Experimental Research

INTERNATIONAL JOURNAL OF SURGERY

OPEN

Development and validation of a deep learning radiomics model with clinical-radiological characteristics for the identification of occult peritoneal metastases in patients with pancreatic ductal adenocarcinoma

Siya Shi, MD[a], Chuxuan Lin, MS[f,c], Jian Zhou, MD[b,e], Luyong Wei, BS[a], Mingjie chen, BS[a], Jian Zhang, PhD[g,d], Kangyang Cao, MS[f,c], Yaheng Fan, MS[f,c], Bingsheng Huang, PhD[f,c,g,*], Yanji Luo, MD, PhD[a,*], Shi-Ting Feng, MD, PhD[a,*]

Background: Occult peritoneal metastases (OPM) in patients with pancreatic ductal adenocarcinoma (PDAC) are frequently overlooked during imaging. The authors aimed to develop and validate a computed tomography (CT)-based deep learning-based radiomics (DLR) model to identify OPM in PDAC before treatment.

Methods: This retrospective, bicentric study included 302 patients with PDAC (training: $n = 167$, OPM-positive, $n = 22$; internal test: $n = 72$, OPM-positive, $n = 9$: external test, $n = 63$, OPM-positive, $n = 9$) who had undergone baseline CT examinations between January 2012 and October 2022. Handcrafted radiomics (HCR) and DLR features of the tumor and HCR features of peritoneum were extracted from CT images. Mutual information and least absolute shrinkage and selection operator algorithms were used for feature selection. A combined model, which incorporated the selected clinical-radiological, HCR, and DLR features, was developed using a logistic regression classifier using data from the training cohort and validated in the test cohorts.

Results: Three clinical-radiological characteristics (carcinoembryonic antigen 19-9 and CT-based T and N stages), nine HCR features of the tumor, 14 DLR features of the tumor, and three HCR features of the peritoneum were retained after feature selection. The combined model yielded satisfactory predictive performance, with an area under the curve (AUC) of 0.853 (95% CI: 0.790–0.903), 0.845 (95% CI: 0.740–0.919), and 0.852 (95% CI: 0.740–0.929) in the training, internal test, and external test cohorts, respectively (all $P < 0.05$). The combined model showed better discrimination than the clinical-radiological model in the training (AUC = 0.853 vs. 0.612, $P < 0.001$) and the total test (AUC = 0.842 vs. 0.638, $P < 0.05$) cohorts. The decision curves revealed that the combined model had greater clinical applicability than the clinical-radiological model.

Conclusions: The model combining CT-based DLR and clinical-radiological features showed satisfactory performance for predicting OPM in patients with PDAC.

Keywords: computed tomography, deep learning radiomics, occult peritoneal metastases, pancreatic ductal adenocarcinoma

---

[a]Department of Radiology, The First Affiliated Hospital, Sun Yat-sen University, [b]State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Department of Radiology, Sun Yat-sen University Cancer Center, Guangzhou, [c]Medical AI Lab, School of Biomedical Engineering, [d]Shenzhen University Medical School, [e]South China Hospital, Medical School, Shenzhen University, [f]Marshall Laboratory of Biomedical Engineering, Shenzhen University and [g]Shenzhen-Hong Kong Institute of Brain Science-Shenzhen Fundamental Research Institutions, Shenzhen, People's Republic of China

Siya Shi, Chuxuan Lin, and Jian Zhou contributed equally to this manuscript.

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

**Corresponding author. Address: Department of Radiology, The First Affiliated Hospital, Sun Yat-sen University, No. 58, Second Zhongshan Road, Yuexiu District, Guangzhou, Guangdong 510080, People's Republic of China. Tel.: +86 208 775 5766, EXT. 8471; fax: +86 208 761 5805. E-mail: fengsht@mail.sysu.edu.cn (S.-T. Feng), and E-mail: luoyj26@mail.sysu.edu.cn (Y. Luo) ; Marshall Laboratory of Biomedical Engineering, Shenzhen University, 3688 Nanhai Avenue, Nanshan District, Shenzhen, Guangdong 518055, People's Republic of China. Tel.: + 86 0755 86172208; fax: +86 86171620. E-mail: huangb@szu.edu.cn (B. Huang).*

International Journal of Surgery (2024) 110:2669–2678

Received 26 November 2023; Accepted 5 February 2024

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.

http://dx.doi.org/10.1097/JS9.0000000000001213

Introduction

The poor prognosis of pancreatic ductal adenocarcinoma (PDAC), with a 5-year survival rate of ~11%[1], may be attributed in part to 80–85% of patients being diagnosed with advanced disease, which manifests as unresectable or metastatic disease, at presentation[2]. Peritoneum is the second common pathway for the metastasis of PDAC, which is observed in ~10–20% of patients at presentation[3,4] and one-third of patients die of recurrence following potentially curative resection[5]. Postoperative complications and delayed chemotherapy may worsen the prognosis of patients with occult peritoneal metastases (OPM)[6]. Thus, identifying the presence of peritoneal metastases before treatment can help in selecting appropriate treatments and avoiding unnecessary surgery.

Computed tomography (CT) is the most commonly used modality for the preoperative diagnosis of PDAC. However, early peritoneal metastasis is often overlooked owing to the lack of visible signs. OPM is confirmed during subsequent laparoscopy in up to 34% of patients with locally unresectable tumors without radiological peritoneal metastases[7]. Diagnostic staging laparoscopy is recommended for the diagnosis of peritoneal metastasis of PDAC. However, a recent cost-benefit analysis suggested that invasive diagnostic staging laparoscopy is not cost-effective unless used selectively[8]. The factors that improve patient selection for diagnostic staging laparoscopy remain unclear. Some clinical and radiological features, such as carcinoembryonic antigen 19-9 (CA19-9) and regional lymph node enlargement, are predictors of OPM[9]; however, their accuracy for predicting OPM has not been verified by other centers. Therefore, there is a need to develop an individualized model for predicting OPM in patients with PDAC.

Artificial intelligence techniques, such as radiomics and deep learning (DL) methods, are remarkable tools for predicting OPM in gastric cancer[10,11], colorectal cancer[12], and epithelial ovarian cancer[13]. Huang *et al.*[14] proposed an effective deep learning-based radiomics (DLR) framework, which could automatically extract effective deep semantic features of the lesions from a DL segmentation network and integrate information using a feature selection module. The DLR framework incorporating radiomics and DL algorithms overcomes the shortcomings of overfitting and has demonstrated advanced performance in some studies[14–18], such as the differentiation of pancreatic neuroendocrine tumors. The handcrafted radiomics (HCR) and DLR features are promising tools for predicting OPM in PDAC; however, these features have not been used to identify the OPM status in patients with PDAC.

Therefore, this study aimed to develop and validate a model that combines HCR and DLR features of the tumor, HCR features of the peritoneum with clinical-radiological characteristics to predict the presence of OPM before treatment in patients with PDAC. We hypothesized that the incorporation of HCR and DLR features would help to predict the status of the OPM, and the combined model would show superior performance compared to the model based only clinical-radiological characteristics.

Materials and methods

Study population

This retrospective study was approved by the Institutional Review Board of our Hospital, and the requirement for obtaining informed consent from the patients was waived. Patients with suspected pancreatic tumors who had undergone contrast-enhanced CT and pathological examinations were screened at Center 1 between January 2012 and October 2022 and Center 2 between January 2014 and October 2021.

