



OPEN

Comprehensive analysis of clinical data and radiomic features from contrast enhanced CT for differentiating benign and malignant pancreatic intraductal papillary mucinous neoplasms

Fengxiang Lou^{1,3}, Mingyang Li^{2,3}, Tongjia Chu¹, Haoyu Duan¹, Huan Liu¹, Jian Zhang¹, Kehang Duan¹, Han Liu¹ & Feng Wei¹✉

The primary aim of this investigation was to leverage radiomics features derived from contrast-enhanced abdominal computed tomography (CT) scans to devise a predictive model to discern the benign and malignant nature of intraductal papillary mucinous neoplasms (IPMNs). Radiomic signatures were meticulously crafted to delineate benign from malignant IPMNs by extracting pertinent features from contrast-enhanced CT images within a designated training cohort ($n = 84$). Subsequent validation was conducted with data from an independent test cohort ($n = 37$). The discriminative ability of the model was quantitatively evaluated through receiver operating characteristic (ROC) curve analysis, with the integration of carefully selected clinical features to improve the comparative analysis. Arterial-phase images were utilized to construct a model comprising 8 features for distinguishing between benign and malignant cases. The model achieved an accuracy of 0.891 [95% confidence interval (95% CI), 0.816–0.996] in the cross-validation set and 0.553 (95% CI 0.360–0.745) in the test set. Conversely, employing 9 features from the venous-phase resulted in a model with a cross-validation accuracy of 0.862 (95% CI 0.777–0.946) and a test set accuracy of 0.801 (95% CI 0.653–0.950). Integrating the identified clinical features with imaging features yielded a model with a cross-validation accuracy of 0.934 (95% CI 0.879–0.990) and a test set accuracy of 0.904 (95% CI 0.808–0.999), thereby further improving its discriminatory ability. Our findings distinctly illustrate that venous-phase radiomics features eclipse arterial-phase radiomic features in terms of predictive accuracy regarding the nature of IPMNs. Furthermore, the synthesis and meticulous screening of clinical features with radiomic data significantly increased the diagnostic efficacy of our model, underscoring the pivotal importance of a comprehensive and integrated approach for accurate risk stratification in IPMN management.

Pancreatic ductal adenocarcinoma (PDAC) is recognized as a profoundly malignant neoplasm, and is the sixth leading cause of oncologic mortality in China^{1,2}. Statistical analysis of data from 2015 to 2019 revealed an annual increase in the incidence of pancreatic cancer of 0.6% to 1%. Projections for 2024 suggest that 611,720 individuals in the United States will succumb to cancer, translating to approximately 1680 deaths per day, with a significant proportion of these deaths attributed to pancreatic cancer³. Globally, there has been a surge in the incidence and mortality rates of pancreatic cancer, accompanied by a disheartening 5-year overall survival rate ranging

¹Department of Hepatobiliary and Pancreatic Surgery, General Surgery Center, The First Hospital of Jilin University, Changchun 130000, China. ²Department of Radiology, The First Hospital of Jilin University, Changchun 130000, China. ³These authors contributed equally: Fengxiang Lou and Mingyang Li. ✉email: wei_feng@jlu.edu.cn

from 9 to 13%⁴. The insidious nature of pancreatic cancer typically results in the late emergence of nonspecific symptoms such as abdominal pain, jaundice, vomiting, and hiccups, which are often apparent only advanced stages. Regrettably, only approximately 25% of these cases are amenable to surgical intervention⁵. Consequently, proactive monitoring of individuals at risk for pancreatic cancer is imperative.

