


RESEARCH ARTICLE

Open Access



Preoperative prediction of early recurrence in resectable pancreatic cancer integrating clinical, radiologic, and CT radiomics features

Jeong Hyun Lee^{1†}, Jaeseung Shin^{1†}, Ji Hye Min^{1*} , Woo Kyoung Jeong¹, Honsoul Kim¹, Seo-Youn Choi^{1,2}, Jisun Lee³, Sungjun Hong⁴ and Kyunga Kim^{4,5}

Abstract

Objectives To use clinical, radiographic, and CT radiomics features to develop and validate a preoperative prediction model for the early recurrence of pancreatic cancer.

Methods We retrospectively analyzed 190 patients (150 and 40 in the development and test cohort from different centers) with pancreatic cancer who underwent pancreatectomy between January 2018 and June 2021. Radiomics, clinical-radiologic (CR), and clinical-radiologic-radiomics (CRR) models were developed for the prediction of recurrence within 12 months after surgery. Performance was evaluated using the area under the curve (AUC), Brier score, sensitivity, and specificity.

Results Early recurrence occurred in 36.7% and 42.5% of the development and test cohorts, respectively ($P=0.62$). The features for the CR model included carbohydrate antigen 19-9 > 500 U/mL (odds ratio [OR], 3.60; $P=0.01$), abutment to the portal and/or superior mesenteric vein (OR, 2.54; $P=0.054$), and adjacent organ invasion (OR, 2.91; $P=0.03$). The CRR model demonstrated significantly higher AUCs than the radiomics model in the internal (0.77 vs. 0.73; $P=0.048$) and external (0.83 vs. 0.69; $P=0.038$) validations. Although we found no significant difference between AUCs of the CR and CRR models (0.83 vs. 0.76; $P=0.17$), CRR models showed more balanced sensitivity and specificity (0.65 and 0.87) than CR model (0.41 and 0.91) in the test cohort.

Conclusions The CRR model outperformed the radiomics and CR models in predicting the early recurrence of pancreatic cancer, providing valuable information for risk stratification and treatment guidance.

Keywords Pancreatic cancer, Prognosis, Radiomics, Tomography, X-ray computed, Machine learning

[†]Jeong Hyun Lee and Jaeseung Shin contributed equally as co-first author to this work.

*Correspondence:
Ji Hye Min
minjh1123@gmail.com

¹Department of Radiology and Center for Imaging Science, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro Gangnam-gu, Seoul 06351, Republic of Korea

²Department of Radiology, Soonchunhyang University Bucheon Hospital, Soonchunhyang University College of Medicine, Bucheon, Republic of Korea

³Department of Radiology, College of Medicine, Chungbuk National University, Chungbuk National University Hospital, Cheongju, Republic of Korea

⁴Department of Digital Health, Samsung Advanced Institute of Health Sciences and Technology (SAIHST), Sungkyunkwan University, Seoul, Republic of Korea

⁵Biomedical Statistics Center, Research Institute for Future Medicine, Samsung Medical Center, Seoul, Republic of Korea



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Pancreatic cancer is a highly aggressive malignancy, which is challenging to treat, and has a 5-year survival rate of only 12% [1]. The poor prognosis of pancreatic cancer is primarily attributed to the high incidence of early postoperative recurrence, which up to 80% of the patients experience within 12 months after surgery [2]. Patients who experience early recurrence after surgical resection are unlikely to benefit from upfront surgery. Therefore, identifying patients at a high risk of early tumor recurrence before surgery may aid in selecting optimal treatment plan. Although several studies have developed preoperative predictive models based on clinical and radiological data, such as tumor size, lymphadenopathy, tumor differentiation, serum carbohydrate antigen 19-9 (CA19-9), and vascular abutment [3–5], reliable biomarkers indicative of the early recurrence of resectable pancreatic cancer are still lacking.

Radiomics is a subfield of radiology that involves the extraction and analysis of a large number of quantitative features from medical images [6]. As radiomics models have been applied to a variety of organ systems, these models are being developed in attempt to diagnose and determine the prognosis of pancreatic cancer [7–11]. These models, however, suffer from limitations, such as insufficient validation, small sample size, and the inability to capture information outside the tumor [7–11].

In the present multicenter study, we hypothesized that the combination of clinical, radiologic, and CT radiomics features would yield a more accurate prediction of early recurrence than clinical and/or radiologic features. We aimed to develop and validate a model, which integrated clinical, radiologic, and CT radiomics features, for the preoperative prediction of the early recurrence of pancreatic cancer.

Methods

The present multicenter study was approved by the institutional review boards of the three tertiary referral centers (Samsung Medical Center IRB No: 2022-11-121, Soon Chun Hyang University Hospital Bucheon IRB No: 2023-02-024, Chungbuk National University IRB No.: 2023-02-020-001), and the requirement for informed consent was waived due to the retrospective study design. Our study adhered to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis guidelines [12].

Patients

We searched the electronic databases of three hospitals for consecutive patients with pancreatic cancer who underwent pancreatectomy between January 2018 and June 2021. The inclusion criteria were as follows: (a) curative surgery for resectable pancreatic cancer

as determined by a multidisciplinary team discussion among surgeons, oncologists, and radiologists in line with the National Comprehensive Cancer Network (NCCN) guidelines for pancreatic cancer version 2.2021 [13], (b) availability of a preoperative CT scan within 3 months of surgery, and (c) a follow-up period ≥ 12 months. The exclusion criteria were as follows: (a) borderline resectable pancreatic cancer according to the NCCN guidelines for pancreatic cancer version 2.2021, (b) suboptimal CT image quality, (c) history of prior pancreatic surgery, and (d) missing clinical data. The three centers included the following: (center 1) Samsung Medical Center, Seoul, South Korea formed the development cohort and (center 2) Soonchunhyang University College of Medicine, Bucheon Hospital, Bucheon, South Korea and (center 3) Chungbuk National University Hospital, Cheongju, South Korea, formed the test cohort. As detailed in Fig. 1, a total of 150 and 40 patients were included in the development and test cohorts, respectively.

Clinical and pathological data

Clinical characteristics, such as age, sex, and preoperative serum CA19-9 levels obtained within one month before surgery, were collected from electronic medical records. Pathological results, including tumor size, tumor differentiation, T and N stages according to the American Joint Committee on Cancer tumor-node-metastasis 8th edition, and resection margin status, were also documented. Recurrence was defined as the presence of radiologic evidence of recurrent disease either at or adjacent to the surgical bed, including the remnant pancreas and locoregional nodes, or as evidence of recurrence detected outside these areas [14].

CT acquisition and evaluation

Multidetector contrast-enhanced CT examinations were performed by each institution's protocols. All CT examinations included the portal venous phase, by scanning the patient 70–80 s after the initiation of the contrast injection. The details of CT imaging techniques are summarized in Table S1.