HIGHLIGHTS

- To our best knowledge, this is the first bicentric, deep learning radiomics study to identify occult peritoneal metastases in patients with pancreatic ductal adenocarcinoma before treatment.
- The performance of our noninvasive tool, which combines deep learning-based radiomics features and clinical-radiological variables, was significantly outperformed that of the clinical-radiological model in training and test cohorts, and can be used as a supplement to current risk stratification strategies for patients with pancreatic ductal adenocarcinoma.

The inclusion criteria were as follows: (a) diagnosis of PDAC with peritoneal examination based on the surgical (for tumors treated with surgery) or diagnostic staging laparoscopy findings (for tumors treated with radiotherapy/chemotherapy); (b) an interval of <2 weeks between contrast-enhanced CT and pathological examinations; and (c) absence of typical peritoneal metastasis signs on CT images (diffuse omental nodule or omental cake, large amount of ascites, and marked irregular thickening with high peritoneal enhancement)[19]. The exclusion criteria were as follows: (a) history of pancreatic surgery, (b) history of pancreatic malignancy, (c) poor CT image quality that undermined peritoneal lesion assessment, (d) history of recurrent chronic pancreatitis, and (e) synchronous other malignancies. A total of 239 patients were included at Center 1 and was randomly divided into a training cohort (167 patients) and an internal test cohort (72 patients) in a 7:3 ratio. Center 2 included 63 patients as an external test cohort. Figure 1 presents the flowchart of the patient recruitment process.

CT protocol

Dual-phase enhanced CT examinations were performed using Aquilion One Vision (Canon Medical Systems), Aquilion Prime (Canon Medical Systems), IQon Spectral (Philips Healthcare), and Brilliance iCT (Philips Healthcare). The CT scanning protocol is detailed in Supplementary A1 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43) and Supplementary Table S1 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43).

Collection of clinical and radiological data

The clinical and radiological parameters, including sex, age, location, serum carcinoembryonic antigen (CEA) and CA19-9 levels, CT-based N stage, and CT-based T stage, were retrieved from the clinical database. CT-based T/N stage was defined according to the American Joint Committee on Cancer Version 8[20].

Development and validation of the clinical-radiological, DLR-T, HCR-P, and combined models

Volume of interest/region of interest segmentation and feature extraction and selection

A radiologist (Radiologist 1, with 7 years of experience in pancreatic diagnosis) performed three-dimensional volumes of interest

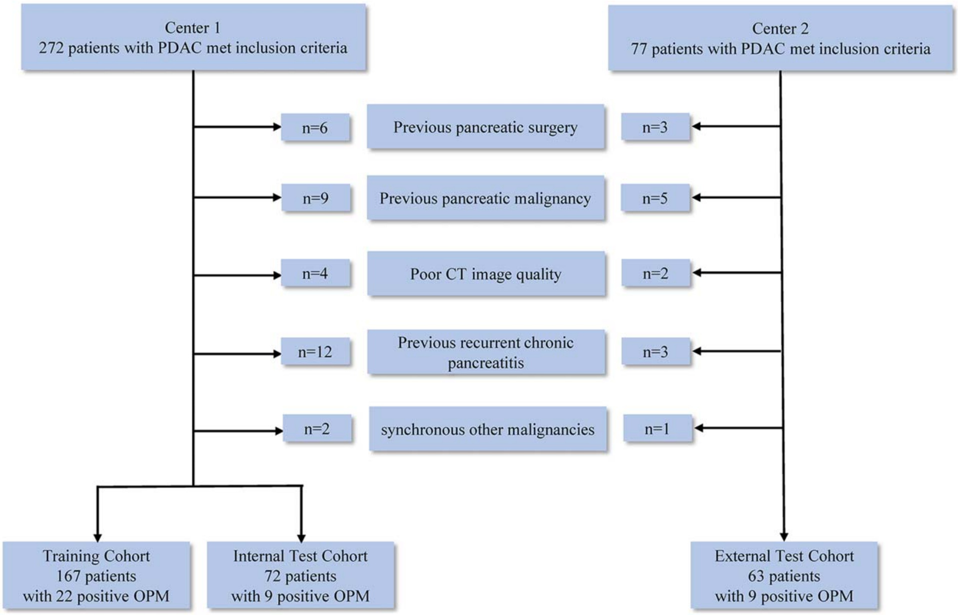

Figure 1. Flowchart of patient inclusion. CT, computer tomography; OPM, occult peritoneal metastases; PDAC, pancreatic ductal adenocarcinoma. A total of 239 patients were included at Center 1 and was randomly divided into a training cohort (167 patients) and an internal test cohort (72 patients) in a 7:3 ratio. Center 2 included 63 patients as an external test cohort.

(VOIs) segmentation of the tumor lesions and two-dimensional regions of interest (ROIs) segmentation of the peritoneal regions on pretreatment CT images in the three cohorts using MITK software (version 2018.04; https://www.mitk.org/). After 3 months, 30 patients were randomly selected from the training cohort, and their VOIs were segmented by Radiologist 1 and another radiologist (Radiologist 2, with 13 years of experience in pancreatic diagnosis) in the same manner. The radiologists were blinded to the clinical and pathological findings. The radiologists manually segmented all the slices of the tumor on the CT images and mapped the tumor