In its precancerous stages, PDAC primarily arises from pancreatic intraepithelial neoplasia (PanIN) and mucinous cystic neoplasms (MCNs), with the latter including intraductal papillary mucinous neoplasms (IPMNs) and mucinous cystic tumours. IPMNs are of clinical importance because of their high prevalence and potential for malignant transformation, which significantly impacts patient outcomes. According to the latest World Health Organization (WHO) classification of gastrointestinal tumours, IPMNs are stratified into three categories on the basis of the degree of dysplasia: low-grade, high-grade (also termed in situ carcinoma), and IPMN-related invasive carcinoma⁶. Clinical evidence suggests that low-grade IPMNs with dysplasia are benign and associated with a lower risk of malignant progression, and a 5-year disease-specific survival rate of 97%, whereas high-grade and invasive IPMNs are associated with 5-year specific survival rates of 84% and 39%, respectively⁷. Notably, longitudinal study findings indicate that a mere 2.8% of patients with IPMNs develop malignant tumours over time⁸. Additionally, surgical treatment carries the inherent risk of postoperative complications, subsequently increasing mortality rates⁹. Therefore, a nuanced approach that optimizes diagnostic accuracy and surgical decision-making is required for effective management of IPMNs to minimize unnecessary surgeries while preventing malignancy progression. Thus, increasing diagnostic precision is critical for advancing therapeutic strategies and improving outcomes for patients with pancreatic IPMNs.¹⁰

Traditional imaging techniques have exhibited inherent limitations in effectively delineating the risks associated with IPMNs. In an effort to overcome these limitations, the concept of “radiomics” was coined by the American scholar Gillies¹¹. In radiomics, critical features are extracted from radiographic images to create high-dimensional datasets, thereby increasing the accuracy of identifying the risk factors associated with IPMNs. The radiomic workflow involves four essential steps: data acquisition, preprocessing, feature extraction, and feature selection¹². Within this framework, the processes of data handling and analysis are instrumental in refining model development, providing a foundation for predictive accuracy.

As radiomic technology has progressed, radiomics has burgeoned into a vital tool in multiple applications, including tumour diagnosis, automated structured reporting, and the evaluation of treatment efficacy. The widespread use of radiomics in clinical applications has proven particularly advantageous in the diagnosis of conditions such as pulmonary nodules, breast cancer, prostate cancer, and rectal cancer. The incorporation of radiomics has significantly augmented diagnostic precision, ultimately facilitating better patient outcomes.

The aims of this study were to evaluate the preoperative predictive accuracy of a model constructed using radiomic features extracted from pancreatic contrast-enhanced CT images in distinguishing between benign and malignant IPMNs; to explore the optimal imaging phase for contrast-enhanced CT that maximizes predictive performance; and to investigate whether the integration of both qualitative and quantitative clinical features can refine the accuracy of pathological stratification diagnoses. Ultimately, the goal is to deliver more precise treatment recommendations, thereby increasing both the immediate and long-term survival prospects of patients diagnosed with IPMNs.

Method

Study population and inclusion criteria

A retrospective analysis of patients with IPMNs who underwent surgical treatment at the First Hospital of Jilin University from January 2013 to January 2023 was conducted. Patients were categorized into benign and malignant groups on the basis of postoperative pathological results. The data collected included demographic, clinical, and radiological parameters. The exclusion criteria were not receiving neoadjuvant therapy prior to surgery and not undergoing contrast-enhanced CT scanning within our facility. A total of 121 patients were deemed suitable for inclusion in the study.

We confirm that all methods employed in this study strictly adhered to relevant guidelines and regulations. The research protocols had been previously reviewed and approved by the Ethics Committee of the First Hospital of Jilin University, ensuring compliance with established ethical standards. Furthermore, informed consent was obtained from all individual participants and/or their legal guardians prior to their involvement in the study.

Image acquisition

In this study, preoperative pancreatic contrast-enhanced CT examinations were conducted with three different models of spiral CT machines: PHILIPS Ict 64/256, SIEMENS Cardiac 64, and GE Revolution 64 (the choice of machine was likely based on the availability and performance characteristics of these machines to ensure the acquisition of the most accurate images). The specific details regarding image acquisition are as follows: Scanning range extended from the xiphoid process level to the umbilical level, covering the entire abdomen, particularly the region containing the pancreas. This comprehensive coverage ensured a complete dataset for diagnostic purposes. Scan parameter settings: Tube voltage (influencing image contrast), 120 kVp; tube current, adjusted automatically to accommodate variations in patient size; pitch (determining the distance the CT machine moves per complete rotation), 5; matrix (defining image resolution), 521 × 521; and reconstructed slice thickness, between 1 and 1.5 mm, with thinner slices providing higher image resolution. All images were processed to create 1-mm-thick, thin-slice CT images to increase image clarity. The images were archived in DICOM format, the standard for medical images in diagnosis and research. These detailed parameters and processing steps are crucial to ensuring the accuracy and repeatability of the study, especially in analysing patients with IPMNs who are undergoing hepato-pancreato-biliary surgical treatment. Accurate image data are fundamental for diagnosis and research in this context.

