Model Construction: Selection of radiomic features for intra-, peri-, and intra-peri-radiomics models, with the integration of clinical factors into predictive modeling using six machine learning algorithms. (C) Model Validation: Internal and external validation of intra-, peri-, and intra-peri-radiomics models. Model performance is evaluated through the area under the receiver operating characteristic curves (AUROC), calibration curves, and decision curve analysis. (D) Exploring Explainability: Application of the SHAP method for model interpretation, including analysis of feature importance, feature contributions to the model output, and interactions between radiomic and clinical features to elucidate the model's predictive dynamics.

platelet-lymphocyte ratio (PLR), neutrophil-lymphocyte ratio (NLR), lymphocyte-monocyte ratio (LMR), and prognostic nutrition index (PNI) [serum albumin (g/L) + 5 × total lymphocyte count (10^9/L)][25], were calculated. Tumor location and size were measured using preoperative CT; the presence of concurrent pancreatitis at diagnosis was confirmed using CT as well as clinical and laboratory examinations. Pathological findings from the surgical specimens were extracted from electronic medical records, including resection margin status (R0/R1), tumor staging according to the eighth edition of the American Joint Committee on Cancer staging system, and details of tumor differentiation, perineural invasion, and lymphovascular invasion (LVI).

Follow-up

Postoperative follow-up was initially conducted every 3 months during the first 2 years, every 6 months during years 3 and 4, and annually thereafter. The study was censored on 1 January 2025. The surveillance protocol included physical examinations, serum tumor marker assessments, and abdominal and pelvic contrast-enhanced CT (CE-CT) scans. Postoperative treatments were stratified into two groups:[1] with adjuvant therapy: patients who received chemotherapy (e.g., gemcitabine, S-1) or chemoradiotherapy (chemotherapy followed by radiotherapy with radiosensitizers when applicable);[2] without adjuvant therapy: patients who did not receive systemic chemotherapy or chemoradiotherapy due to poor performance status or patient preference. When imaging features suggested potential cancer recurrence, additional magnetic resonance imaging and/or fluorodeoxyglucose positron emission tomography (FDG-PET) scans were performed to further evaluate ambiguous CT findings. Recurrence was defined as a relapse confirmed by either radiological or pathological evidence. Radiological recurrence was characterized by the detection of new lesions in anatomical sites, including the surgical bed, lungs, liver, and peritoneum[26]. When imaging findings were consistent with recurrence, biopsy was rarely performed. The primary outcome was recurrence-free survival, defined as the time from surgery to recurrence or last follow-up, with ER defined as recurrence within 1 year after curative resection[5]. The secondary outcome was overall survival, defined as the interval from surgery to death or last follow-up.

Body composition measurements

Body composition parameters were measured from transverse non-contrast CT images obtained at the third lumbar vertebra level using SliceOMatic V5.0 (TomoVision, Canada) (Supplementary Digital content, **Fig. S2**, Available at: http://links.lww.com/JS9/E783). The measurements were based on tissue-specific Hounsfield unit (HU) thresholds (–29 to + 150 HU for skeletal muscle, –190 to –30 HU for subcutaneous and intermuscular fat, –150 to –50 HU for visceral fat)[27], as illustrated in Supplementary Digital content, **Fig. S2**, Available at: http://links.lww.com/JS9/E783. The cross-sectional areas of body composition were generated and normalized to height squared to derive indices, including the skeletal muscle index (SMI, cm^2/m^2), subcutaneous fat index (SFI, cm^2/m^2), intermuscular fat index (IMFI, cm^2/m^2), and visceral fat index (VFI, cm^2/m^2). The visceral-subcutaneous fat tissue ratio (VSR) was calculated by dividing the VFI by the SFI; the visceral fat-muscle tissue ratio (VMR) was calculated by dividing the VFI by the SMI. Skeletal muscle density (SMD), also known as muscle radiodensity, served as a proxy for muscle quality. The optimal cutoff values were determined through receiver operating characteristic (ROC) curve analysis using the maximum Youden index method. Given the significant sex-related differences in body composition, patients were categorized into high and low body composition groups according to sex-specific thresholds (Supplementary Digital content, **Table S1**, Available at: http://links.lww.com/JS9/E783). The cutoff values for age, tumor size, CA199, CA125, CEA, PLR, NLR, LMR, and PNI were also calculated.

Volume of interest segmentation and feature extraction

The 3D tumor volumes of interest (VOIs) were manually segmented and subsequently validated by two experienced radiologists using the open-source software ITK-SNAP (www.itksnap.org) on axial portal venous phase CT images. Following a previous study, the segmented masks were automatically expanded by 3 mm outward to define the peritumoral regions[15]. Prior to feature extraction, the CT images underwent preprocessing in the following sequence:[1] resampling to a standardized voxel size of $1.0 \times 1.0 \times 1.0$ mm^3;[2] discretization of gray levels into 25 fixed bins;[3] intensity normalization by scaling all CT images to a range of 1000 (additional preprocessing details were provided in the supplementary materials, Available at: http://links.lww.com/JS9/E783). Radiomic features were then extracted using PyRadiomics (version 3.8). A total of 1106 radiomic features were derived from both intratumoral and peritumoral VOIs, comprising 216 first-order features, 264 gray-level co-occurrence matrix (GLCM) features, 168 gray-level dependence matrix (GLDM) features, 192 gray-level run-length matrix (GLRLM) features, 192 gray-level size zone matrix (GLSZM) features, 60 neighboring gray-tone difference matrix (NGTDM) features, and 14 shape-based features.

Feature selection and radiomics signature construction

To ensure the reproducibility and reliability of radiomic features, two radiologists repeated the segmentation on 30 randomly selected cases after 2 weeks. Intra- and interobserver reproducibility was evaluated using intraclass correlation coefficients (ICCs). Features with good reproducibility (both ICCs ≥ 0.75) were retained for the subsequent analysis. To control the false discovery rate (FDR), the Benjamini–Hochberg (BH) procedure was applied to adjust the *P*-values from the Mann–Whitney U test. Only features with a statistically significant *q*-value (adjusted *P*-value) < 0.05 were retained. Subsequently, correlation analysis was performed, and features with a correlation coefficient >0.9 were excluded to reduce redundancy. To ensure the stability of features, maximum relevance and minimum redundancy (mRMR) with 1000-fold bootstrap resampling was conducted, and the top 20 features with the highest recurrence frequencies were selected (Supplementary Digital content, **Table S2**, Available at: http://links.lww.com/JS9/E783). The least absolute shrinkage and selection operator (LASSO) regression was then employed for dimensionality reduction, using 5-fold cross-validation to determine the optimal regularization parameter λ. The selected radiomic features were used to construct radiomics signatures. Six ML algorithms, including logistic regression (LR), SVM, RF, Extremely Randomized Trees (ExtraTrees), Light Gradient Boosting Machine (LightGBM), and multilayer perceptron (MLP), were evaluated to identify the best-performing classifier. A grid search with 5-fold cross-validation and manual fine-tuning was carried out to optimize the final model hyperparameters.

Clinical model and combined model construction

In addition to the radiomics model, clinical, clinical-radiomics, and combined clinicopathological-radiomics models were also developed. Clinical variables that were statistically significant in the univariate analysis were included in a multivariable logistic regression analysis to identify independent predictors of ER. The clinical model was constructed using these predictors via logistic regression. The clinical-radiomics and combined clinicopathological-radiomics models were constructed by integrating the independent predictors with radiomics signatures through SVM analysis. Model performance was evaluated by calculating the area under the receiver operating characteristic curve (AUC) for the training, internal validation, and external validation cohorts. Calibration curves were used to evaluate model goodness-of-fit, while decision curve analysis was employed to assess clinical utility across datasets.

Statistical analysis

Statistical analyses were conducted using R software (version 4.3.3) and Python (version 3.8). Continuous variables were presented as medians and interquartile ranges and compared between groups using the Mann–Whitney U test. The BH procedure was applied to adjust *P*-values and control the FDR. Categorical variables were reported as frequencies and percentages and compared using the chi-square test or Fisher's exact test, as appropriate. Logistic regression analysis was used to identify potential risk factors and to construct predictive models. All analyses were considered statistically significant with a two-sided *P*-value < 0.05.

Results

Clinicopathological characteristics

The clinical characteristics, CT-quantified body composition metrics, and postoperative data were summarized in Table 1.