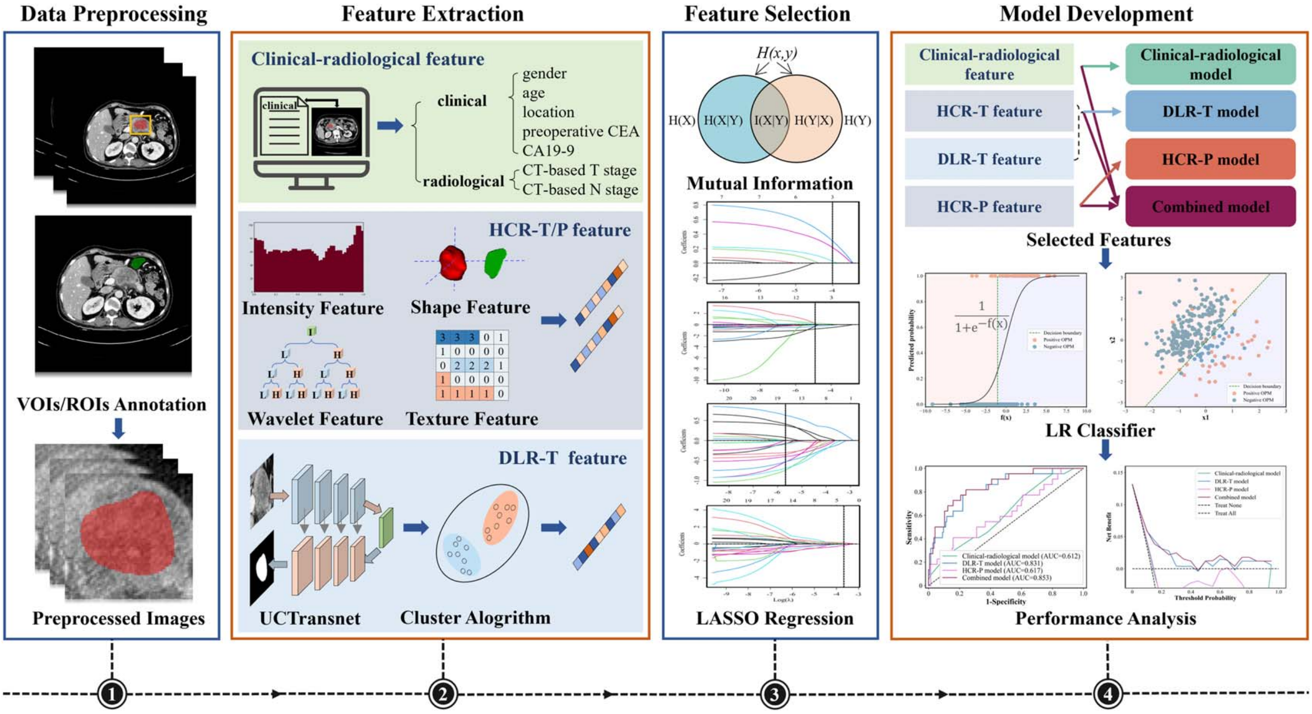

Figure 2. Workflow of model development. CEA, carcinoembryonic antigen; CA19-9, carbohydrate antigen 19-9; CT, computed tomography; DLR, deep learning-based radiomics; HCR, handcrafted radiomics; LASSO, least absolute shrinkage and selection operator; LR, logistic regression; ROIs, regions of interest; VOIs, volumes of interest.

Table 1

Characteristics of patients in different cohorts.

Characteristics	Training cohort (*n*=167)	Test cohorts			*P*
		Internal test cohort (*n*=72)	External test cohort (*n*=63)	Total test cohort (*n*=135)	
Sex, No. (%)					0.804[a]
Male	105 (62.9)	41 (56.9)	42 (66.7)	83 (61.5)	
Female	62 (37.1)	31 (43.1)	21 (33.3)	52 (38.5)	
Age, median (IQR), years	60.0 (53.0–67.0)	62.5 (53.0–70.0)	60.0 (50.0–65.0)	60.0 (50.0–68.0)	0.687[b]
Locations, No. (%)					0.795[a]
Head/uncinate	138 (82.6)	59 (81.9)	51 (81.0)	110 (81.5)	
Body/tail	29 (17.4)	13 (18.1)	12 (19.0)	25 (18.5)	
CEA, No. (%)					0.170[a]
Normal	128 (76.6)	51 (70.8)	43 (68.3)	94 (69.6)	
Elevated	39 (23.4)	21 (29.2)	20 (31.7)	41 (30.4)	
CA 19-9, No. (%)					0.921[a]
Normal	40 (24.0)	19 (26.4)	14 (22.2)	33 (24.4)	
Elevated	127 (76.0)	53 (73.6)	49 (77.8)	102 (75.6)	
CT-based T stage, No. (%)					0.119[a]
T1/T2	98 (58.7)	52 (72.2)	39 (61.9)	91 (67.4)	
T3/T4	69 (41.3)	20 (27.8)	24 (38.1)	44 (32.6)	
CT-based N stage, No. (%)					
N0	89 (53.3)	46 (63.9)	28 (44.4)	74 (54.8)	0.792[a]
others	78 (46.7)	26 (36.1)	35 (55.6)	61 (45.2)	

P-value indicated comparisons of the difference between training cohort and total test cohort.

[a] By chi-square test.

[b] By Mann–Whitney *U* test.

CA19-9, carbohydrate antigen 19-9; CEA, carcinoembryonic antigen; CT, computed tomography; No., number.

contours to generate a three-dimensional VOI. For the peritoneum, the radiologist selected the slice with the largest primary tumor area and delineated the ROI of the nearest peritoneal area (area > 2 cm^2) on this slice to extract peritoneal HCR features[10]. The inter-class and intra-class correlation coefficients (ICCs) were calculated to ensure inter-observer and intra-observer reproducibility of the HCR feature extraction (see Supplementary A2, Supplemental Digital Content 1, http://links.lww.com/JS9/C43). For the peritoneum, we also assess the feature reproducibility on the agreement of the selected slice and adjacent slices using ICCs

Table 2

Prediction performance of the clinical-radiological, DLR and combined model.

	AUC (95% CI)	ACC	SEN	SPEC	TP	TN	FP	FN	*P*
Training cohort (n = 167[a])									
Clinical-radiological model	0.612 (0.534-0.686)	0.437	0.773	0.386	17	56	89	5	0.076
DLR-T model	0.831 (0.765-0.884)	0.766	0.818	0.759	18	110	35	4	< 0.001
HCR-P model	0.617 (0.539-0.691)	0.784	0.409	0.841	9	122	23	13	0.084
Combined model	0.853 (0.790-0.903)	0.766	0.818	0.759	18	110	35	4	< 0.001
Total test cohort (n = 135[a])									
Clinical-radiological model	0.638 (0.551-0.719)	0.600	0.667	0.590	12	69	48	6	0.078
DLR-T model	0.817 (0.741-0.878)	0.733	0.722	0.735	13	86	31	5	< 0.001
HCR-P model	0.664 (0.577-0.743)	0.637	0.556	0.650	10	76	41	8	0.041
Combined model	0.842 (0.769-0.899)	0.770	0.722	0.778	13	91	26	5	< 0.001
Internal test cohort (n = 72[a])									
Clinical-radiological model	0.643 (0.521-0.752)	0.694	0.667	0.698	6	44	19	3	0.231
DLR-T model	0.811 (0.702-0.894)	0.722	0.778	0.714	7	45	18	2	0.002
HCR-P model	0.679 (0.559-0.784)	0.528	0.667	0.508	6	32	31	3	0.113
Combined model	0.845 (0.740-0.919)	0.722	0.778	0.714	7	45	18	2	< 0.001
External test cohort (n = 63[a])									
Clinical-radiological model	0.630 (0.499-0.748)	0.492	0.667	0.463	6	25	29	3	0.278
DLR-T model	0.840 (0.725-0.920)	0.746	0.667	0.759	6	41	13	3	< 0.001
HCR-P model	0.665 (0.534-0.779)	0.762	0.444	0.815	4	44	10	5	0.171
Combined model	0.852 (0.740-0.929)	0.825	0.667	0.852	6	46	8	3	< 0.001

Note: ACC, SEN, SPEC, TP, TN, FP, FN of the models in training and total test cohort were calculated with the corresponding cut-off values, which maximize the Youden index in the training cohort.

[a] The numbers in the parentheses represent the number of PDAC patients.

P value is the significance level of comparison of AUC with that of a random case (AUC = 0.500).

Abbreviations: AUC, area under receiver operating characteristic curve; CI, confidence interval; ACC, accuracy; DLR, deep learning-based radiomics; FN, false negative; FP, false positive; SEN, sensitivity; SPEC, specificity; TN, true negative; TP, true positive.