Two board-certified abdominal radiologists (J.S. and J.H.L., with 10 and 9 years of experience, respectively), both blinded to the patients' clinicopathological information other than the diagnosis of pancreatic cancer, independently reviewed the CT images. After their initial review, they reevaluated the images together to reach a consensus. They evaluated the following radiologic features: tumor location (pancreatic head/neck vs. body/tail), tumor abutment to the portal vein (PV) and/or superior mesenteric vein (SMV), peripancreatic tumor infiltration, adjacent organ invasion, lymph node enlargement, findings of obstructive pancreatitis, upstream parenchymal atrophy, and dilatation of the

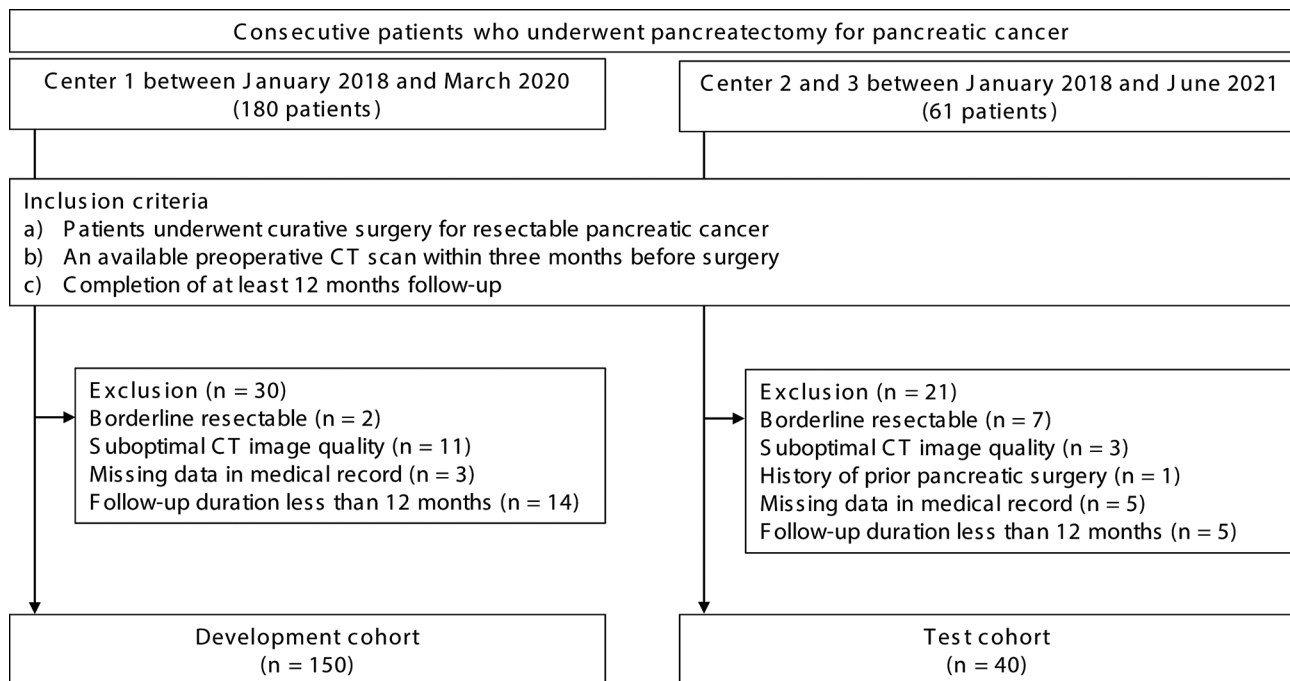


Fig. 1 Flowchart of the study population

main pancreatic duct [4, 15–18]. Detailed definitions of the imaging features are provided in Appendix E1.

Outcome measurement

Early recurrence was defined as recurrence of pancreatic cancer within the 12 months after surgery [2, 4]. After surgery, patients were followed up according to each institution's protocol, through October 2022. In general, the patients' follow-up included a clinical assessment, serum tumor marker evaluation, and CT or MRI every 3–6 months and at any time when clinically indicated.

Development and validation of predictive models for the early recurrence of pancreatic cancer

We developed the three models for the preoperative prediction of the early recurrence of pancreatic cancer in the development cohort (Fig. 2): (1) the radiomics model, based on CT radiomics features only; (2) clinical-radiologic (CR) model, based on clinical and radiologic features; and (3) the clinical-radiologic-radiomics (CRR) model, which combined the CR and radiomics models. The performance of each model was then validated externally using the test cohort.

Segmentation

A radiologist (J.H.M., 13 years of experience in abdominal radiology), blinded to the patients' clinical data except for pancreatic cancer diagnosis, manually segmented the pancreatic masses in both cohorts on the portal venous phase CT images by tracing along the tumor margins to

create a volume of interest (VOI), using commercial software (Aview, version 1.0.38.6; Coreline Soft). To verify the feature stability of the interobserver variance of segmentation, another radiologist (J.H.L.) performed tumor segmentation in 20 randomly selected cases from the development cohort.

Radiomics feature extraction and selection

Radiomics feature extraction was performed using the VOI of each tumor with PyRadiomics software (version 3.0.1; <https://www.radiomics.io/pyradiomics.html>) [19], in compliance with the Imaging Biomarker Standardization Initiative [20]. The original, Laplacian of the Gaussian (LoG)-filtered (sigma values of 2.0, 3.0, 4.0, and 5.0), and wavelet-filtered images were utilized to extract first-order, shape, and higher-order features. A total of 572 features were extracted, including 9 shape, 14 first-order, and 549 higher-order and filtered features.

Using data from the 20 cases, which were segmented independently by the aforementioned radiologists, the intraclass correlation coefficient (ICC) was calculated for all extracted features. Features with an interobserver ICC value >0.75 were selected for subsequent process. Radiomics features were then selected using the exhaustive variant of the minimum redundancy maximum relevance ensemble [21], the top five features were selected for the radiomics model, based on the selection frequency.