Table 1

Patient characteristics of the training, internal validation, and external validation cohorts

Characteristics	Training cohort (*n* = 320)	Internal validation (*n* = 138)	External validation (*n* = 131)	*P* value
Age (years)				0.285
<60	149(46.56)	65(47.10)	51(38.93)	
≥60	171(53.44)	73(52.90)	80(61.07)	
Sex				0.891
Women	129(40.31)	54(39.13)	55(41.98)	
Men	191(59.69)	84(60.87)	76(58.02)	
BMI (kg/m^2)				0.218
<18.5	21(6.56)	7(5.07)	13(9.92)	
18.5-23.9	220(68.75)	86(62.32)	87(66.41)	
24-27.9	70(21.88)	37(26.81)	29(22.14)	
≥28	9(2.81)	8(5.80)	2(1.53)	
Smoking status				0.470
Smoker	229(71.56)	93(67.39)	97(74.05)	
Never smoker	91(28.44)	45(32.61)	34(25.95)	
Diabetes				0.405
Without	244(76.25)	97(70.29)	97(74.05)	
With	76(23.75)	41(29.71)	34(25.95)	
Pancreatitis at diagnosis				0.976
Without	288(90.00)	124(89.86)	117(89.31)	
With	32(10.00)	14(10.14)	14(10.69)	
Tumor location				0.552
Head and neck	240(75.00)	110(79.71)	100(76.34)	
Body and tail	80(25.00)	28(20.29)	31(23.66)	
Tumor size (cm)				0.673
<3	162(50.62)	76(55.07)	67(51.15)	
≥3	158(49.38)	62(44.93)	64(48.85)	
Adjuvant therapy				0.147
No	109(34.06)	60(43.48)	51(38.93)	
Yes	211(65.94)	78(56.52)	80(61.07)	
Resection margin				0.937
R0	276(86.25)	118(85.51)	114(87.02)	
R1	44(13.75)	20(14.49)	17(12.98)	
Lymphovascular invasion				0.348
Absent	173(54.06)	74(53.62)	80(61.07)	
Present	147(45.94)	64(46.38)	51(38.93)	
Perineural invasion				0.391
Absent	22(6.88)	12(8.70)	14(10.69)	
Present	298(93.12)	126(91.30)	117(89.31)	
Tumor grade				0.848
Well or moderate	222(69.38)	92(66.67)	90(68.70)	
Poor	98(30.63)	46(33.33)	41(31.30)	
Tumor TNM stage				0.308
Stage ≤ IIA	147(45.94)	63(45.65)	70(53.44)	
IIA < stage < IV	173(54.06)	75(54.35)	61(46.56)	
CA199 (U/ml)	201(62.81)	84(60.87)	83(63.36)	0.899
<472.2	119(37.19)	54(39.13)	48(36.64)	
≥472.2				
CA125 (U/ml)				
<18.9	181(56.56)	78(56.52)	75(57.25)	0.958
≥18.9	139(43.44)	60(43.48)	56(42.75)	
CEA (μg/L)				0.848
<3.2	167(52.19)	70(50.72)	71(54.20)	
≥3.2	153(47.81)	68(49.28)	60(45.80)	
PLR				0.879
<160.58	201(62.81)	90(65.22)	84(64.12)	
≥160.58	119(37.19)	48(34.78)	47(35.88)	
NLR				0.221
<2.33	135(42.19)	65(47.10)	48(36.64)	
≥2.33	185(57.81)	73(52.90)	83(63.36)	
LMR				0.153
<4.78	233(72.81)	111(80.43)	103(78.63)	
≥4.78	87(27.19)	27(19.57)	28(21.37)	
PNI index				0.532
<48.55	244(76.25)	108(78.26)	95(72.52)	
≥48.55	76(23.75)	30(21.74)	36(27.48)	
VFI (cm^2/m^2)				0.956
Low	176(55.00)	77(55.80)	74(56.49)	
High	144(45.00)	61(44.20)	57(43.51)	
SFI (cm^2/m^2)				0.521
Low	129(40.31)	62(44.93)	59(45.04)	
High	191(59.69)	76(55.07)	72(54.96)	
SMI (cm^2/m^2)				0.927
Low	182(56.88)	80(57.97)	77(58.78)	
High	138(43.12)	58(42.03)	54(41.22)	
IMFI (cm^2/m^2)				0.423
Low	82(25.62)	35(25.36)	41(31.30)	
High	238(74.38)	103(74.64)	90(68.70)	
SMD (HU)				0.714
Low	168(52.50)	78(56.52)	69(52.67)	
High	152(47.50)	60(43.48)	62(47.33)	
VSR				0.582
Low	137(42.81)	60(43.48)	63(48.09)	
High	183(57.19)	78(56.52)	68(51.91)	
VMR				0.886
Low	137(42.81)	57(41.30)	53(40.46)	
High	183(57.19)	81(58.70)	78(59.54)	
Early recurrence				0.712
No	166(51.88)	69(50.00)	72(54.96)	
Yes	154(48.12)	69(50.00)	59(45.04)	
Follow-up (months)	50.7 (44.7, 54.8)	51.5 (44.7, 55.3)	29 (19.8,35.5)	<0.001

Unless otherwise indicated, data are the number of patients, with percentages in parentheses. BMI, body mass index; CA19-9, carbohydrate antigen 19-9; CA12-5, carbohydrate antigen 12-5; CEA, carcinoembryonic antigen; LMR, lymphocyte to monocyte ratio; NLR, neutrophil to lymphocyte ratio; PLR, platelet to lymphocyte ratio; PNI, Prognostic Nutritional Index; IMFI, intermuscular fat index; SFI, subcutaneous fat index; SMD, skeletal muscle density; SMI, skeletal muscle index; VFI, visceral fat index; VSR: VFA-to-SFA ratio; VMR: VFA-to-SMA ratio.

The training, internal validation, and external validation cohorts comprised 320 patients (mean age: 60.4 ± 8.3 years; 191 men), 138 patients (mean age: 60.7 ± 8.9 years; 84 men), and 131 patients (mean age: 61.7 ± 10.9 years; 76 men), respectively. The follow-up duration in the external validation cohort was significantly shorter than that in the training and internal validation cohorts ($P < 0.001$). There were no statistically significant differences in other clinicopathological variables among the three groups. During the follow-up, ER occurred in 48.12% (154/320), 50.0% (69/138), and 45.04% (59/131) of patients in the training, internal validation, and external validation cohorts, respectively ($P = 0.712$). Univariate analysis showed that preoperative clinical variables – including tumor size, CA199,

Table 2

Univariate and multivariable analysis in the training cohort

	Univariate analysis		Multivariable analysis	
Characteristics	**OR (95%CI)**	***P* value**	**OR (95%CI)**	***P* value**
Age ≥ 60 years	0.766 (0.493-1.19)	0.235		
Sex (Male)	1.526 (0.973-2.395)	0.066		
BMI (vs. < 18.5 kg/m^2)				
18.5-23.9 kg/m^2	1.333 (0.54-3.292)	0.533		
24-27.9 kg/m^2	0.943 (0.352-2.529)	0.907		
≥ 28 kg/m^2	2.667 (0.521-9.656)	0.239		
Smoking	0.952 (0.586-1.549)	0.844		
Diabetes	1.103 (0.659-1.847)	0.708		
Pancreatitis at diagnosis	1.437 (0.689-2.999)	0.334		
Tumor location	0.905 (0.545-1.502)	0.698		
Tumor size ≥3 cm	**2.497 (1.592-3.916)**	**<0.001**	**2.013(0.76-2.335)**	**0.015**
Adjuvant therapy	**0.289 (0.177-0.471)**	**<0.001**	**0.265(0.745-2.48)**	**<0.001**
Resection margin (R1)	2.083 (1.079-4.024)	0.029	1.763(0.664-3.535)	0.184
Lymphovascular invasion	**3.34 (2.109-5.29)**	**<0.001**	**2.46(0.755-2.381)**	**0.002**
Perineural invasion	0.837 (0.499-1.403)	0.499		
Tumor grade (Poor)	2.434 (1.491-3.972)	<0.001	1.669(0.737-2.563)	0.107
Tumor TNM stage ≥ IIA	**2.359 (1.503-3.704)**	**<0.001**	**2.28(0.754-2.388)**	**0.005**
CA199 ≥ 472.2 U/ml	**3.738 (2.311-6.045)**	**<0.001**	**2.668(0.748-2.446)**	**0.001**
CA125 ≥ 18.9 U/ml	**2.696 (1.71-4.253)**	**<0.001**	**1.962(0.755-2.378)**	**0.021**
CEA ≥ 3.2 μg/L	1.687 (1.084-2.626)	0.021	1.041(0.752-2.408)	0.891
PLR ≥ 160.58	**1.599 (1.013-2.524)**	**0.044**	**2.218(0.747-2.456)**	**0.009**
NLR ≥ 2.33	1.432 (0.916-2.238)	0.115		
LMR ≥ 4.78	0.604 (0.366-0.997)	0.049	0.713(0.732-2.617)	0.297
PNI index ≥ 48.55	0.837 (0.499-1.403)	0.499		
VFI (High group)	2.234 (1.425-3.501)	<0.001	2.25(0.703-2.963)	0.027
SFI (High group)	0.735 (0.469-1.15)	0.177		
SMI (High group)	0.84 (0.539-1.309)	0.441		
IMFI (High group)	0.904 (0.547-1.494)	0.694		
SMD (High group)	0.992 (0.64-1.54)	0.973		
VSR (High group)	**2.168(0.723-2.715)**	**0.022**	**2.168(0.723-2.715)**	**0.022**
VMR (High group)	1.757 (1.122-2.752)	0.014	0.645(0.674-3.368)	0.285

BMI, body mass index; CA19-9, carbohydrate antigen 19-9; CA12-5, carbohydrate antigen 12-5; CEA, carcinoembryonic antigen; LMR, lymphocyte to monocyte ratio; NLR, neutrophil to lymphocyte ratio; PLR, platelet to lymphocyte ratio; PNI, Prognostic Nutritional Index; IMFI, intermuscular fat index; SFI, subcutaneous fat index; SMD, skeletal muscle density; SMI, skeletal muscle index; VFI, visceral fat index; VSR: VFA-to-SFA ratio; VMR: VFA-to-SMA ratio.