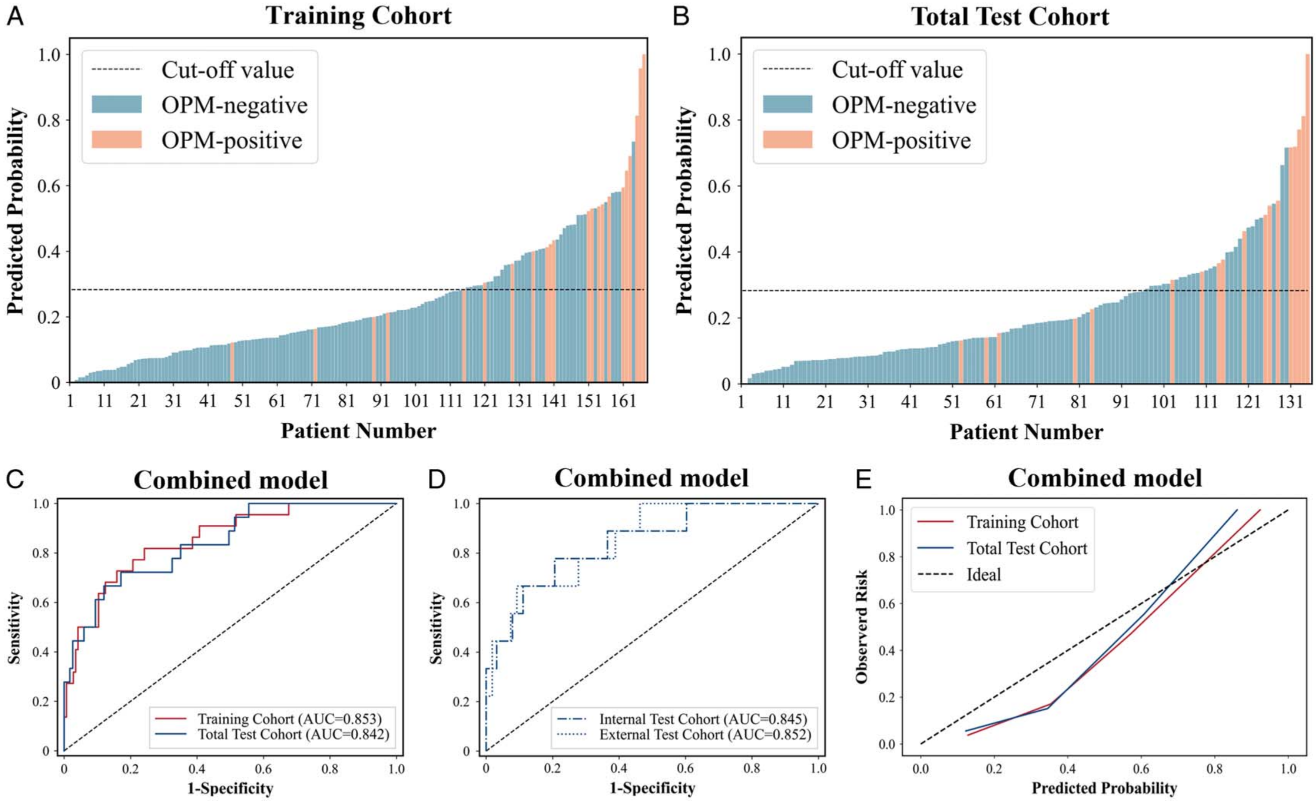

Figure 3. Predicted probability of combined model for each PDAC patient with a cut-off value of 0.283 (indicated by gray dotted line) for (A) the training cohort and (B) total test cohort. The diagnostic performances of combined model in the training cohort and total test cohort (C), as well as internal and external test cohorts (D). The calibration curves for the combined model in the training cohort and total test cohort (E). AUC, area under the receiver operator characteristic curve; OPM, occult peritoneal metastases; PDAC, pancreatic ductal adenocarcinoma.

(see Supplementary A2, Supplemental Digital Content 1, http://links.lww.com/JS9/C43). For each VOI of the tumor lesions, 1130 HCR features (HCR-T) were automatically extracted using the PyRadiomics toolkit (version 3.0.1; https://pyradiomics.readthedocs.io/en/v3.0.1/) (see Supplementary A3, Supplemental Digital Content 1, http://links.lww.com/JS9/C43). The DLR features (DLR-T) were extracted from a DL semantic segmentation model using UCTransNet[21]. A specific preprocessing procedure (Supplementary A4, Supplemental Digital Content 1, http://links.lww.com/JS9/C43) was used to ensure the proper training and testing of the segmentation model. This procedure was performed to align the images with the input requirements of the segmentation network, thereby optimizing the input data and enhancing the overall performance of the model. Subsequently, 50 patients were randomly selected from the training cohort to construct the DL segmentation model, and 512 deep semantic features were extracted using the DLR method[14]. Supplementary A4 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43) and Supplementary Figure S1 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43) illustrate this process. For each ROI of the peritoneal region, 474 HCR features (HCR-P) were automatically extracted using the PyRadiomics toolkit (see Supplementary A2, Supplemental Digital Content 1, http://links.lww.com/JS9/C43).

Continuous clinical-radiological variables and the HCR and DLR features of all patients were standardized using the mean and standard deviation of each corresponding feature in the training cohort. Feature selection was performed in the training cohort. HCR features with ICCs >0.90 were used to guarantee reproducibility. The clinical-radiological variables were selected using least absolute shrinkage and selection operator (LASSO) logistic regression analysis, whereas the retained HCR and DLR features were selected using mutual information and the LASSO algorithm, which ensured that the most relevant features related to the OPM status were retained and improved the accuracy and efficiency of the analysis. Supplementary A5 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43) and Supplementary Figure S2 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43) present the detailed feature selection process.

Development and validation of models

A logistic regression (LR) classifier was used to distinguish between positive and negative OPM based on the selected features. Four models were constructed to determine the performance differences between the following models: a clinical-radiological model constructed using clinical and radiological information, a DLR model of tumor (DLR-T model) constructed using HCR-T and DLR-T features, a HCR model of peritoneum (HCR-P model), and a combined model constructed using all selected features. Figure 2 presents the flowchart of the modeling method. Different features were integrated via direct feature splicing in the DLR-T and combined models to make full use of the feature information.