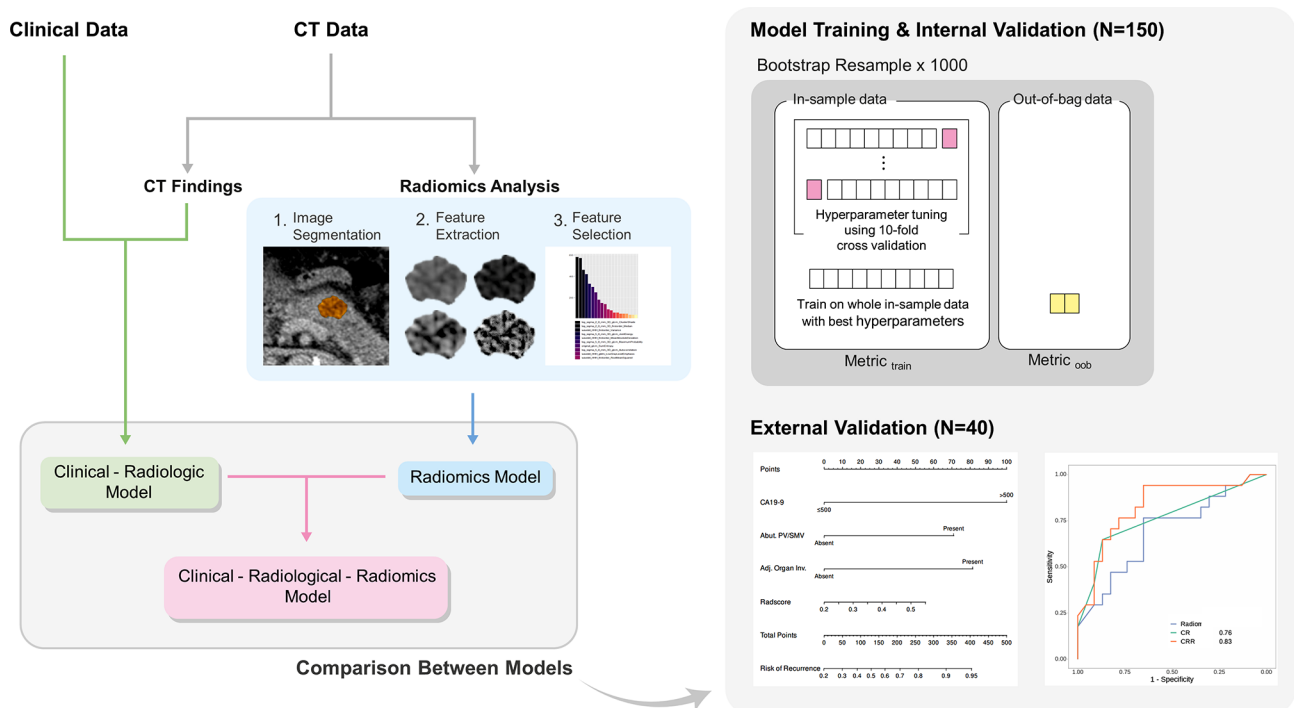


Fig. 2 Training and validation flowchart

Development and selection of radiomics model

The process used for developing the radiomics model is illustrated in Fig. 2. The following three candidate algorithms were considered: logistic regression with L2 regularization (ridge regression); random forest; and Light Gradient Boosting Machine (LightGBM) [22, 23]. To compare the performance of the models, bootstrap resampling was utilized to create 1,000 bootstrap samples. For random forest and LightGBM, a 10-fold cross-validation was performed for each bootstrap sample. The set of hyperparameters that produced the best mean area under the receiver operating characteristic curve (AUC) was determined by cross-validation and was used to train the bootstrap sample [24]. Among the three models developed, the model with the highest AUC was selected as the final predictive model. The probability output from the final radiomics model was designated as the “Radscore”.

Development of the CR model

To develop the CR model, univariable logistic regression was performed for each clinical and radiologic feature, and features with a P -value < 0.05 underwent multivariable logistic regression with backward selection using the Akaike information criterion. Features that survived backward selection were used to build the final CR model.

Development of the CRR model

The CRR model was constructed by combining the predictors from the CR model and Radscore from the radiomics model using ridge regression. The CRR model was presented as a nomogram to calculate the risk of early pancreatic cancer recurrence.

Statistical analysis

Comparisons between the development and test cohorts were done using chi-squared or the Fisher’s exact test for categorical variables, and student’s t -test or the Mann–Whitney U test for continuous variables. Interobserver agreement for the CT imaging features was evaluated using the kappa test as follows: poor, ≤ 0.20 ; fair, 0.21–0.40; moderate, 0.41–0.60; good, 0.61–0.80; and excellent, 0.81–0.99.

The performance of the models was evaluated using the AUC, Brier score, accuracy, sensitivity, and specificity. The Brier score is a measure of prediction accuracy, with lower scores indicating higher accuracy and 0 indicating a perfect prediction [25]. The thresholds for each model were set using the generalized threshold shifting technique, an algorithm-agnostic method for finding the optimal threshold for prediction by evaluating multiple subsets of the training data [26]. Internal validation was performed using 1,000-fold bootstrap samples with 0.632+ method for all metrics [27, 28] and with the percentile bootstrap method for 95% confidence interval (CI) [29].

In the development cohort, pairwise comparisons (radiomics vs. CRR and CR vs. CRR) of the AUCs were performed, and two-sided *P*-values were calculated by inverting the corresponding CIs [30, 31]. In the test cohort, pairwise AUC comparisons were performed using the Delong test. Statistical significance was set at

a *P*-value < 0.05. The software packages used are listed in Appendix E2.

Results

Patients

The clinical and CT radiologic features, as well as the postoperative data, are summarized in Table 1. The development and test cohorts included a total of 150 (mean age, 66.3 years ± standard deviation [SD], 9.3; 69 males) and 40 (mean age, 69.3 years ± SD, 8.9; 18 males) patients, respectively. During the follow-up period, early recurrence occurred in 36.7% (55/150) and 42.5% (17/40) of the patients in the development and test cohort, respectively (*P* = 0.62). There were no significant differences in the demographic features, CA19-9 levels, CT radiologic features, or pathological tumor stages between the two cohorts, although the development cohort had a smaller tumor size (mean 2.6 vs. 2.8 cm; *P* = 0.048) and a lower proportion of head/neck tumors (34.7 vs. 60.0%; *P* = 0.01) than the test cohort. The proportion of patients who received adjuvant therapy showed a borderline difference between the two cohorts (70.7 vs. 87.5%; *P* = 0.050).

Radiomics model

The ICCs for the radiomics features ranged 0.02–1.00, with a mean value of 0.81, and 418 features had an ICC > 0.75. The following top five features were subsequently selected: 3D_cluster_shade of LoG-filtered images (sigma = 2); 3D_first_order_median of LoG-filtered (sigma = 2); 3D_joint_energy of LoG-filtered (sigma = 5); mean_absolute_deviation of wavelet-filtered (HHH); and variance of wavelet-filtered (HHH). The mean AUCs (95% CI) for the random forest, logistic regression, and LightGBM were 0.73 (0.57–0.84), 0.64 (0.51–0.73), and 0.71 (0.54–0.84), respectively, and the mean Brier scores (95% CI) were 0.20 (0.16–0.25), 0.22 (0.20–0.27), and 0.32 (0.30–0.35), respectively. Based on these results, the random forest algorithm was selected for the final radiomics model.