CA125, CEA, PLR, LMR, VFI, VSR, and VMR – as well as pathological factors such as adjuvant therapy, resection margin, LVI, tumor grade, and TNM stage were significantly associated with ER in PDAC (all $P < 0.05$; Table 2). Multivariate analysis identified tumor size, adjuvant therapy, LVI, TNM stage, CA199, CA125, PLR, VFI, and VSR as independent risk factors for ER (all $P < 0.05$; Table 2).

Feature selection and radiomics signature building

A total of 978 out of 1106 intratumoral and 934 out of 1106 peritumoral radiomic features demonstrated good reproducibility, with both inter- and intra-observer ICCs > 0.75. Following feature selection using mRMR and LASSO, 9 intratumoral, 11 peritumoral, and 13 intra-peritumoral features were retained to construct intra-, peri-, and intra-peri-radiomics signatures, respectively, using six ML algorithms (LR, SVM, RF, ExtraTrees, LightGBM, and MLP). The LASSO feature selection process and the corresponding feature weight coefficients were presented in the Supplementary Digital content, **Figs. S3** and **S4**, Available at: http://links.lww.com/JS9/E783.

Model performance comparison

The performance of the intra-, peri-, and intra-peri-radiomics models was evaluated using six ML algorithms. As illustrated in Fig. 2 and 3 and Supplementary Digital content, **Tables S3** and **S4**, Available at: http://links.lww.com/JS9/E783, the intra-peri-radiomics model consistently outperformed the intra- and peri-radiomics models across all cohorts. Among the six algorithms, the RF algorithm demonstrated the best performance when applied to the intra-peri-radiomics model, achieving AUCs of 0.865 (95% CI: 0.825–0.904), 0.849 (95% CI: 0.783–0.915), and 0.839 (95% CI: 0.770–0.907) in the training, internal validation, and external validation cohorts, respectively. To further validate the robustness of this model, 5-fold cross-validation was conducted. As shown in Supplementary Digital content, **Fig. S5**, Available at: http://links.lww.com/JS9/E783, the final RF model achieved a mean AUC of 0.870 ± 0.030. At a threshold of 0.450, it yielded high accuracy (ACC) values of 0.791, 0.812, and 0.771; sensitivities of 0.922, 0.841, and 0.847; specificities of 0.669, 0.783, and 0.708; positive predictive values (PPVs) of 0.721, 0.795, and 0.704, and negative predictive values (NPVs) of 0.902, 0.831, and 0.850, respectively (Table 3;

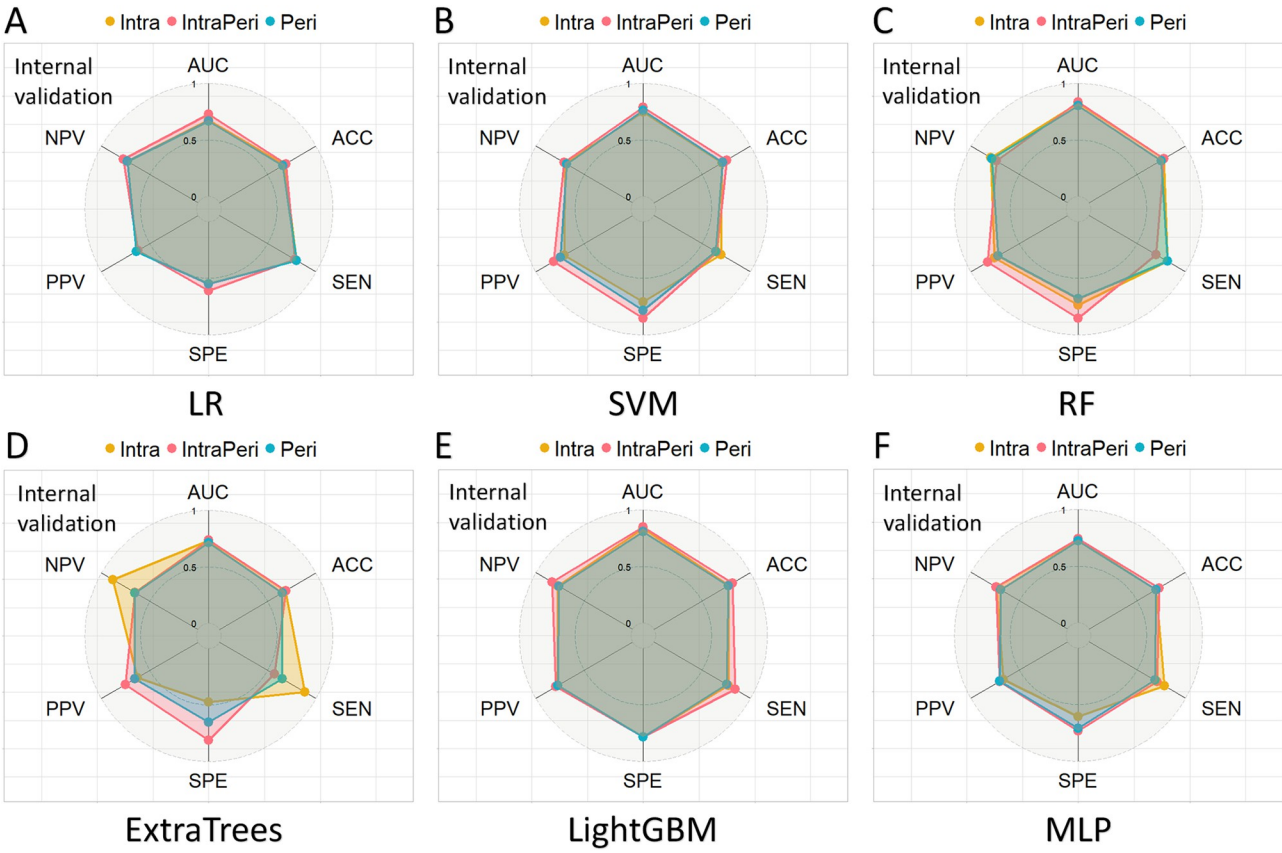

Figure 2. Radar charts demonstrating the performance comparison of intra-, peri-, and intra-peri-radiomics models across six machine learning algorithms in the internal validation cohort. (A) LR algorithm, (B) SVM algorithm, (C) RF algorithm, (D) ExtraTrees algorithm, (E) LightGBM algorithm, and (F) MLP algorithm. AUC, area under the receiver operating characteristic curve; ACC, accuracy; SEN, sensitivity; SPE, specificity; PPV, positive predictive value; NPV, negative predictive value; LR, logistic regression; SVM, support vector machine; RF, random forest; ExtraTrees, extremely randomized trees; LightGBM, light gradient boosting machine; MLP, multilayer perceptron.

Supplementary Digital content, **Fig. S6**, Available at: http://links.lww.com/JS9/E783). Given its superior performance, the RF-based intra-peri-radiomics model was selected for constructing subsequent combined models. A clinical model incorporating independent clinical risk factors (including tumor size, CA199, CA125, PLR, VFI, VSR) achieved AUCs of 0.760 (95% CI: 0.707–0.812), 0.719 (95% CI: 0.634–0.804) and 0.721 (95% CI: 0.635–0.807) in the training, internal validation, and external validation cohorts, respectively (Table 4). A clinical-radiomics model integrating the intra-peri-radiomics model with clinical variables demonstrated improved performance, with AUCs of 0.915 (95% CI: 0.886–0.945), 0.886 (95% CI: 0.831–0.940), and 0.859 (95% CI: 0.796–0.921) across the three cohorts. Furthermore, a comprehensive model integrating the intra-peri-radiomics with clinicopathological data achieved the highest performance, yielding AUCs of 0.936 (95% CI: 0.912–0.961), 0.899 (95% CI: 0.847–0.951), 0.884 (95% CI: 0.828–0.939), respectively. DeLong tests and integrated discrimination improvement (IDI) analyses confirmed that the combined model outperformed the other three models in terms of discrimination (Supplementary Digital content, $P < 0.05$; **Fig. S7**, Available at: http://links.lww.com/JS9/E783). Calibration curves demonstrated excellent agreement between predicted and observed ER probabilities in the training ($P = 0.674$), internal validation ($P = 0.372$), and external validation ($P = 0.217$, Fig. 4A-C) cohorts. In contrast, the clinical model exhibited poor calibration in the external validation cohort ($P = 0.013$), highlighting the robustness of the combined model. Decision curve analysis further indicated that the combined model offered the greatest net benefits across a wide range of threshold probabilities (0.0–1.0 in the training cohort, 0.0–0.85 in the internal validation cohort, and 0.0–0.8 in the external validation cohort), consistently outperforming other models.