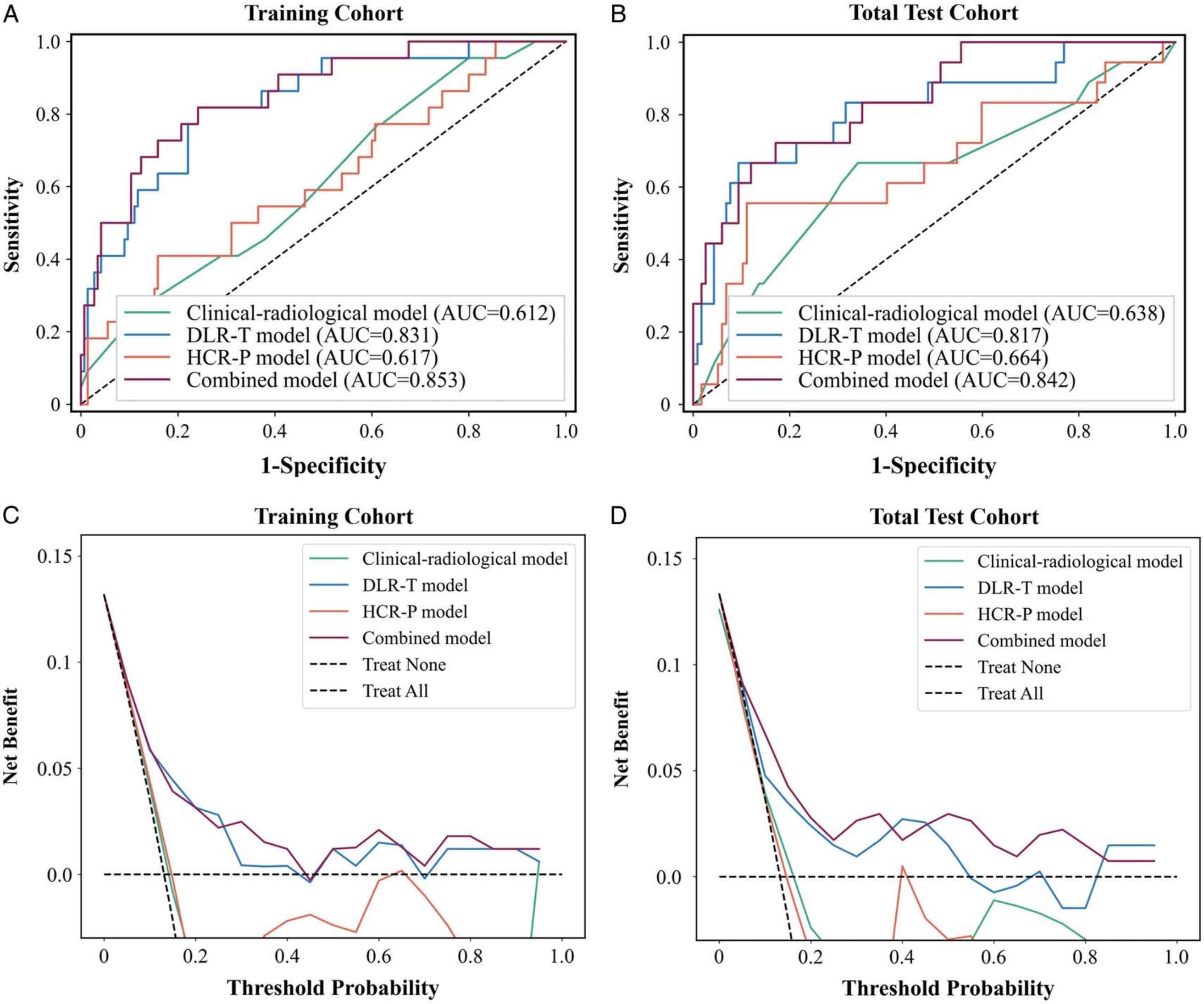

Figure 4. The ROC curves of clinical-radiological model, DLR-T model, HCR-P model, and combined model in the training cohort (A) and the total test cohort (B) are shown. Decision curves of clinical-radiological model, DLR-T model, HCR-P model, and combined model in the training cohort (C) and total test cohort (D) are shown. AUC, area under the receiver operator characteristic curve; DLR, deep learning-based radiomics; ROC, receiver operating characteristic.

The models were developed in the training cohort and validated in internal and external test cohorts. The predictive performance of the models was assessed using a receiver operating characteristic (ROC) curve and expressed as the area under the curve (AUC) with the corresponding 95% CI values. The accuracy (ACC); sensitivity (SEN); specificity (SPEC); and the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) cases were calculated according to the cut-off value derived from maximizing the Youden index from ROC curve analysis. The internal and external test cohorts were combined and analyzed in the subsequent comparison. Stratified analyses were conducted using different CT scanners to further assess the performance stability of the combined model. The sample size evaluation details in the Supplementary A6 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43). The feature visualization and analysis details in the Supplementary A7 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43).

Statistical analysis

Differences in the characteristics were assessed using the independent *t*-test or Mann–Whitney *U* test for continuous variables and Fisher's exact test or χ^2 test for categorical variables. The AUCs were compared using the DeLong test. Calibration curves were plotted to assess the calibration of different models, accompanied by the Hosmer–Lemeshow test. Decision curve analysis was performed to determine the clinical value of the combined model and quantify the net benefits at the threshold probabilities by comparing the clinical-radiological and DLR models in different cohorts.

Statistical analyses were performed using Python (version 3.9; https://www.python.org/), R (version 4.1.2; https://www.r-project.org/), MedCalc (version 15.8; https://www.medcalc.org/) and SPSS (version 26.0; https://www.ibm.com/). Statistical significance was defined as a two-sided P-value < 0.05.