CR model

Table 2 shows the frequency of each clinical and radiologic feature selected, and the results of the logistic regression analysis. Multivariable analysis revealed the following three features, which are predictive of the risk of early pancreatic cancer recurrence: CA19-9 > 500 U/mL (odds ratio [OR], 3.60; 95% CI, 1.39–9.34; *P* < 0.01), abutment to the PV and/or SMV (OR, 2.54; 95% CI, 0.98–6.56; *P* = 0.054), and adjacent organ invasion (OR, 2.91; 95% CI, 1.11–7.62; *P* = 0.03). Other than lymph node enlargement, the radiologic features showed moderate-to-excellent interreader agreement (Table S2).

Table 1 Characteristics of development and test cohorts

	Develop- ment cohort (<i>n</i> = 150)	Test cohort (<i>n</i> = 40)	<i>P</i> -value
Outcome			
Recurrence within 1 year	55 (36.7)	17 (42.5)	0.62
Pre-operative clinical features			
Age, years (mean ± SD)	66.3 ± 9.3	69.3 ± 8.9	0.06
Sex			0.68
Male	69 (46.0)	18 (45.0)	
Female	81 (54.0)	22 (55.0)	
CA19-9 (mean ± SD)	551.0 ± 1,291.6	176.6 ± 320.4	0.07
CT radiologic features			
Tumor location			< 0.01
Head/neck	52 (34.7)	24 (60.0)	
Body/tail	98 (65.3)	16 (40.0)	
Abutment to the PV and/ or SMV	27 (18.0)	6 (15.0)	0.83
Peripancreatic infiltration	72 (48.0)	19 (47.5)	> 0.99
Adjacent organ invasion	27 (18.0)	5 (12.5)	0.56
Enlarged lymph node	18 (12.0)	6 (15.0)	0.81
Obstructive pancreatitis	48 (32.0)	10 (25.0)	0.51
Upstream parenchymal atrophy	42 (28.0)	11 (27.5)	> 0.99
Main pancreatic duct dilatation	104 (69.3)	26 (65.0)	0.74
Post-operative features			
Tumor size, cm	2.6 ± 0.8	2.8 ± 0.7	0.048
Tumor differentiation			0.76
G1	7 (4.7)	2 (5.0)	
G2	117 (78.0)	33 (82.5)	
G3	26 (17.3)	5 (12.5)	
T stage			0.86
Tis	1 (0.7)	0 (0.0)	
T1	31 (20.7)	7 (17.5)	
T2	118 (78.7)	33 (82.5)	
N stage			0.93
N0	73 (48.7)	21 (52.5)	
N1	65 (43.3)	16 (40.0)	
N2	12 (8.0)	3 (7.5)	
Resection margin			0.39
R0	120 (80.0)	35 (87.5)	
R1	30 (20.0)	5 (12.5)	
Adjuvant chemotherapy/ CCRT	106 (70.7)	35 (87.5)	0.050

Note: Unless otherwise indicated, data represent the number of patients, and numbers in parentheses indicate percentages

CA 19-9, carbohydrate antigen 19-9; CCRT, concurrent chemoradiotherapy; PV, portal vein; SD, standard deviation; SMV, superior mesenteric vein

Table 2 Logistic regression analysis for predicting early recurrence of pancreatic cancer in the development cohort

Variables	Prevalence (%)	Univariable		Multivariable			
		Odds ratio (95% CI)	P-value	Clinical-radiologic model		Clinical-radiologic-radiomics model	
				Odds ratio (95% CI)	P-value	Odds ratio (95% CI)	P-value
CA19-9 > 500 U/mL	30 (20.0)	5.90 (2.46–14.17)	< 0.01	3.60 (1.39–9.34)	0.01	2.91 (1.44–6.41)	0.01
Abutment to the PV and/or SMV	27 (18.0)	3.13 (1.33–7.38)	< 0.01	2.54 (0.98–6.56)	0.054	2.14 (1.05–4.39)	0.04
Peripancreatic infiltration	72 (48.0)	1.51 (0.78–2.95)	0.22				
Adjacent organ invasion	27 (18.0)	3.80 (1.59–9.07)	< 0.01	2.91 (1.11–7.62)	0.03	2.39 (1.10–5.19)	0.03
Enlarged lymph node	18 (12.0)	1.11 (0.40–3.06)	0.84				
Obstructive pancreatitis	48 (32.0)	0.92 (0.45–1.89)	0.83				
Upstream parenchymal atrophy	42 (28.0)	0.52 (0.23–1.13)	0.10				
Main pancreatic duct dilatation	104 (69.3)	0.58 (0.28–1.18)	0.13				
Radscore	–	–	–	–	–	5.46 (2.94–9.28)	< 0.01

CA 19-9, carbohydrate antigen 19-9; CI, confidence interval; PV, portal vein; SMV, superior mesenteric vein

CRR model

Multivariable analysis of the combined model showed that a CA19-9 level > 500 U/mL (OR 2.91; 95% CI, 1.44–6.41; $P=0.01$), abutment to the PV and/or SMV (OR, 2.14; 95% CI, 1.05–4.39; $P=0.04$), and adjacent organ invasion (OR, 2.39; 95% CI, 1.10–5.19; $P=0.03$) were predictive of early pancreatic cancer recurrence. Furthermore, the Radscore (OR, 5.46; 95% CI, 2.94–9.28; $P<0.01$) was also found to be an independent predictor of early recurrence. The optimal threshold for the prediction of early recurrence was determined to be 0.41. The nomogram of the final CRR model is depicted in Fig. 3. CT images of representative cases with CR and CRR model discrepancy are shown in Fig. 4.

Internal validation

In the development cohort, the AUCs (95% CI) of the radiomics, CR, and CRR models were 0.73 (0.57–0.84), 0.70 (0.60–0.77), and 0.77 (0.63–0.86), respectively (Table 3). The AUC of the CRR model was higher than that of the radiomics-only model (0.77 vs. 0.73; $P=0.048$). Although the CCR model also had a higher AUC than the CR model, the difference was not statistically significant (0.70 vs. 0.77; $P=0.26$). The sensitivities (95% CI) of the radiomics, CR, and CRR models were 0.75 (0.38–0.96), 0.50 (0.26–0.68), and 0.77 (0.49–0.96), respectively, while the specificities (95% CI) were 0.60 (0.46–0.84), 0.81 (0.66–0.93), and 0.61 (0.47–0.84), respectively.

External validation

In the test cohort, the AUC (95% CI) of each model was as follows (Table 3): 0.69 (0.51–0.85) for radiomics, 0.76 (0.56–0.83) for CR, and 0.83 (0.65–0.94) for CRR. When the CR model was combined with the radiomics model, the resulting CRR model showed a higher AUC value than the radiomics model alone (0.69 vs. 0.83; $P=0.038$). The sensitivity (95% CI) of each model was as follows: 1.00 (1.00–1.00) for radiomics, 0.41 (0.12–0.53) for CR, and 0.65 (0.29–0.77) for CRR. The specificity (95% CI) of each model was as follows: 0.04 (0.00–0.13) for radiomics, 0.91 (0.78–1.00) for CR, and 0.87 (0.70–1.00) for CRR. Additionally, the increase in the Brier score for the radiomics model (0.04; 0.20 vs. 0.24) was greater than those of the CR (0.01; 0.20 vs. 0.21) and CRR (0.01; 0.19 vs. 0.20) models. The ROC curves for the three models are shown in Fig. 5.