Model interpretability

The SHAP method was employed to interpret model outputs by quantifying the contribution of each variable to the prediction. This explainable approach provided two levels of insights: global interpretation at the feature level and local interpretation at the individual level. For global interpretation, SHAP summary plots (**Figs. S8–S10**) illustrated the importance of each radiomic feature in the radiomics model. In the RF-based intra-peri-radiomics model (Supplementary Digital content, **Fig. S10C**, Available at: http://links.lww.com/JS9/E783), features such as peri-wavelet-HHH-ngtdm-Strength and intra-log-sigma-2-0-mm-3D-firstorder-90-Percentile contributed most significantly to the model. In the combined model, the SHAP bar plot (Fig. 5A) showed that

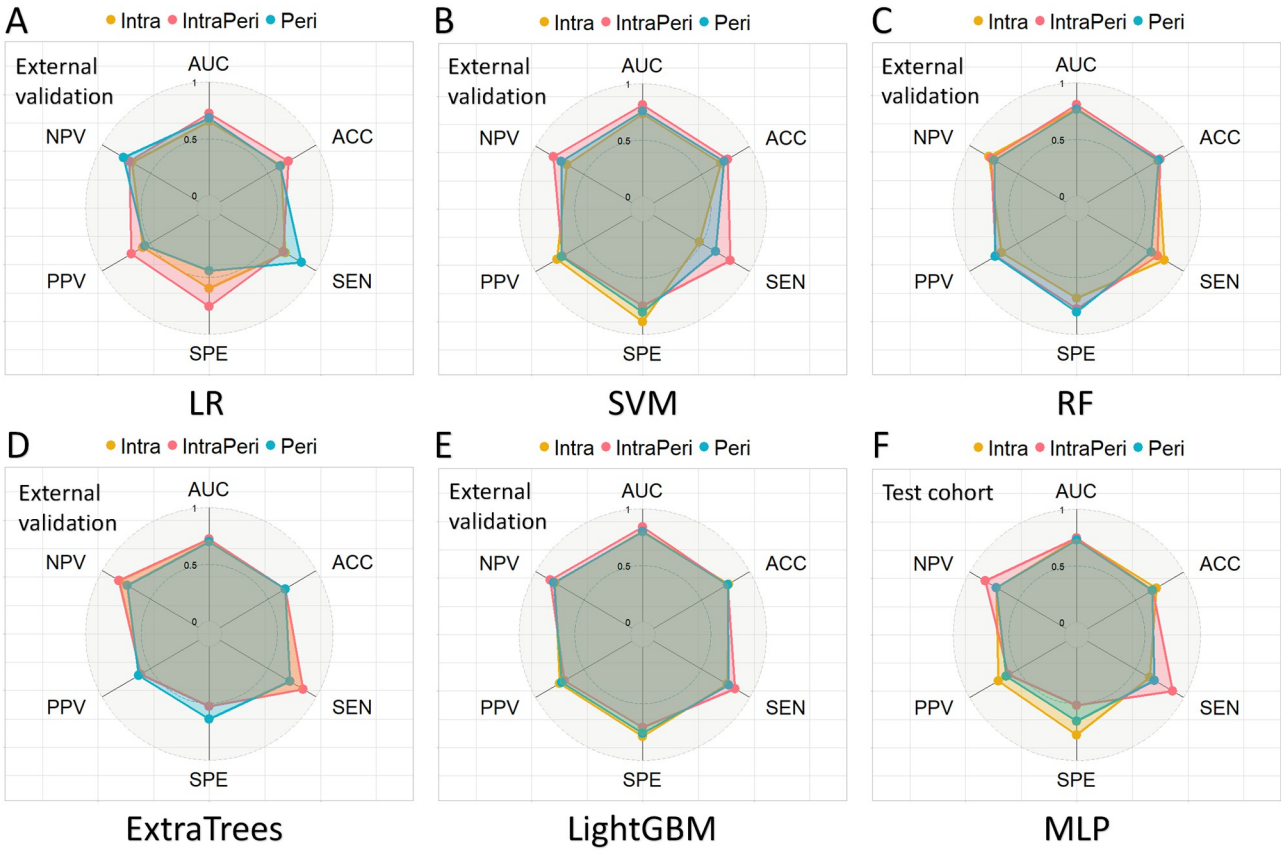

Figure 3. Radar charts demonstrating the performance comparison of intra-, peri-, and intra-peri-radiomics models across six machine learning algorithms in the external validation cohort. (A) LR algorithm, (B) SVM algorithm, (C) RF algorithm, (D) ExtraTrees algorithm, (E) LightGBM algorithm, and (F) MLP algorithm. AUC, area under the receiver operating characteristic curve; ACC, accuracy; SEN, sensitivity; SPE, specificity; PPV, positive predictive value; NPV, negative predictive value; LR, logistic regression; SVM, support vector machine; RF, random forest; ExtraTrees, extremely randomized trees; LightGBM, light gradient boosting machine; MLP, multilayer perceptron.

radiomic features had the greatest contribution to the model output, with an average SHAP value of 0.25. The violin plot (Fig. 5B) and dependence plot (Supplementary Digital content, **Fig. S11A**, Available at: http://links.lww.com/JS9/E783) revealed a clear upward trend in SHAP values with increasing radiomic feature values, indicating a strong association between higher radiomic values and an increased ER risk. Similarly, the absence of adjuvant therapy, high-level CA199, positive LVI, high-level PLR, high-level VFI, high-level CA125, high-level VSR, larger tumor size, and TNM stage ≥ IIA were all associated with an increased ER risk (Supplementary Digital content, **Fig. S10B–J**, Available at: http://links.lww.com/JS9/E783). Moreover, the interaction between radiomics, adjuvant therapy, and VFI, as indicated by the color gradient in Fig. 5C and D, suggested that patients with higher radiomic feature values, who received adjuvant therapy and had elevated VFI, were at higher risk of recurrence.

For local interpretation, Fig 6A and B presents a case of a patient who did not experience ER of PDAC during the follow-up period. According to the prediction model, this patient was classified into the "ER" category with a predicted probability of 18.0%, which was significantly lower than the baseline probability of 50.1%. The individual contributions of each feature were illustrated in the corresponding waterfall plots (Fig. 6A and B). As shown, radiomic features, adjuvant therapy, low-level CA199, TNM ≤ IIA, smaller tumor size, and low-level CA125 contributed to the prediction of "non-ER," whereas the presence of LVI, high-level PLR, high VFI, and high VSR did not. Fig. 6C and D presents another case of a patient who experienced ER. According to the model, this patient was classified into the "ER" category with a predicted probability of 99.9%, markedly higher than the baseline probability. Fig. 6E displays a force plot for ER prediction. The x-axis represented individual patients, and the y-axis denoted feature contributions. A larger red segment for a patient indicated a higher predicted probability of ER.

Discussion

In this study, we developed and validated an ML prediction model that integrated both intra- and peritumoral radiomic features to predict the risk of ER after radical resection of PDAC. The model incorporating both intra- and peritumoral radiomic features outperformed the models using only intra- or peritumoral radiomic features. Furthermore, integrating radiomics signatures with clinicopathological factors–including tumor size, CA199, CA125, PLR, VFI, VSR, TNM stage, LVI, and adjuvant therapy–further enhanced predictive accuracy. Notably, model interpretability was facilitated through SHAP,

Table 3

Comparison of the performance of intra-peri-radiomics ML models

Models		AUC	ACC	SEN	SPE	PPV	NPV
Training set	LR	0.737 (0.683-0.790)	0.688	0.740	0.639	0.655	0.726
	SVM	0.834 (0.790-0.879)	0.784	0.825	0.747	0.751	0.821
	RF	0.865 (0.825-0.904)	0.791	0.922	0.669	0.721	0.902
	ExtraTrees	0.756 (0.703-0.808)	0.716	0.779	0.657	0.678	0.762
	LightGBM	0.853 (0.811-0.894)	0.787	0.883	0.699	0.731	0.866
	MLP	0.766 (0.715-0.817)	0.691	0.805	0.584	0.642	0.764
Internal validation	LR	0.730 (0.644-0.816)	0.687	0.780	0.611	0.622	0.772
	SVM	0.791 (0.714-0.867)	0.754	0.652	0.855	0.818	0.711
	RF	0.849 (0.783-0.915)	0.812	0.841	0.783	0.795	0.831
	ExtraTrees	0.738 (0.657-0.820)	0.688	0.569	0.812	0.750	0.653
	LightGBM	0.834 (0.764-0.904)	0.775	0.696	0.855	0.828	0.737
	MLP	0.740 (0.655-0.824)	0.725	0.703	0.734	0.703	0.743
External validation	LR	0.723 (0.637-0.810)	0.710	0.656	0.757	0.700	0.718
	SVM	0.811 (0.737-0.886)	0.771	0.797	0.750	0.723	0.818
	RF	0.839 (0.770-0.907)	0.771	0.847	0.708	0.704	0.850
	ExtraTrees	0.723 (0.635-0.810)	0.678	0.864	0.528	0.600	0.826
	LightGBM	0.810 (0.735-0.884)	0.756	0.729	0.778	0.729	0.778
	MLP	0.746 (0.662-0.830)	0.679	0.881	0.514	0.598	0.841

AUC, area under the receiver operating characteristic curve; ACC, accuracy; SEN, sensitivity; SPE, specificity; PPV, positive predictive value; NPV, negative predictive value; LR, logistic regression; SVM, support vector machine; RF, random forest; ExtraTrees, extremely randomized trees; LightGBM, Light Gradient Boosting Machine; MLP, Multilayer Perceptron

which visually elucidated the contribution of each variable to the predicted outcome.