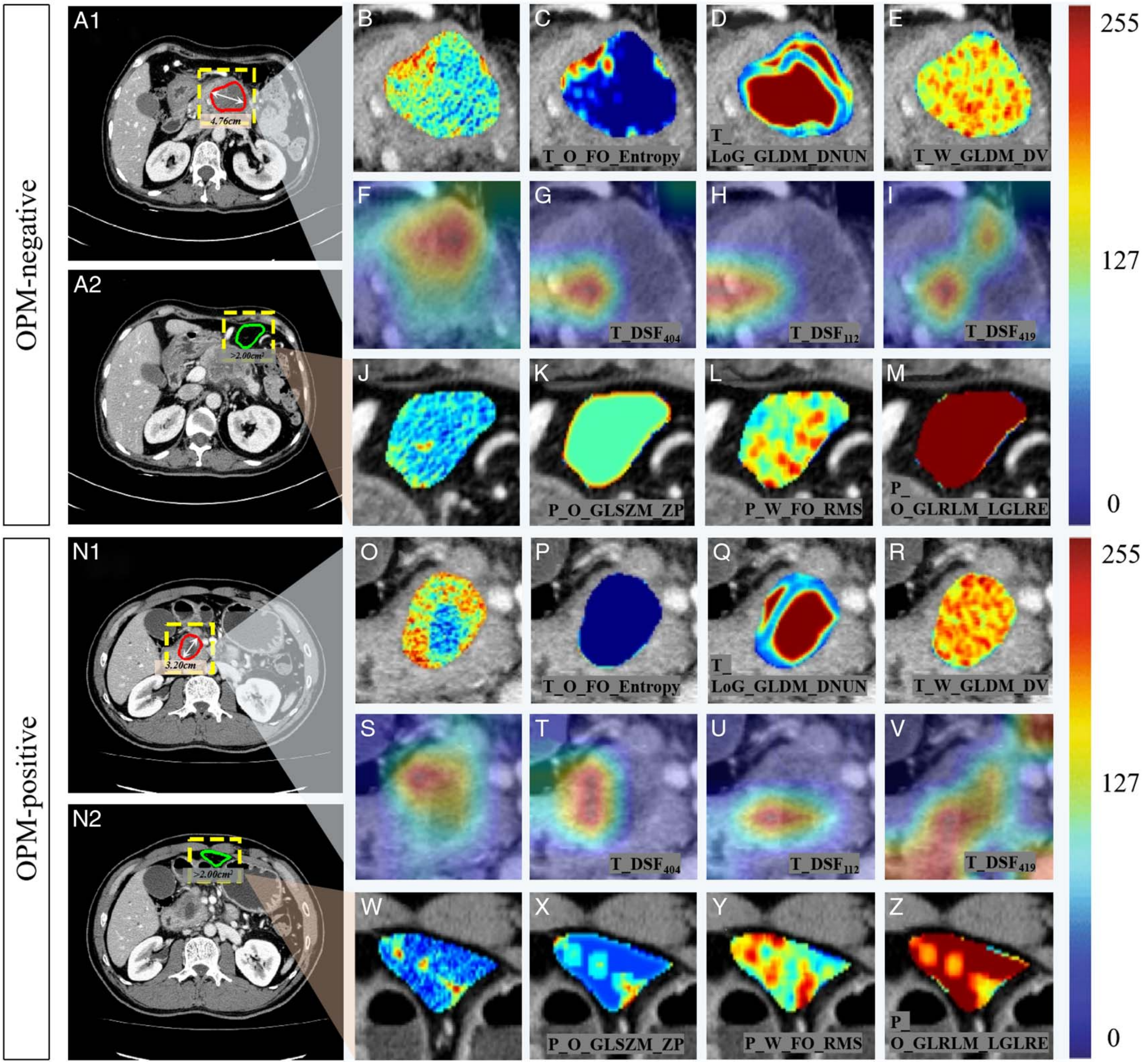

Figure 5. Visualization of the HCR and DLR features of a 53-year-old female patient with PDAC who was confirmed OPM-negative (A_1 , A_2) and a 47-year-old male patient with PDAC who was confirmed OPM-positive (N_1 , N_2). (B–E; O–R) show the tumors overlaid with pseudo color of pixel intensity and the top three HCR-T features. The red and yellow regions indicate higher pixel feature values, whereas the blue and green regions represent lower pixel feature values. The intensity pattern (B) within the tumor demonstrates heightened complexity near the edge of the tumor, as indicated by T_O_FO-entropy (C). A diminished uniformity is noticeable within the tumors, as evidenced by T_LoG_GLDM-DNUN (D, Q). Furthermore, textural changes were more pronounced in OPM-positive tumor than in OPM-negative tumor, according to T_W_GLDM-DV (E, R). (F–I; S–V) show the tumors overlaid with Grad-CAM heatmaps of all deep semantic features and the top three DLR features. The red and yellow regions represent the areas with higher activation, whereas the blue and green regions represent the areas with lower activation. All deep semantic features and the top three DLR-T features focused on the tumor region (F, I, S, U), or both tumor region and peritumoral area (G, H, T, V), respectively. In addition, (J–M; W–Z) show the peritoneal regions on the selected slices overlaid with pseudo color of pixel intensity and all three HCR-P features. Meanwhile, P_O_GLSZM_ZP (K, X) and P_O_GLRLM_LGLRE (M, Z) obviously reflect lower uniformity and higher heterogeneity in the peritoneum of OPM-negative patient. The OPM status of cases in the examples was incorrectly predicted by the clinical-radiological model but correctly predicted by the combined model. DSF, deep semantic feature; DLR, deep learning-based radiomics; Grad-CAM, gradient-weighted class activation mapping; HCR, handcrafted radiomics; OPM, occult peritoneal metastases; PDAC, pancreatic ductal adenocarcinoma.

Results

Patient characteristics and clinical-radiological model

The characteristics of the patients with PDAC in the training and test cohorts are shown in Table 1. No significant differences were found between the training and total test cohorts in terms of sex, age, location, pretreatment CEA, CA19-9, CT-based T stage, and CT-based N stage (all $P > 0.05$). Among the seven clinical and radiological characteristics, CA19-9 level, CT-based T stage, and CT-based N stage were retained using the LASSO algorithm.

The LR classifier was trained to achieve the best performance based on the selected characteristics to build the clinical-radiological model. As shown in Table 2, the model predicted OPM with AUCs of 0.612 (95% CI: 0.534–0.686, $P = 0.076$) and 0.638 (95% CI: 0.551–0.719, $P = 0.078$) in the training and total test cohorts, respectively. The AUCs of internal and external test cohorts were 0.643 (95% CI: 0.521–0.752, $P = 0.231$) and 0.630 (95% CI: 0.499–0.748, $P = 0.278$), respectively.

Development and validation of the DLR-T, HCR-P, and combined models

Among the HCR features, 577 HCR-T features of the tumor and 190 HCR-P features of the peritoneum with inter-observer and intra-observer ICCs > 0.90 were retained as candidates for further selection. Nine HCR-T features, 14 DLR-T features, and three HCR-P features were selected based on the mutual information and LASSO algorithms to build the DLR-T, HCR-P, and combined models using the LR classifier. Further details regarding the selected features in the LASSO algorithm are presented in Supplementary Table S2 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43). Supplementary A8 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43) presents the predicted-probability formula for the combined model.

The DLR-T model achieved AUCs of 0.831 (95% CI: 0.765–0.884; $P < 0.001$) and 0.817 (95% CI: 0.741–0.878; $P < 0.001$) in the training and total test cohorts, respectively, for predicting OPM. The AUCs of the internal and external test cohorts were 0.811 (95% CI: 0.702–0.894; $P = 0.002$) and 0.840 (95% CI: 0.725–0.920; $P < 0.001$), respectively (Table 2). The AUCs of the HCR-P model were 0.617 (95% CI: 0.539–0.691, $P = 0.084$) in the training cohort, 0.664 (95% CI: 0.557–0.743, $P = 0.041$) in the total test cohort, 0.679 (95% CI: 0.559–0.784, $P = 0.113$) in the internal test cohort, and 0.665 (95% CI: 0.534–0.779, $P = 0.171$) in the external test cohort. The combined model showed excellent predictive performance, with AUCs of 0.853 (95% CI: 0.790–0.903; $P < 0.001$) and 0.842 (95% CI: 0.769–0.899; $P < 0.001$) in the training and total test cohorts, respectively. The AUCs of the internal and external test cohorts were 0.845 (95% CI: 0.740–0.919; $P < 0.001$) and 0.852 (95% CI: 0.740–0.929; $P < 0.001$), respectively. Figure 3A and B present the predicted probabilities of the combined model for each patient in the training and total test cohorts.

DeLong's test revealed no significant difference in the ROC curves between the training and total test cohorts ($P = 0.826$; Fig. 3C) or between the internal and external test cohorts ($P = 0.943$; Fig. 3D) for the combined model, indicating the robustness of the model. As shown in Supplementary Table S3 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43) and Supplementary Figure S3 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43), the stratified analysis revealed that the performance of the combined model was not significantly affected by the type of CT scanner used (Delong's test, all $P > 0.05$). The Hosmer–Lemeshow test yielded χ^2 values of 16.947 ($P = 0.076$), and 12.810 ($P = 0.234$) for the training, and total test cohorts, respectively (Fig. 3E), demonstrating the good calibration power of the combined model.

Comparison of the clinical-radiological, DLR-T, HCR-P, and combined models

DeLong's test revealed that the discriminatory power of the combined model was significantly better than that of the clinical-radiological model in the training (AUC = 0.853 vs. 0.612, $P < 0.001$) and the total test (AUC = 0.842 vs. 0.638, $P < 0.05$) cohorts (Table 2 and Fig. 4A and B). The predictive performances of the combined model were better than the DLR-T model though there were no statistical significances in the training cohort (AUC = 0.853 vs. 0.831, $P = 0.081$), and the total test cohort (AUC = 0.842 vs. 0.817, $P = 0.495$) (Table 2 and Fig. 4A and B). The discriminatory power of the combined model was significantly better than that of the HCR-P model in the training (AUC = 0.853 vs. 0.617, $P < 0.001$) and the total test (AUC = 0.842 vs. 0.664, $P < 0.05$) cohorts (Table 2 and Fig. 4A and B). The decision curves (Fig. 4C and D) illustrated that the combined model surpassed the clinical-radiological, DLR-T, and HCR-P models in terms of the overall net benefit for the majority of threshold probabilities considered reasonable within both the training and total test cohorts.

Feature visualization and analysis

The top three HCR-T (T_O_FO_Entropy, T_LoG_GLDM_DNUN, and T_W_GLDM_DV) and DLR-T (T_DSF_{404}, T_DSF_{404}, and T_DSF_{418}) features with the highest coefficients and all three HCR-P features (P_O_GLSZM_ZP, P_W_FO_RMS, P_O_GLRLM_LGLRE) are highlighted in bold in Supplementary Table S2 (Supplemental Digital Content 1, http://links.lww.com/JS9/C43), and were selected to improve the comprehension of the operational mechanism of the combined model. Figure 5 presents the radiomics feature maps and gradient-weighted class activation mapping (Grad-CAM) heat maps of the chosen features overlaid with the CT images of a patient with PDAC who tested negative for OPM and a patient with PDAC who tested positive for OPM.