Discussion

We developed radiomics-, CR-, and CRR-based models for the prediction of the early recurrence of pancreatic cancer within 12 months after surgical resection. The CRR model demonstrated better performance than the radiomics model alone, with a better AUC in both the development and test cohorts. External validation supported the effectiveness of the CRR model, indicating that integrating radiomics with clinical and radiologic factors enhanced the predictive accuracy for the early recurrence of pancreatic cancer.

Patients at high risk of early recurrence of pancreatic cancer might not benefit from an immediate surgical

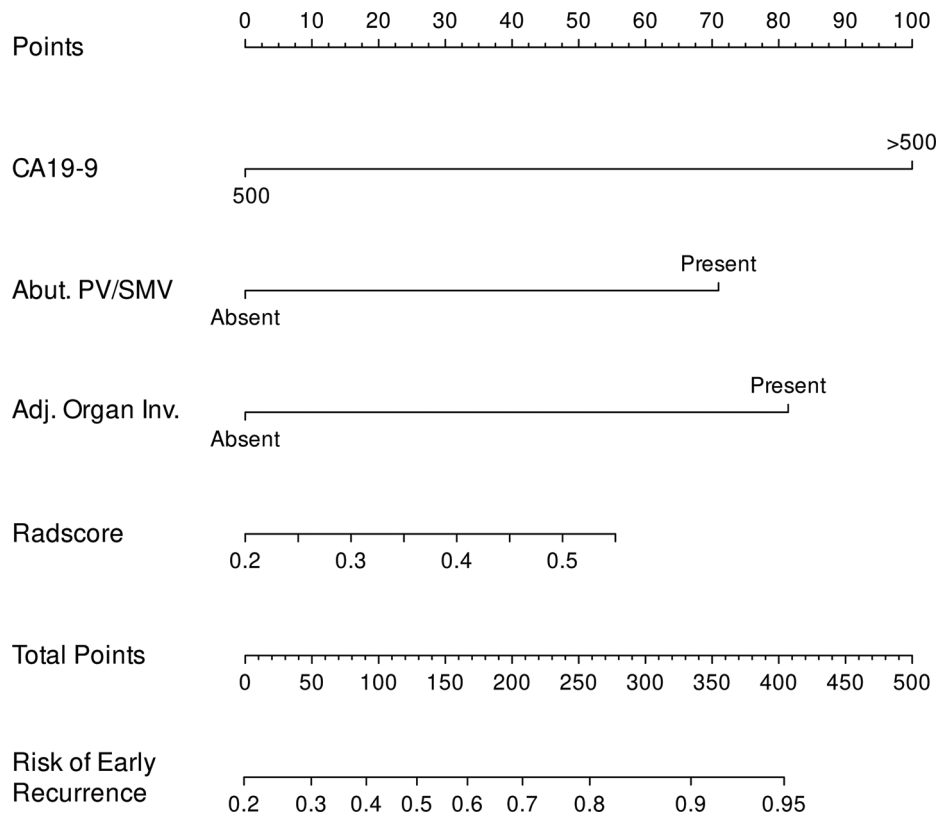


Fig. 3 Nomogram of the clinical-radiologic-radiomics model for predicting the early recurrence of pancreatic cancer after surgery. The nomogram based on the clinical-radiologic-radiomics model for the prediction of early recurrence after pancreatectomy for resectable pancreatic cancer. To use the nomogram, the value for each variable is located on the corresponding axis and a line is drawn upward to determine the corresponding points value. The sum of these point values for all four predictive features is located on the total points axis, and a line is drawn downward to the survival axis to determine the likelihood of early cancer recurrence for an individual patient

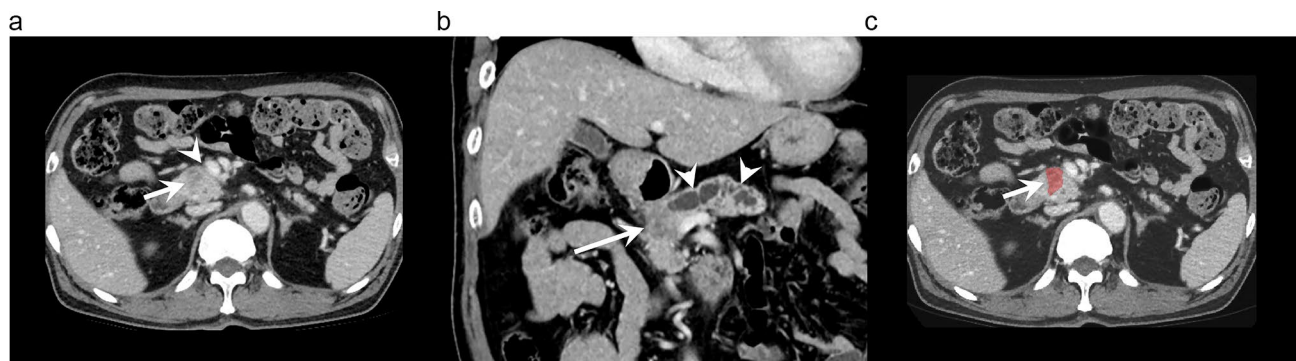


Fig. 4 CT images of a 64-year-old male with pancreatic cancer. The patient’s preoperative CA19-9 level was 16.7 U/ml. **(a)** Axial portal venous phase CT image shows a 3.2 cm-sized pancreatic cancer (arrow) at the pancreas head. There was abutment of the tumor to the superior mesenteric vein (arrow-head), but no evidence of adjacent organ invasion. **(b)** Coronal portal venous phase shows the pancreatic cancer (arrow) with upstream main pancreatic duct dilatation (arrowheads). **(c)** Portal venous phase with segmentation overlay. An R0 resection was performed, with a pathological stage of T2N1. The CR model predicted non-early recurrence, while the CRR model predicted early recurrence. The patient experienced recurrence within 12 months, in the form of liver metastasis

resection, which can result in significant morbidity [32]. The accurate prediction of early recurrence, therefore, can guide treatment decisions, such as determining which patients may benefit from neoadjuvant chemotherapy [33–36]. Unfortunately, there is no widely

accepted model for predicting the early recurrence of pancreatic cancer after curative resection. Previous models have limited clinical utility because they rely only on postoperative or pathological findings as predictive features. Predictions using those models, therefore, can only

Table 3 Model performance in predicting early recurrence of pancreatic cancer in the development and test cohorts