Models combining both intra- and peritumoral features outperformed those relying on any single group of features in predicting PDAC[15,16,28]. This advantage stemmed from the fact that tumor heterogeneity exists not only within the tumor but is also closely associated with the surrounding microenvironment[29,30]. While conventional imaging techniques may fail to directly capture the tumor microenvironment, radiomics, through quantitative analysis, can detect subtle changes and provide important complementary information regarding tumor heterogeneity[31]. Additionally, this study compared the predictive performance of multiple ML algorithms, including LR, SVM, RF, ExtraTrees, LightGBM, and MLP. Among them, the RF-based intra-peri-radiomics model exhibited the best predictive capability, achieving AUCs of 0.865 in the training cohort, 0.849 in the internal validation cohort, and 0.839 in the external validation cohort. As an ensemble learning technique, the RF algorithm can effectively handle high-dimensional data and has demonstrated superior performance across many predictive modeling studies[32–34].

Several studies have focused on predicting ER after surgery in patients with PDAC. For instance, Lee *et al*[35] developed a nomogram incorporating radiomic features, preoperative CA199 levels, and CT-derived tumor characteristics from 190 patients, achieving AUCs of 0.77 and 0.83 in the training and validation cohorts, respectively, showing good predictive performance. Similarly, Li *et al*[28] constructed a nomogram using radiomic features from both intra- and peritumoral regions, along with preoperative CA199, in 220 patients to predict recurrence risk within 1 year after surgery, with AUCs of 0.916 and 0.764. However, these studies primarily focused on preoperative clinical and imaging variables and did not incorporate postoperative pathological information, which was crucial for predicting ER in PDAC. By integrating preoperative tumor radiomic features, body composition metrics, and postoperative pathological variables, our study further improved predictive performance, outperforming previous models across training, internal validation, and external validation cohorts. The innovations of our study compared with previous research are as follows:[1] it considered not only imaging features of the tumor itself but also individual-level differences, particularly the impact of body composition and postoperative pathology on ER in PDAC;[2] multiple ML algorithms were employed to comprehensively explore the complex relationships between predictors and outcomes, thereby improving model accuracy and flexibility;[3] the SHAP methodology was used to enhance model interpretability, providing more intuitive insights into the contribution of each feature to prediction and supporting individualized clinical decision-making.

In addition to radiomics, body composition reflected a patient's nutritional and metabolic status and had been identified as an important factor influencing cancer recurrence and survival outcomes[17,18]. High visceral fat was associated with ER and poor prognosis in several cancers, including PDAC[17,36–39]. Our findings supported this observation, demonstrating that patients with high visceral fat were more likely to experience recurrence within 1 year after surgery. This association may be attributed to the accumulation of visceral fat promoting pancreatic fat infiltration, which altered the tumor microenvironment and facilitated tumor spread[40–42]. High VSR was another independent risk factor for poor prognosis, consistent with a previous report[17]. This may relate to the differing physiological functions and tumor-regulating roles of visceral and subcutaneous fat. Visceral fat secreted pro-inflammatory cytokines, such as interleukin-6 and tumor necrosis factor-α, which modulated the tumor microenvironment and promoted immune evasion[43,44], whereas subcutaneous fat primarily served as an energy reserve, supporting metabolic function and facilitating postoperative recovery[45]. Therefore, an elevated VSR indicated increased visceral fat and decreased subcutaneous fat, predisposing patients to ER. Additionally, LVI, TNM stage, and adjuvant therapy were key determinants of PDAC recurrence[5,46], justifying their inclusion in the predictive model. Although

Table 4

Comparison of the performance of different models in training, internal validation, and external validation cohorts

Models		AUC	ACC	SEN	SPE	PPV	NPV
Training	Clinical	0.760 (0.707-0.812)	0.688	0.760	0.620	0.650	0.736
	Intra-Peri-radiomics	0.865 (0.825-0.904)	0.791	0.922	0.669	0.721	0.902
	Clinical-radiomics	0.915 (0.886-0.945)	0.834	0.825	0.843	0.830	0.838
	Combined	0.936 (0.912-0.961)	0.863	0.883	0.843	0.840	0.886
Internal validation	Clinical	0.719 (0.634-0.804)	0.667	0.710	0.623	0.653	0.683
	Intra-Peri-radiomics	0.849 (0.783-0.915)	0.812	0.841	0.783	0.795	0.831
	Clinical-radiomics	0.886 (0.831-0.940)	0.814	0.884	0.739	0.772	0.864
	Combined	0.899 (0.847-0.951)	0.848	0.855	0.841	0.843	0.853
External validation	Clinical	0.721 (0.635-0.807)	0.672	0.898	0.486	0.589	0.854
	Intra-Peri-radiomics	0.839 (0.770-0.907)	0.771	0.847	0.708	0.704	0.850
	Clinical-radiomics	0.859 (0.796-0.921)	0.794	0.712	0.861	0.808	0.785
	Combined	0.884 (0.828-0.939)	0.809	0.746	0.861	0.815	0.805

AUC, area under the receiver operating characteristic curve; ACC, accuracy; SEN, sensitivity; SPE, specificity; PPV, positive predictive value; NPV, negative predictive value.

postoperative adjuvant chemotherapy was the standard treatment for PDAC[8], approximately 30-40% of patients in this study did not receive it, resulting in a significantly higher rate of ER. Multiple studies have confirmed that adjuvant chemotherapy effectively delays recurrence and improves overall survival[47,48], further highlighting its importance in the predictive model. Including adjuvant therapy in the model helped identify high-risk patients requiring intensified postoperative care and facilitated the formulation of individualized treatment strategies. CA199 was a well-established biomarker for PDAC recurrence and poor prognosis, and our study has reaffirmed its predictive value. Notably, we also found that CA125 was significantly associated with ER in PDAC and was included in the ML model. Unlike CA199, which reflected total tumor burden, CA125 may indicate "metastasis-associated burden"[49,50]. Elevated preoperative CA125 levels may signal the presence of subclinical metastases undetectable on conventional imaging, thereby increasing recurrence risk after surgery[51]. This could explain why some patients relapse within 1 year despite curative resection. Among inflammation-based prognostic markers, including LMR, NLR, and PLR, only PLR was significantly correlated with ER in our cohorts. Shirai similarly reported that elevated preoperative PLR was associated with both poor overall survival and ER in PDAC[52]. Finally, the integrated model combining radiomic features with clinicopathological indicators showed