Discussion

This bicentric study developed and validated a model combining DLR features with clinical-radiological information of patients with PDAC to predict occult peritoneal metastasis with good robustness and generalization. The performance of the combined model was comparable with that of the DLR model and significantly outperformed the clinical-radiological model, providing an easy-to-use, and individualized tool for OPM detection that can help OPM-positive patients avoid inappropriate surgical procedures and identify the best candidates for laparoscopic exploration.

The 'seed and soil' theory proposed that the initiation of peritoneal metastasis depends on the synergistic effect of tumor cells (seed) and peritoneal microenvironment (soil)[22]. Therefore, the incorporation of radiomics feature derived from the tumor region and the peritoneal region resulted in incremental predictive performance of occult peritoneal metastasis, similar to previous study[10]. Peritoneal micro-metastases are prone to be overlooked during imaging, but radiomics features derived from the peritoneum will help to detect OPM at a very early stage[10]. Previous studies have shown that CA19-9 is a risk factor for peritoneal metastasis of PDAC[9,23], which in line with the findings of the present study. The HIF1α and c-Myc-mediated metabolic signaling pathways and gene-related reactive oxygen species production reflecting antioxidant stress responses were enhanced in patients with high CA19-9, indicating their association with high metastasis[24]. T stage, N stage, and peritoneal metastasis were consistently associated with the aggressiveness and prognosis of PDAC, and the CT-based T and N stages were included in our combined model for predicting OPM.

The proposed combined model, which merges the HCR and DLR features, can noninvasively and robustly characterize intratumoral heterogeneity from medical images at different levels[25,26], which helps improve the performance of the model. Moreover, combining high-dimensional features can provide further details regarding cancer, making the model more sensitive for disease diagnosis and prediction[13,27,28]. Furthermore, the clinical and radiological features can provide a more comprehensive description of tumor characteristics. The present study achieved a better diagnostic performance by combining DLR features with clinical-radiological features that complemented the image features with more information and made the model more robust by limiting the features extracted from the image[29]. This combination took real-world clinical practice into account and could provide meaningful guidance for clinicians who can benefit patients.

Although the accuracy of diagnostic imaging in determining the stage of PDAC has been improved, OPM that were not detected preoperatively can still be observed intraoperatively[30,31]. Surgery is not a first-line treatment option for these patients, and both postoperative complications and delayed chemotherapy may worsen prognosis[6]. In addition, diagnostic staging laparoscopy, the recommended examination for the diagnosis of peritoneal metastasis, is invasive and should be used selectively[8]. Furthermore, the application of neoadjuvant therapy eliminates preoperatively undetected micro-metastases, prevents recurrence, prolongs progression-free survival and overall survival, and currently plays an important role in the treatment of PDAC[32]. All in all, pretreatment noninvasive prediction of OPM can avoid unnecessary surgery and allow for a more individualized diagnostic and treatment plan, but it is still a challenge[6,33]. The present study attempted to construct a DLR-combined model to predict OPM in patients with PDAC. Our DLR model can noninvasively detect OPM in patients with PDAC prior to treatment, therefore, not only avoiding inappropriate surgery, but also screening more patients suitable for diagnostic staging laparoscopy and neoadjuvant therapy treatment. It is worth noting that the model showed satisfactory prediction accuracy in the training, internal, and external test cohorts. It has the potential to be broadly applied in clinical practice owing to its relatively large sample size and external validation. Furthermore, its predictive efficacy remained robust in stratified analyses of different CT scanners, highlighting its potential for generalization. These results indicated that the proposed model provides a complementary value for individualized diagnostic procedure and treatment decision-making.

In addition to radiomics, genomics, and pathomics are gradually attracting attention in surgery[34]. There is no doubt that combining radiomics, genomics, and pathomics can yield more information about tumors and can better help make personalized treatment decision and predict patient prognosis. However, genomic tests used to identify targetable variants can be expensive, adding to the financial burden on patients[35,36]. Pathological specimens are obtained by surgery or by biopsy, but not all patients are candidates for surgical treatment, and biopsy is limited by tumor heterogeneity and invasive procedures. Thus, we have selected the most readily available and relatively cost-effective radiological images to develop and validate the model for predicting the OPM in PDAC. There are still limitations and challenges in the clinical application of radiomics, including data curation and annotation, interpretability, and regulatory and reimbursement issues[37,38]. However, the successful application of radiomics in diagnostic purposes has prompted the application of radiomics-based cancer imaging analyses to meet more complex clinical needs[37].

Nevertheless, the present study had some limitations. First, the retrospective study design may have introduced selection bias. Second, there were relatively few OPM-positive samples, resulting in an unbalanced sample proportion, which may have influenced the predictive performance of the model. Furthermore, we only segmented the region of the peritoneum adjacent to the largest slice of the tumor to extract two dimension radiomics features. Although our results indicated that the consecutive three-layer peritoneal features show good agreement, further in-depth studies are needed in the future to explore the predictive efficacy of three-dimensional peritoneal regional features, as three-dimensional would be more representative than two-dimensional peritoneal segmentation.

Conclusion

In conclusion, the proposed noninvasive tool, which combines OPM-specific DLR features and clinical-radiological variables, for predicting OPM in patients with PDAC before treatment had good robustness and generalizability.

Ethical approval

This retrospective study was approved by the Institutional Review Board of the First Affiliated Hospital, Sun Yat-sen University, Guangdong, China on 13 January 2021 (approval number: 2021 [025]).

Consent

The requirement for obtaining informed consent from the patients was waived due to the retrospective design.

Sources of funding

This work was supported in part by the National Natural Science foundation of China (Grant Number 81801761 to Yanji Luo and Grant Number 82271958 to Shi-Ting Feng), Natural Science Foundation of Guangdong Province (2023A1515011097 to Shi-Ting Feng), and the Shenzhen-Hong Kong Institute of Brain Science-Shenzhen Fundamental Research Institutions of China (Grant Number 2023SHIBS0003 to Jian Zhang).

Author contribution

All authors contributed to the study conception and design. S.S., C.L., J.Z., B.H., Y.L., S.-T.F., L.W., M.C., J.Z., K.C., and Y.F.: conception and design of the study; S.S., Y.L., J.Z., S.-T.F., L.W., and M.C.: data collection; L.W. and M.C.: CT images acquisition; C.L., J.Z., K.C., Y.F., and B.H.: data analyses; S.S., C.L., J.Z., Y.L., B.H., and S.-T.F.: manuscript preparation; S.-T.F., Y.L., and B.H.: supervision and project administration.

Conflicts of interest disclosure

The author declares no conflict of interest.

Research registration unique identifying number (UIN)

This retrospective study only analyzed clinical and radiological data and did not intervene in the diagnosis and treatment of patients.

Guarantor

All the authors took responsibility of the final manuscript and approved it for publication.

Data sharing statement

All data generated for this study are included in the article or Supplementary Material and from the corresponding author upon reasonable request.

Provenance and peer review

Our paper was not invited.

Presentation

None.

References

[1] Siegel RL, Miller KD, Fuchs HE, *et al.* Cancer statistics, 2022. CA Cancer J Clin 2022;72:7–33.