	AUC	P-value	Brier score	Accuracy	Sensitivity	Specificity
Development cohort						
Radiomics model	0.73 (0.57–0.84)	0.048*	0.20 (0.16–0.25)	0.64 (0.50–0.77)	0.75 (0.38–0.96)	0.60 (0.46–0.84)
CR model	0.70 (0.60–0.77)	0.26*	0.20 (0.18–0.24)	0.70 (0.61–0.76)	0.50 (0.26–0.68)	0.81 (0.66–0.93)
CRR model	0.77 (0.63–0.86)	–	0.19 (0.14–0.24)	0.66 (0.51–0.78)	0.77 (0.49–0.96)	0.61 (0.47–0.84)
Test cohort						
Radiomics model	0.69 (0.51–0.85)	0.038†	0.24 (0.22–0.25)	0.45 (0.43–0.50)	1.00 (1.00–1.00)	0.04 (0.00–0.13)
CR model	0.76 (0.56–0.83)	0.17†	0.21 (0.20–0.26)	0.70 (0.53–0.78)	0.41 (0.12–0.53)	0.91 (0.78–1.00)
CCR model	0.83 (0.65–0.94)	–	0.20 (0.18–0.24)	0.78 (0.60–0.85)	0.65 (0.29–0.77)	0.87 (0.70–1.00)

Data are presented as mean (95% CI)

*Pairwise comparison with CCR model using bootstrap

†Pairwise comparison with CCR model using Delong’s test

AUC, area under the receiver operating characteristic curve; CI, confidence interval; CR, clinical-radiologic; CRR, clinical-radiologic-radiomics

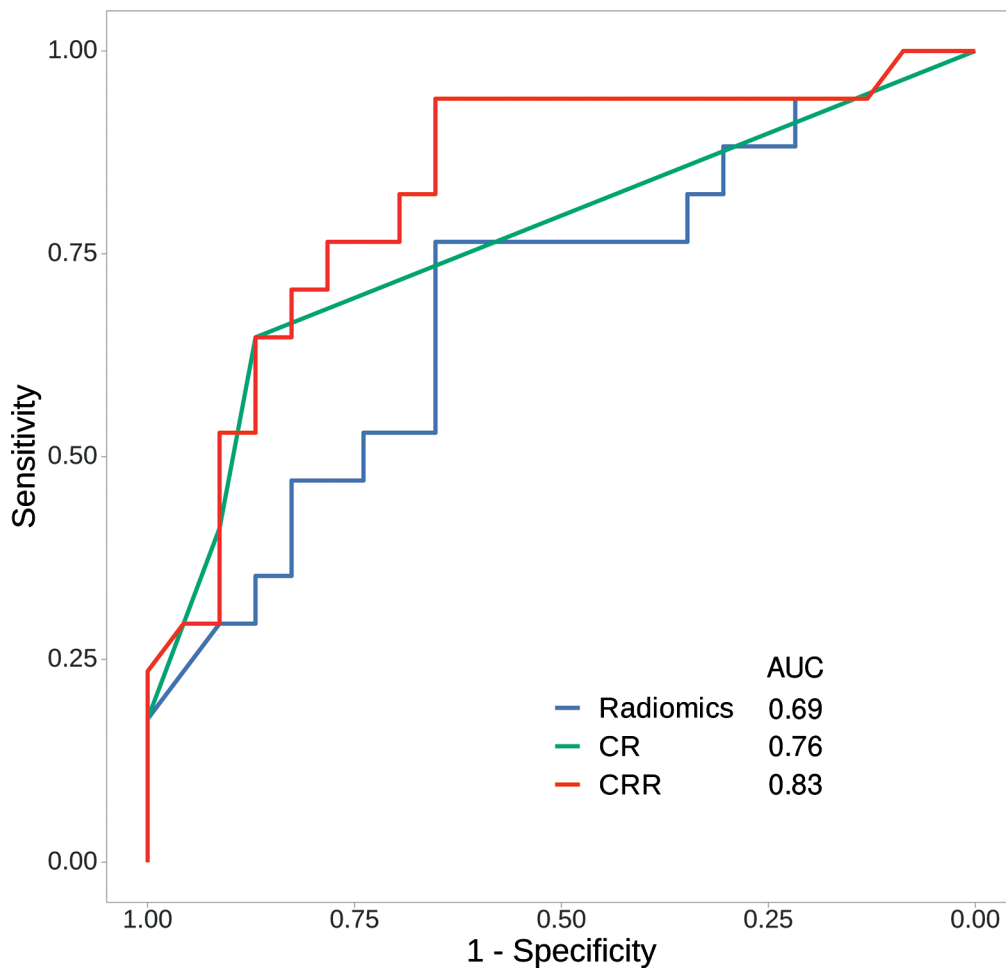


Fig. 5 Receiver operating characteristic (ROC) curves of the radiomics, CR, and CRR models in the test cohort. The AUC of the CRR model in the test cohort was higher than that of the radiomics model (0.69 vs. 0.83; $P=0.038$). AUC, area under the curve; CR, clinical-radiologic; CRR, clinical-radiologic-radiomics

be made after surgery and as such are not beneficial to the decision to preclude upfront surgical resection [3, 37, 38]. CA19-9 level has been the most consistent and widely accepted prognostic factor [17, 39], but its performance is limited, as its ability to predict early recurrence was reported to have an AUC of 0.68 [17]. Another issue with CA19-9 levels as a predictive factor is that they are heavily influenced by obstructive jaundice [39, 40]. Meanwhile, radiologic features, such as tumor size, hypodensity during the portal venous phase, invasion of adjacent organs, and contact with the PV and/or SMV are associated with pancreatic cancer prognosis [4, 15, 16]; however, their utility is limited due to the moderate interobserver agreement [41].

Previous studies have explored the potential applications of radiomics-based models in determining the prognosis of pancreatic cancer [42–46]; however, this approach is unable to capture information outside of the tumor, whereas the CR model may provide valuable peritumoral information, such as PV and/or SMV abutment or invasion of adjacent organs. Accordingly, the radiomics model in the present study did not demonstrate sufficient predictive ability for the early recurrence of pancreatic cancer. A previous study attempted to include perilesional information by increasing the segmentation boundary; however, this approach could not take into account broader contextual information, such as adjacent organ invasion and vascular abutment, which were also demonstrated to be significant in the present study [42]. One promising approach for enhancing radiomics-based models involves providing additional information, such as radiologic features and clinical data, that cannot be obtained from a segmented image alone [47]. The results of the present study support this hypothesis, as the combined CRR model performed better than the radiomics-only model.

The present study compared several machine-learning algorithms in developing a radiomics-based prediction model. Random forest, a model built on the bootstrap aggregation algorithm, seeks to avoid overfitting by an ensemble of multiple decision trees trained on different subsets of training data, thereby offering robust performance and noise resistance [48]. The random forest algorithm showed the best performance in the present study, which is consistent with previous studies on various organs [49–51]; however, compared to the CR and CRR models, the random forest-based radiomics model exhibited a significant decrease in accuracy and specificity with the largest increase in the Brier score on external validation. Potential causes for this decreased performance include variations in CT scanners and protocols, as well as differences in patient and tumor characteristics across different cohorts [52]. Accordingly, while modern machine learning algorithms deliver impressive results,

their performance may deteriorate significantly when applied to different clinical scenarios. Therefore, it is crucial to assess these models with various clinical settings and imaging protocols to ensure their efficacy in real world practice.