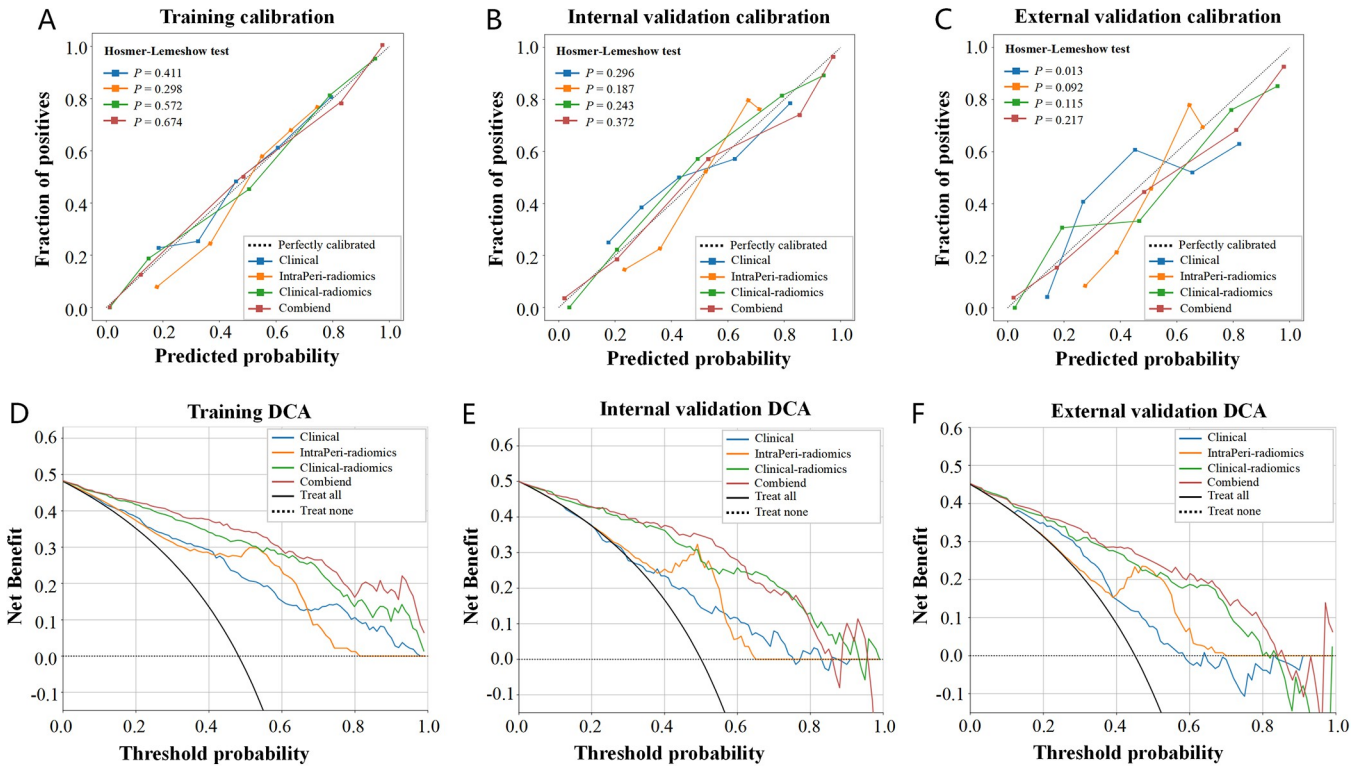

Figure 4. Performance comparison of clinical, radiomics, clinical-radiomics, and combined models for predicting recurrence in PDAC patients. (A-C) Calibration curves for different models in the training, internal validation, and external validation sets, respectively. (D-F) Decision curves for different models in the training, internal validation, and external validation sets, respectively.

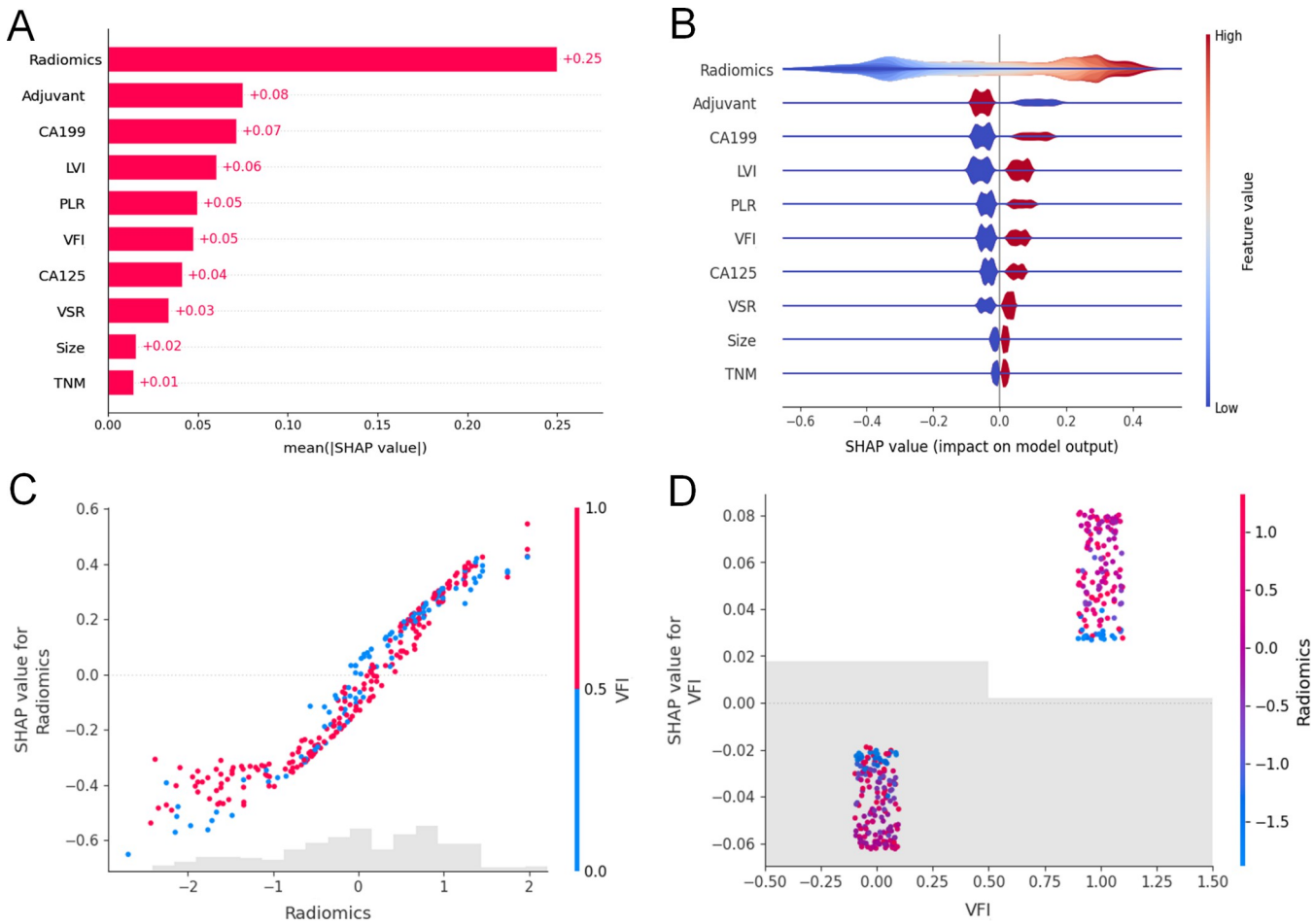

Figure 5. Global model explanation by the SHAP method. (A) Bar plot showing the mean absolute SHAP values for each feature, indicating the overall importance of features in predicting tumor recurrence in PDAC patients. (B) Violin plot illustrating the distribution of SHAP values for each feature. The color gradient represents the feature values (red for high values and blue for low values). (C) Dependence plot showing the relationship between the radiomic feature and its SHAP value, with adjuvant as a color-coded interaction feature. The x-axis represents the radiomics feature value, and the y-axis represents its SHAP value. (D) Dependence plot showing the relationship between the VFI feature and its SHAP value, with radiomics as a color-coded interaction feature. The x-axis represents the VFI feature value, and the y-axis represents its SHAP value.

excellent performance, with AUCs of 0.936, 0.899, and 0.884 in the training, internal validation, and external validation cohorts, respectively. This highlighted the prognostic value of integrating multimodal data. The model demonstrated favorable clinical utility, identifying high-risk patients while avoiding overtreatment at low thresholds (nd external validation cohorts, ciated wit intermediate-to-high thresholds (0.4–0.8). Similarly, the clinical-radiomics model exhibited robust discriminative ability and positive net benefits. Overall, the clinical-radiomics and the combined models served complementary roles across the treatment timeline: the former supported preoperative planning, while the latter enabled postoperative risk stratification and adjuvant therapy optimization, together forming a comprehensive decision-support system throughout the patient's treatment cycle.

The interpretability of ML models was crucial for their clinical application. In this study, we utilized the SHAP values to quantify the contribution of each feature to the model's predictions. The SHAP analysis revealed that intra-peritumoral radiomics signature was the most influential predictor of ER, followed by adjuvant therapy, CA199, LVI, PLR, VFI, CA125, VSR, tumor size, and TNM staging. Moreover, the model's local interpretability emphasized its potential to inform personalized treatment strategies. By understanding the specific impact of each variable, clinicians can better assess recurrence risk and make informed decisions about postoperative surveillance and adjuvant therapy. The clinical utility of the model is further demonstrated through two representative case studies. In Case 1 (Fig. 6A and B), the predicted probability of ER was 18.0%, substantially lower than the baseline probability of 50.1%, supporting the continuation of standard treatment and routine follow-up. This approach not only facilitated effective monitoring but also minimized unnecessary interventions. In contrast, Case 2 (Fig. 6C and D) had a predicted ER probability of 99.9%, far exceeding the baseline, suggesting the need for more postoperative management. Strategies such as intensifying chemotherapy and strengthening the monitoring of tumor markers should be implemented, underscoring the significance of such individualized measures in optimizing the clinical care of patients with PDAC.