[2] Mizrahi JD, Surana R, Valle JW, *et al.* Pancreatic cancer. Lancet 2020; 395:2008–20.

[3] Badic B, Morvan M, Quénéhervé L, *et al.* Real world data for pancreatic adenocarcinoma from a population-based study in France. Cancers (Basel) 2023;15:525.

[4] Takeda T, Sasaki T, Mie T, *et al.* Improved prognosis of pancreatic cancer patients with peritoneal metastasis. Pancreatology 2021;21:903–11.

[5] Hishinuma S, Ogata Y, Tomikawa M, *et al.* Patterns of recurrence after curative resection of pancreatic cancer, based on autopsy findings. J Gastrointest Surg 2006;10:511–8.

[6] Oba A, Inoue Y, Ono Y, *et al.* Radiologically occult metastatic pancreatic cancer: how can we avoid unbeneficial resection? Langenbecks Arch Surg 2020;405:35–41.

[7] Liu RC, Traverso LW. Diagnostic laparoscopy improves staging of pancreatic cancer deemed locally unresectable by computed tomography. Surg Endosc 2005;19:638–42.

[8] Li K, Cannon JGD, Jiang SY, *et al.* Diagnostic staging laparoscopy in gastric cancer treatment: a cost-effectiveness analysis. J Surg Oncol 2018; 117:1288–96.

[9] 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.

[10] Dong D, Tang L, Li ZY, *et al.* Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer. Ann Oncol 2019;30:431–8.

[11] Liu S, He J, Liu S, *et al.* Radiomics analysis using contrast-enhanced CT for preoperative prediction of occult peritoneal metastasis in advanced gastric cancer. Eur Radiol 2020;30:239–46.

[12] Yuan Z, Xu T, Cai J, *et al.* Development and validation of an image-based deep learning algorithm for detection of synchronous peritoneal carcinomatosis in colorectal cancer. Ann Surg 2022;275:e645–51.

[13] Wei M, Zhang Y, Ding C, *et al.* Associating peritoneal metastasis with T2-Weighted MRI images in epithelial ovarian cancer using deep learning and radiomics: a multicenter study. J Magn Reson Imaging 2023;59: 122–31.

[14] Huang B, Tian J, Zhang H, *et al.* Deep semantic segmentation feature-based radiomics for the classification tasks in medical image analysis. IEEE J Biomed Health Inform 2021;25:2655–64.

[15] Wang J, Wu D, Sun M, *et al.* Deep segmentation feature-based radiomics improves recurrence prediction of hepatocellular carcinoma. BME Front 2022;2022:9793716.

[16] Li Y, Fan Y, Xu D, *et al.* Deep learning radiomic analysis of DCE-MRI combined with clinical characteristics predicts pathological complete response to neoadjuvant chemotherapy in breast cancer. Front Oncol 2022;12:1041142.

[17] Guo Y, Xie X, Tang W, *et al.* Noninvasive identification of HER2-low-positive status by MRI-based deep learning radiomics predicts the disease-free survival of patients with breast cancer. Eur Radiol 2023;34: 899–913.

[18] Li M, Fan Y, You H, *et al.* Dual-Energy CT deep learning radiomics to predict macrotrabecular-massive hepatocellular carcinoma. Radiology 2023;308:e230255.

[19] Kim SJ, Kim HH, Kim YH, *et al.* Peritoneal metastasis: detection with 16- or 64-detector row CT in patients undergoing surgery for gastric cancer. Radiology 2009;253:407–15.

[20] Saka B, Balci S, Basturk O, *et al.* Pancreatic ductal adenocarcinoma is spread to the peripancreatic soft tissue in the majority of resected cases, rendering the AJCC T-stage protocol (7th Edition) inapplicable and insignificant: a size-based staging system (pT1: ≤ 2, pT2: >2-≤ 4, pT3: >4 cm) is more valid and clinically relevant. Ann Surg Oncol 2016;23:2010–8.

[21] Wang H, Cao P, Wang J, *et al.* Uctransnet: rethinking the skip connections in u-net from a channel-wise perspective with transformer. In Proceedings of the AAAI conference on artificial intelligence 2022;36: 2441–9.

[22] Fidler IJ, Poste G. The "seed and soil" hypothesis revisited. Lancet Oncol 2008;9:808.

[23] Karabicak I, Satoi S, Yanagimoto H, *et al.* Risk factors for latent distant organ metastasis detected by staging laparoscopy in patients with radiologically defined locally advanced pancreatic ductal adenocarcinoma. J Hepatobiliary Pancreat Sci 2016;23:750–5.

[24] Ji S, Zhang B, Liu J, *et al.* ALDOA functions as an oncogene in the highly metastatic pancreatic cancer. Cancer Lett 2016;374:127–35.

[25] Zhu Y, Man C, Gong L, *et al.* A deep learning radiomics model for preoperative grading in meningioma. Eur J Radiol 2019;116:128–34.

[26] Wang J, Zeng J, Li H, *et al.* A deep learning radiomics analysis for survival prediction in esophageal cancer. J Healthc Eng 2022;2022:4034404.

[27] An C, Li D, Li S, *et al.* Deep learning radiomics of dual-energy computed tomography for predicting lymph node metastases of pancreatic ductal adenocarcinoma. Eur J Nucl Med Mol Imaging 2022;49:1187–99.

[28] Liang HY, Yang SF, Zou HM, *et al.* Deep learning radiomics nomogram to predict lung metastasis in soft-tissue sarcoma: a multi-center study. Front Oncol 2022;12:897676.

[29] Xie Y, Zhang J, Xia Y, *et al.* Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT. Inform Fusion 2018;42:102–10.

[30] Scialpi M, Reginelli A, D'Andrea A, *et al.* Pancreatic tumors imaging: an update. Int J Surg 2016;28(suppl 1):S142–55.

[31] Cheng H, Luo G, Jin K, *et al.* Predictive values of preoperative markers for resectable pancreatic body and tail cancer determined by MDCT to detect occult metastases. World J Surg 2021;45:2185–90.

[32] Zhao B, Xia C, Xia T, *et al.* Development of a radiomics-based model to predict occult liver metastases of pancreatic ductal adenocarcinoma: a multicenter study. Int J Surg 2024;110:740–9.

[33] Hata T, Mizuma M, Iseki M, *et al.* Circulating tumor DNA as a predictive marker for occult metastases in pancreatic cancer patients with radiographically non-metastatic disease. J Hepato-Bil-Pan Sci 2021;28:648–58.

[34] Gumbs AA, Croner R, Abu-Hilal M, *et al.* Surgomics and the Artificial intelligence, Radiomics, Genomics, Oncopathomics and Surgomics (AiRGOS) Project. Art Int Surg 2023;3:180–5.

[35] National Cancer Institute. Tumor DNA sequencing in cancer treatment. Accessed 15 January 2019. https://www.cancer.gov/about-cancer/treatment/types/precision-medicine/tumor-dna-sequencing

[36] Yabroff K, Zhao J, de Moor J, *et al.* What factors are associated with oncologist discussions of the costs of genomic testing and related treatment? J Clin Oncol 2018;36(30 _suppl):95.

[37] Bera K, Braman N, Gupta A, *et al.* Predicting cancer outcomes with radiomics and artificial intelligence in radiology. Nat Rev Clin Oncol 2022;19:132–46.

[38] Taher H, Grasso V, Tawfik S, *et al.* The challenges of deep learning in artificial intelligence and autonomous actions in surgery: a literature review. Art Int Surg 2022;2:144–58.