The present study had several limitations. First, due to the retrospective nature of the study, selection bias might have been introduced, despite efforts to minimize it. Second, our model did not account for postoperative factors that may affect recurrence, such as the resection margin status and the administration of adjuvant therapy. As we focused on the development of preoperative model for early recurrence to identify patients who may benefit from alternative treatments, those potential factors, which can be obtained after surgery, were not included in our model. Third, our cohorts exhibited a high average CA19-9 level, which was a significant variable in the final model. Therefore, additional research is needed to determine whether our model can be generalized to patients with normal CA 19-9 levels. Fourth, the development and testing cohorts were limited in size. Although bootstrapping was employed, a larger sample size would have produced more reliable results and potentially improved the radiomics model. Fifth, although interobserver variance in tumor segmentation was addressed using the ICC, more accurate automatic segmentation may decrease potential measurement errors. Unfortunately, automatic segmentation of the pancreas is not sufficiently reliable as of now.

Conclusion

In conclusion, we have developed and compared the performance of radiomics, CR, and CRR models for predicting early recurrence in operable pancreatic cancer. The CRR model outperformed the radiomics and CR models in predicting the early recurrence of pancreatic cancer, providing valuable information for risk stratification and treatment guidance prior to surgical resection of pancreatic cancer.

Abbreviations

AUC	area under the curve
CA19-9	carbohydrate antigen 19-9
CI	confidence interval
CR	clinical-radiologic
CRR	clinical-radiologic-radiomics
CT	computed tomography
GBM	gradient boosting machine
ICC	intraclass correlation coefficient
LoG	Laplacian of the Gaussian
OR	odds ratio
PV	portal vein
SD	standard deviation
SMV	superior mesenteric vein
VOI	volume of interest

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40644-024-00653-3>.

Supplementary Tables and Appendix

Acknowledgements

The authors thank Seol Eui Lee for providing excellent technical support with a medical illustration.

Author contributions

J.H.M conceived the presented idea. J.H.M, J.H.L, and J.S developed the methodology and performed the computations. J.H.L and J.S were the main contributors in manuscript writing. W.K.J, H.K, S.C, J.L verified the analytical methods and provided feedback. S.H and K.K helped with the statistical methodology.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the South Korean government (MSIT) (No. 2020R1F1A1071531 and No. 2022R1F1A1074362) and a grant of Korea Government Grant Program for Education and Research in Medical AI through the Korea Health Industry Development Institute (KHIDI), funded by the Korea government (MOE, MOHW).

Data availability

The data supporting the conclusions of this article is(are) available in the author's github repository, [https://github.com/sigjhl/pdac_prognosis]. Our patient data is not included in the repository, but available upon reasonable request.

Declarations

Ethics approval and consent to participate

The present multicenter study was approved by the institutional review boards of the three tertiary referral centers (Samsung Medical Center IRB No: 2022-11-121, Soon Chun Hyang University Hospital Bucheon IRB No: 2023-02-024, Chungbuk National University IRB No.: 2023-02-020-001), and the requirement for informed consent was waived due to the retrospective study design.

Consent for publication

All authors consented to the submission of this manuscript.

Competing interests

Not applicable.

Received: 26 July 2023 / Accepted: 29 December 2023

Published online: 08 January 2024

References

1. American Cancer Society. *Cancer Facts & Fig.* 2023. Atlanta: American Cancer Society; 2023.
2. Groot VP, Gemenetzis G, Blair AB et al. Defining and predicting early recurrence in 957 patients with resected pancreatic ductal adenocarcinoma. *Ann Surg* 2019.
3. Strijker M, Chen J, Mungroop T, et al. Systematic review of clinical prediction models for survival after surgery for resectable pancreatic cancer. *J Br Surg*. 2019;106:342–54.
4. Kim N, Han IY, Ryu Y, et al. Predictive nomogram for early recurrence after pancreatectomy in resectable pancreatic cancer: risk classification using preoperative clinicopathologic factors. *Cancers*. 2020;12:137.
5. Sugawara T, Ban D, Nishino J, et al. Prediction of early recurrence of pancreatic ductal adenocarcinoma after resection. *PLoS ONE*. 2021;16:e0249885.
6. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278:563–77.
7. Preuss K, Thach N, Liang X, et al. Using quantitative imaging for personalized medicine in pancreatic cancer: a review of radiomics and deep learning applications. *Cancers*. 2022;14:1654.
8. Gao Y, Cheng S, Zhu L, et al. A systematic review of prognosis predictive role of radiomics in pancreatic cancer: heterogeneity markers or statistical tricks? *Eur Radiol*. 2022;32:8443–52.
9. Mukherjee S, Patra A, Khasawneh H, et al. Radiomics-based machine learning models can detect pancreatic Cancer on Prediagnostic CTs at a substantial lead time prior to clinical diagnosis. *Gastroenterology*. 2022;S0016–5085(0022):00728.
10. Palumbo D, Mori M, Prato F, et al. Prediction of early distant recurrence in upfront resectable pancreatic adenocarcinoma: a multidisciplinary, machine learning-based approach. *Cancers*. 2021;13:4938.
11. Abunahel BM, Pontre B, Kumar H, Petrov MS. Pancreas image mining: a systematic review of radiomics. *Eur Radiol*. 2021;31:3447–67.
12. Moons KG, Altman DG, Reitsma JB, et al. Transparent reporting of a multi-variable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162:W1–W73.
13. Network NCC. Datum pristupa. NCCN Guidelines Version 2.2021 Pancreatic Adenocarcinoma. Dostupno na adresi: https://www.nccn.org/professionals/physician_gls/pdf/pancreatic.pdf. 2021; 23.
14. Jones RP, Psarelli EE, Jackson R, et al. Patterns of recurrence after resection of pancreatic ductal adenocarcinoma: a secondary analysis of the ESPAC-4 randomized adjuvant chemotherapy trial. *JAMA Surg*. 2019;154:1038–48. <https://doi.org/10.1001/jamasurg.2019.3337>.
15. Kim DW, Lee SS, Kim S-O, et al. Estimating recurrence after upfront surgery in patients with resectable pancreatic ductal adenocarcinoma by using pancreatic CT: development and validation of a risk score. *Radiology*. 2020;296:541–51.
16. Li D, Wang L, Cai W, Liang M, Ma X, Zhao X. Prognostic stratification in patients with pancreatic ductal adenocarcinoma after curative resection based on preoperative pancreatic contrast-enhanced CT findings. *Eur J Radiol*. 2022;151:110313.
17. Nishio K, Kimura K, Amano R, et al. Preoperative predictors for early recurrence of resectable pancreatic cancer. *World J Surg Oncol*. 2017;15:1–10.
18. Ferrone CR, Finkelstein DM, Thayer SP, Muzikansky A, Fernandez-del Castillo C, Warsaw AL. Perioperative CA19-9 levels can predict stage and survival in patients with resectable pancreatic adenocarcinoma. *J Clin Oncol: Off J Am Soc Clin Oncol*. 2006;24.
19. Van Griethuysen JJ, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res*. 2017;77:e104–7.
20. Zwanenburg A, Vallières M, Abdalah MA, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*. 2020;295:328–38.
21. De Jay N, Papillon-Cavanagh S, Olsen C, El-Hachem N, Bontempi G, Haibe-Kains B. mRMRe: an R package for parallelized mRMRe ensemble feature selection. *Bioinformatics*. 2013;29:2365–8.
22. Breiman L. Random forests. *Mach Learn*. 2001;45:5–32.
23. Ke G, Meng Q, Finley T et al. Lightgbm: a highly efficient gradient boosting decision tree. *Adv Neural Inf Process Syst*. 2017;30.
24. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: a next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, 2019:2623–2631.
25. Rufibach K. Use of Brier score to assess binary predictions. *J Clin Epidemiol*. 2010;63:938–9.
26. Esposito C, Landrum GA, Schneider N, Stiefel N, Riniker S. GHOST: adjusting the decision threshold to handle imbalanced data in machine learning. *J Chem Inf Model*. 2021;61:2623–40.
27. Efron B, Tibshirani R. Improvements on cross-validation: the 632+ bootstrap method. *J Am Stat Assoc*. 1997;92:548–60.
28. Iba K, Shinozaki T, Maruo K, Noma H. Re-evaluation of the comparative effectiveness of bootstrap-based optimism correction methods in the development of multivariable clinical prediction models. *BMC Med Res Methodol*. 2021;21:1–14.
29. Rousseelet GA, Pernet CR, Wilcox RR. The percentile bootstrap: a primer with step-by-step instructions in R. *advances in methods and practices in. Psychol Sci*. 2021;4:2515245920911881.
30. Liu RY, Singh K. Notions of limiting P values based on data depth and bootstrap. *J Am Stat Assoc*. 1997;92:266–77.
31. Wilcox RR. *Introduction to robust estimation and hypothesis testing*. Academic press; 2011.