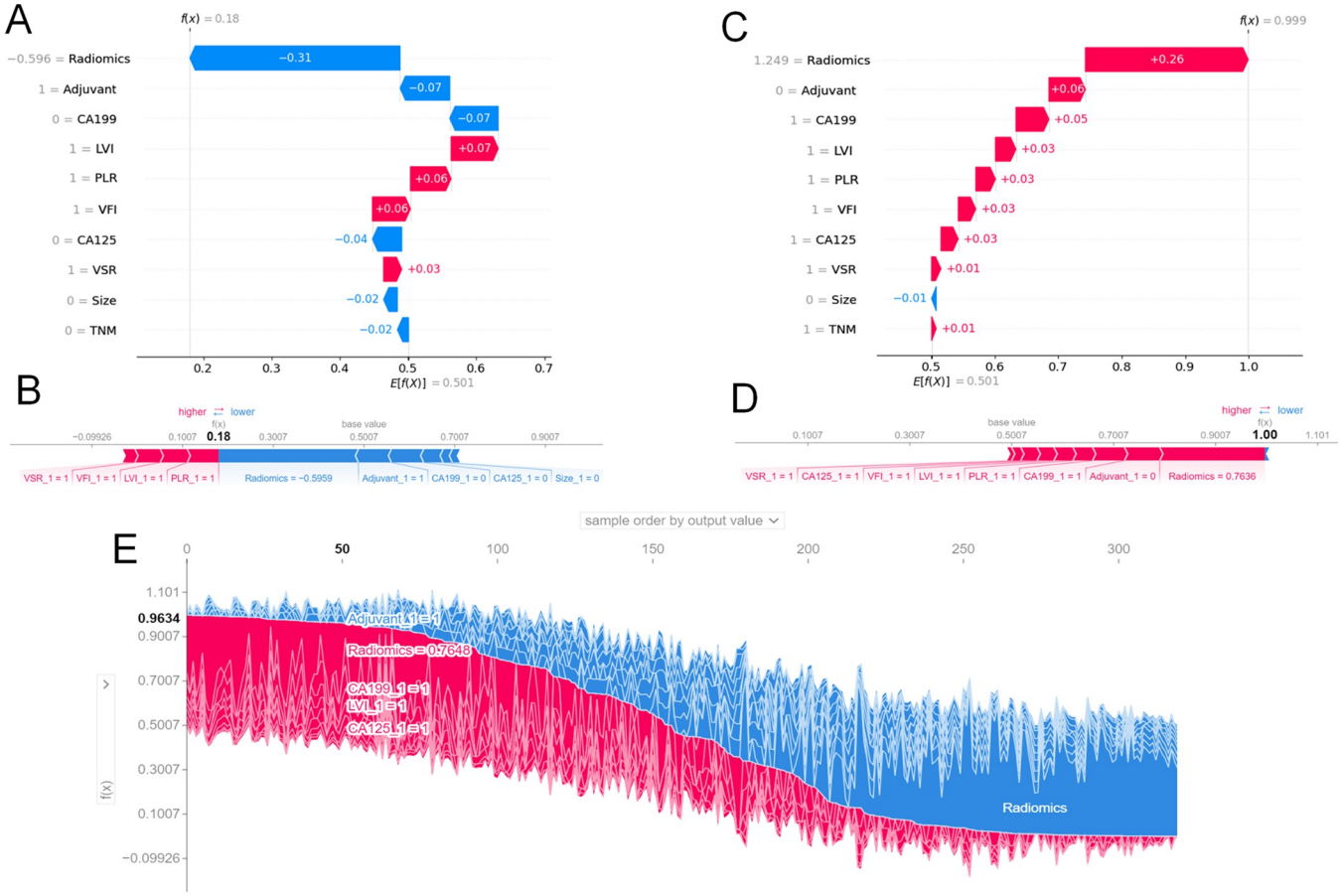

Figure 6. Local model explanation by the SHAP method. (A and B) Waterfall and force plots showing the evolution of risks contributed by each feature for individuals with PDAC at low risk of tumor recurrence. (C and D) Waterfall and force plots showing the evolution of risks contributed by each feature for individuals with PDAC at high risk of tumor recurrence. The base value (gray vertical line) represents the average model output, while the arrows indicate how each feature contributes to pushing the prediction [f(x)] toward either the recurrence (red) or non-recurrence (blue) class. (E) Force plot for the training set. Each patient is represented on the x-axis, while the magnitude of features' contributions is represented by the red (recurrence) and blue (non-recurrence) segments: a larger red portion for an individual indicates a higher probability of recurrence.

This study had several limitations. First, its retrospective design may introduce selection bias. Second, radiomic features extracted from CT images reflected tumor characteristics and body composition at a single time point, without accounting for temporal changes during treatment. Third, although the SHAP method improved model interpretability, the biological relevance of radiomic features in PDAC recurrence remained unclear. Future research should refine the model by incorporating additional data modalities, such as genomic and proteomic information, to provide deeper insights into the mechanisms underlying recurrence.

Conclusion

This study demonstrated the feasibility and clinical utility of integrating intratumoral and peritumoral radiomic features with CT-derived body composition parameters to predict ER in patients with PDAC. The RF-based intra-peritumoral radiomics model outperformed traditional clinical predictors and provided a promising tool for personalized management. The interpretability of the model, enabled by SHAP analysis, enhanced its clinical applicability and supported its potential integration into routine clinical practice. Further prospective validation and model refinement are warranted to confirm its efficacy and generalizability across diverse patient populations.

Ethical approval

This study was approved by the Institutional Review Board of each center. All procedures involving human participants were performed in accordance with the 1975 Declaration of Helsinki and its amendments.

Consent

Written informed consent was waived by the Institutional Review Board.

Sources of funding

This study was supported by the National Key Research and Development Program of China (2023YFC2413500 to C.Z.), the National Natural Science Foundation of China (Grant no. 82402411 to L.C., U22A20352 to C.Z., 82101996 to M.

H.), the Fundamental Research Funds for the Central Universities (YCJJ20252425 to L.W.), and Science and Technology Projects in Guangzhou (2024A04J5048 to C.C.).

Author contributions

Conceptualization: C.Z., L.C., P.H.; Data curation: D.O., L.Z.; Formal analysis: L.W., C.C.; Funding acquisition: C.Z., L.C., M.H., L.W., C.C.. Investigation: X.L.; Methodology: L.W., C. C.; M.H.; Project administration: C.Z.; Resources: H.W.; X.L.; Supervision: C.Z., L.C., P.H.; Visualization: L.W., C.C.; Writing – original draft: L.W.; Writing – review & editing: C. Z., L.C., P.H., and W.T.

Conflicts of interest disclosure

The authors have no relevant financial or non-financial interests to disclose.

Research registration unique identifying number (UIN)

The study was registered in the Chinese Clinical Trial Registry. The registration number: ChiCTR2300079198. Hyperlink to the registration: https://www.chictr.org.cn/showproj.html?proj=214941.

Guarantor

Chuansheng Zheng.

Provenance and peer review

Not commissioned, externally peer-reviewed.

Data availability statement

The data used in the study can be available from the corresponding author (Chuansheng Zheng, hqzcsxh@sina.com).

Acknowledgements

We would like to thank all colleagues for helping us during the current study, and all the selfless volunteers who participated in the study.