32. Dusch N, Lietzmann A, Barthels F, Niedergethmann M, Rückert F, Wilhelm T. International study group of pancreatic surgery definitions for postpancreatectomy complications: applicability at a high-volume center. *Scand J Surg*. 2017;106:216–23.
33. Miyasaka Y, Ohtsuka T, Kimura R, et al. Neoadjuvant chemotherapy with gemcitabine plus nab-paclitaxel for borderline resectable pancreatic cancer potentially improves survival and facilitates surgery. *Ann Surg Oncol*. 2019;26:1528–34.
34. Ye M, Zhang Q, Chen Y, et al. Neoadjuvant chemotherapy for primary resectable pancreatic cancer: a systematic review and meta-analysis. *HPB*. 2020;22:821–32.
35. Versteijne E, Vogel JA, Besselink M, et al. Meta-analysis comparing upfront surgery with neoadjuvant treatment in patients with resectable or borderline resectable pancreatic cancer. *J Br Surg*. 2018;105:946–58.
36. Jang J-Y, Han Y, Lee H, et al. Oncological benefits of neoadjuvant chemoradiation with gemcitabine versus upfront surgery in patients with borderline resectable pancreatic cancer: a prospective, randomized, open-label, multi-center phase 2/3 trial. *Ann Surg*. 2018;268:215–22.
37. Brennan MF, Kattan MW, Klimstra D, Conlon K. Prognostic nomogram for patients undergoing resection for adenocarcinoma of the pancreas. *Ann Surg*. 2004;240:293.
38. Adamu M, Nitschke P, Petrov P, et al. Validation of prognostic risk scores for patients undergoing resection for pancreatic cancer. *Pancreatol*. 2018;18:585–91.
39. Turrini O, Schmidt C, Moreno J, et al. Very high serum CA 19-9 levels: a contraindication to pancreaticoduodenectomy? *J Gastrointest Surg*. 2009;13:1791–7.
40. Tsen A, Barbara M, Rosenkranz L. Dilemma of elevated CA 19-9 in biliary pathology. *Pancreatol*. 2018;18:862–7.
41. Joo I, Lee JM, Lee ES, et al. Preoperative CT classification of the resectability of pancreatic cancer: interobserver agreement. *Radiology*. 2019;293:343–9.
42. 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.
43. Cassinotto C, Chong J, Zogopoulos G, et al. Resectable pancreatic adenocarcinoma: role of CT quantitative imaging biomarkers for predicting pathology and patient outcomes. *Eur J Radiol*. 2017;90:152–8.
44. Chakraborty J, Langdon-Embry L, Cunanan KM, et al. Preliminary study of tumor heterogeneity in imaging predicts two year survival in pancreatic cancer patients. *PLoS ONE*. 2017;12:e0188022.
45. Attiye MA, Chakraborty J, Doussot A, et al. Survival prediction in pancreatic ductal adenocarcinoma by quantitative computed tomography image analysis. *Ann Surg Oncol*. 2018;25:1034–42.
46. He M, Xue H, Jin Z. Radiomics in pancreatic ductal adenocarcinoma: a state of art review. *J Pancreatol*. 2020;3:195–200.
47. Wen Q, Yang Z, Dai H, Feng A, Li Q. Radiomics study for predicting the expression of PD-L1 and tumor mutation burden in non-small cell lung cancer based on CT images and clinicopathological features. *Front Oncol*. 2021;11:620246.
48. Borstelmann SM. Machine learning principles for radiology investigators. *Acad Radiol*. 2020;27:13–25.
49. McKinley R, Häni L, Gralla J, et al. Fully automated stroke tissue estimation using random forest classifiers (FASTER). *J Cereb Blood Flow Metabolism*. 2017;37:2728–41.
50. Oakden-Rayner L, Carneiro G, Bessen T, Nascimento JC, Bradley AP, Palmer LJ. Precision radiology: predicting longevity using feature engineering and deep learning methods in a radiomics framework. *Sci Rep*. 2017;7:1648.
51. Raman SP, Chen Y, Schroeder JL, Huang P, Fishman EK. CT texture analysis of renal masses: pilot study using random forest classification for prediction of pathology. *Acad Radiol*. 2014;21:1587–96.
52. Healy GM, Salinas-Miranda E, Jain R et al. Pre-operative radiomics model for prognostication in resectable pancreatic adenocarcinoma with external validation. *Eur Radiol* 2021: 1–14.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.