References

[1] Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. CA Cancer J Clin 2023;73:17–48.
[2] Park W, Chawla A, O'Reilly EM. Pancreatic cancer: a review. Jama 2021;326:851–62.
[3] Strobel O, Lorenz P, Hinz U, *et al*. Actual five-year survival after upfront resection for pancreatic ductal adenocarcinoma: who beats the odds? Ann Surg 2022;275:962–71.
[4] Strobel O, Neoptolemos J, Jäger D, Büchler MW. Optimizing the outcomes of pancreatic cancer surgery. Nat Rev Clin Oncol 2019;16:11–26.
[5] Groot VP, Gemenetzis G, Blair AB, *et al*. Defining and predicting early recurrence in 957 patients with resected pancreatic ductal adenocarcinoma. Ann Surg 2019;269:1154–62.
[6] Leonhardt CS, Gustorff C, Klaiber U, *et al*. Prognostic factors for early recurrence after resection of pancreatic cancer: a systematic review and meta-analysis. Gastroenterology 2024;167:977–92.
[7] van Goor I, Schouten TJ, Verburg DN, *et al*. Predicting long-term disease-free survival after resection of pancreatic ductal adenocarcinoma: a nationwide cohort study. Ann Surg 2024;279:132–37.
[8] Daly MB, Pal T, Maxwell KN, *et al*. NCCN guidelines® insights: genetic/familial high-risk assessment: breast, ovarian, and pancreatic, version 2.2024. J Natl Compr Canc Netw 2023;21:1000–10.
[9] Aerts HJ, Velazquez ER, Leijenaar RT, *et al*. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
[10] Lambin P, Leijenaar RTH, Deist TM, *et al*. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749–62.
[11] Bian Y, Zheng Z, Fang X, *et al*. Artificial intelligence to predict lymph node metastasis at CT in pancreatic ductal adenocarcinoma. Radiology 2023;306:160–69.
[12] Li Q, Song Z, Li X, *et al*. Development of a CT radiomics nomogram for preoperative prediction of Ki-67 index in pancreatic ductal adenocarcinoma: a two-center retrospective study. Eur Radiol 2024;34:2934–43.
[13] Tikhonova VS, Karmazanovsky GG, Kondratyev EV, *et al*. Radiomics model-based algorithm for preoperative prediction of pancreatic ductal adenocarcinoma grade. Eur Radiol 2023;33:1152–61.
[14] Xie N, Fan X, Chen D, *et al*. Peritumoral and intratumoral texture features based on multiparametric MRI and multiple machine learning methods to preoperatively evaluate the pathological outcomes of pancreatic cancer. J Magn Reson Imaging 2023;58:379–91.
[15] Zhang X, Gao A, Ma L, Yu N. Integrating intratumoral and peritumoral radiomics with clinical risk factors for prognostic prediction in pancreatic ductal adenocarcinoma patients undergoing combined chemotherapy and HIFU ablation. Int J Hyperthermia 2024;41:2410342.
[16] Miao Q, Wang X, Cui J, *et al*. Artificial intelligence to predict T4 stage of pancreatic ductal adenocarcinoma using CT imaging. Comput Biol Med 2024;171:108125.
[17] Okumura S, Kaido T, Hamaguchi Y, *et al*. Visceral adiposity and sarcopenic visceral obesity are associated with poor prognosis after resection of pancreatic cancer. Ann Surg Oncol 2017;24:3732–40.
[18] van Dijk DP, Bakens MJ, Coolsen MM, *et al*. Low skeletal muscle radiation attenuation and visceral adiposity are associated with overall survival and surgical site infections in patients with pancreatic cancer. J Cachexia, Sarcopenia Muscle 2017;8:317–26.
[19] Greener JG, Kandathil SM, Moffat L, Jones DT. A guide to machine learning for biologists. Nat Rev Mol Cell Biol 2022;23:40–55.
[20] Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. J Intern Med 2018;284:603–19.
[21] Azodi CB, Tang J, Shiu S-H. Opening the black box: interpretable machine learning for geneticists. Trends Genet 2020;36:442–55.
[22] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Proceedings of the 31st International Conference on Neural Information Processing Systems; 2017.
[23] Harris AL. REporting recommendations for tumour MARKer prognostic studies (REMARK). Br J Cancer 2005;93:385–86.
[24] Agha RA, Mathew G, Rashid R, *et al*. Transparency in the reporting of artificial intelligence – the TITAN guideline. Premier Journal of Science 2025;10:100082.
[25] Hackner D, Merkel S, Weiß A, *et al*. Neutrophil-to-lymphocyte ratio and prognostic nutritional index are predictors for overall survival after primary pancreatic resection of pancreatic ductal adenocarcinoma: a single centre evaluation. Cancers (Basel) 2024;16:2911.
[26] Groot VP, Rezaee N, Wu W, *et al*. Patterns, timing, and predictors of recurrence following pancreatectomy for pancreatic ductal adenocarcinoma. Ann Surg 2018;267:936–45.
[27] Peng Y-C, Wu C-H, Tien Y-W, Lu T-P, Wang Y-H, Chen -B-B. Preoperative sarcopenia is associated with poor overall survival in pancreatic cancer patients following pancreaticoduodenectomy. Eur Radiol 2021;31:2472–81.
[28] Li X, Wan Y, Lou J, *et al*. Preoperative recurrence prediction in pancreatic ductal adenocarcinoma after radical resection using radiomics of diagnostic computed tomography. EClinicalMedicine 2022;43:101215.
[29] Fyfe J, Dye D, Razak NBA, Metharom P, Falasca M. Immune evasion on the nanoscale: small extracellular vesicles in pancreatic ductal adenocarcinoma immunity. Semin Cancer Biol 2023;96:36–47.

[30] Wu C, Liu Y, Wei D, *et al*. Gene coexpression network characterizing microenvironmental heterogeneity and intercellular communication in pancreatic ductal adenocarcinoma: implications of prognostic significance and therapeutic target. Front Oncol 2022;12:840474.
[31] Zhao B, Cao B, Xia T, *et al*. Multiparametric MRI for assessment of the biological invasiveness and prognosis of pancreatic ductal adenocarcinoma in the era of artificial intelligence. J Magn Reson Imaging 2025;62:9–19.
[32] Hu J, Xu J, Li M, *et al*. Identification and validation of an explainable prediction model of acute kidney injury with prognostic implications in critically ill children: a prospective multicenter cohort study. EClinical Medicine 2024;68:102409.
[33] Hassan AM, Biaggi AP, Asaad M, *et al*. Development and assessment of machine learning models for individualized risk assessment of mastectomy skin flap necrosis. Ann Surg 2023;278:e123–e30.
[34] Chen X, Ishwaran H. Random forests for genomic data analysis. Genomics 2012;99:323–29.
[35] Lee JH, Shin J, Min JH, *et al*. Preoperative prediction of early recurrence in resectable pancreatic cancer integrating clinical, radiologic, and CT radiomics features. Cancer Imaging 2024;24:6.
[36] Kwon MR, Ko ES, Park MS, *et al*. Impact of skeletal muscle loss and visceral obesity measured using serial CT on the prognosis of operable breast cancers in Asian patients. Korean J Radiol 2022;23:159–71.
[37] Iyer K, Ren S, Pu L, *et al*. A graph-based approach to identify factors contributing to postoperative lung cancer recurrence among patients with non-small-cell lung cancer. Cancers (Basel) 2023;15:3472.
[38] Lee JW, Yoo ID, Hong SP, *et al*. Prognostic impact of visceral adipose tissue imaging parameters in patients with cholangiocarcinoma after surgical resection. Int J Mol Sci 2024;25:3939.
[39] Lee CW, Tsai HI, Hsu HM, *et al*. Do low skeletal muscle bulk and disturbed body fat mass impact tumor recurrence in stage I/II hepatocellular carcinoma undergoing surgery? An observational cohort study. Int J Surg 2024;110:7067–79.
[40] Mathur A, Zyromski NJ, Pitt HA, *et al*. Pancreatic steatosis promotes dissemination and lethality of pancreatic cancer. J Am Coll Surg 2009;208:989–94.
[41] Targher G, Rossi AP, Zamboni GA, *et al*. Pancreatic fat accumulation and its relationship with liver fat content and other fat depots in obese individuals. J Endocrinol Invest 2012;35:748–53.
[42] White PB, True EM, Ziegler KM, *et al*. Insulin, leptin, and tumoral adipocytes promote murine pancreatic cancer growth. J Gastrointest Surg 2010;14:1888–93.
[43] Kirchgessner TG, Uysal KT, Wiesbrock SM, Marino MW, Hotamisligil GS. Tumor necrosis factor-alpha contributes to obesity-related hyperleptinemia by regulating leptin release from adipocytes. J Clin Invest 1997;100:2777–82.
[44] O'Sullivan J, Lysaght J, Donohoe CL, Reynolds JV. Obesity and gastrointestinal cancer: the interrelationship of adipose and tumour microenvironments. Nat Rev Gastroenterol Hepatol 2018;15:699–714.
[45] Hwang I, Jo K, Shin KC, *et al*. GABA-stimulated adipose-derived stem cells suppress subcutaneous adipose inflammation in obesity. Proc Natl Acad Sci U S A 2019;116:11936–45.
[46] Lee KS, Jang JY, Yu YD, *et al*. Usefulness of artificial intelligence for predicting recurrence following surgery for pancreatic cancer: retrospective cohort study. Int J Surg 2021;93:106050.
[47] Hayashi T, Nakamura T, Kimura Y, *et al*. Phase 2 study of neoadjuvant treatment of sequential S-1-based concurrent chemoradiation therapy followed by systemic chemotherapy with gemcitabine for borderline resectable pancreatic adenocarcinoma (HOPS-BR 01). Int J Radiat Oncol Biol Phys 2019;105:606–17.
[48] Sinn M, Bahra M, Liersch T, *et al*. CONKO-005: adjuvant chemotherapy with gemcitabine plus erlotinib versus gemcitabine alone in patients after R0 resection of pancreatic cancer: a multicenter randomized phase III trial. J Clin Oncol 2017;35:3330–37.
[49] Liu L, Xu HX, Wang WQ, *et al*. Serum CA125 is a novel predictive marker for pancreatic cancer metastasis and correlates with the metastasis-associated burden. Oncotarget 2016;7:5943–56.
[50] Ge J, Li L, Ma Z, *et al*. A nomogram of preoperative predictors for occult metastasis in patients with PDAC during laparoscopic exploration. Gland Surg 2021;10:279–89.
[51] Luo G, Liu C, Guo M, *et al*. Potential biomarkers in Lewis negative patients with pancreatic cancer. Ann Surg 2017;265:800–05.
[52] Shirai Y, Shiba H, Sakamoto T, *et al*. Preoperative platelet to lymphocyte ratio predicts outcome of patients with pancreatic ductal adenocarcinoma after pancreatic resection. Surgery 2015;158:360–65.