Collection of radiological features in CT images

To ensure the accuracy and reliability of the radiological data, each CT image underwent meticulous independent review by two radiologists, each with over a decade of experience in clinical image analysis. Discrepancies between their initial assessments prompted a structured discussion, complemented by an additional review from a senior physician, to achieve a consensus on the evaluations. This rigorous procedure was imperative for the precise identification, classification, and localization of IPMN lesions, as well as for evaluating critical indicators such as dilation of the main pancreatic duct (MPD). The incorporation of such detailed and expert-driven analysis guarantees the robustness and reliability of the radiological features collected for subsequent analyses.

Radiomic analysis of CT images

All patient images in DICOM format were imported into RIAS software and meticulously processed by two radiologists with two and ten years of clinical reading experience. These radiologists carefully delineated the complete and surrounding boundaries of the pancreas and the lesion, layer by layer, to generate three-dimensional volumes of interest (VOIs). This manual process ensured a precise depiction of the pancreas and lesions while avoiding major blood vessels and dilated bile ducts adjacent to the lesion. In cases of evaluative discrepancies, the radiologists engaged in discussions to reach a consensus. After completing the delineation, the image regions of interest (ROIs) were saved in an independent “nii” file format and categorized on the basis of their benign or malignant characteristics for further analysis.

RIAS software uses standardization technology to extract radiological features within the region of interest. These features typically encompass four categories: (1) shape-based characteristics; (2) first-order intensity attributes; (3) second-order texture properties; and (4) high-order texture features. Shape-based characteristics primarily include tumour volume, maximum diameter, surface area, tumour density, and sphericity. The first-order intensity feature is derived from the histogram, with evaluation of the ROI on the basis of the mean, median, maximum, minimum, kurtosis, entropy, and skewness values of the voxel intensity. Texture features are computed via various grey matrices, such as the grey co-occurrence matrix (GLCM), and grey dependence matrix (GLDM). Higher-order features are typically generated via wavelet or Laplacian Gaussian filtering. After extraction, a standardization process was applied to ensure consistent scaling across all the features. The LASSO algorithm was employed for dimensionality reduction, selecting the most significant parameters pertinent to pancreatic lesions.

The comprehensive dataset was divided into a cross-validation set and a test set at a 7:3 ratio. Selected radiomic features were utilized to construct predictive models via a random forest classifier (RFC), support vector machine (SVM), and logistic regression (LR) on the RIAS platform, ensuring efficient data classification and prediction. The SVM model, which demonstrated superior performance, was selected for validation. A fivefold cross-validation method involving receiver operating characteristic (ROC) curve plotting and evaluation of metrics such as the area under the curve (AUC), accuracy, sensitivity, specificity, and other relevant indicators was applied to the cross-validation set to assess model performance.

Separate radiomic models for the arterial and venous phases were developed and compared against various evaluation indicators to determine the optimal predictive model. The model exhibiting the best prediction performance was subsequently integrated with selected clinical features to increase diagnostic accuracy and efficacy.

Statistical analysis

In this study, statistical analysis was performed on data collected from 121 patients using IBM SPSS 26.0 software. Continuous variables, including patient attributes such as height, weight, and lesion size, were evaluated. Differences in categorical variables, such as sex, lesion type, pancreatic duct dilation, and the solid-cystic nature of the lesion, were also evaluated. For normally distributed metric data (e.g., patient age, lesion diameter, and volume), mean values and standard deviations are reported. Nonnormally distributed data, are reported as median values along with upper and lower quartiles. Independent sample *t* tests were used to compare two sets of data for differences, whereas Mann-Whitney *U* tests were used for nonnormally distributed data. ROC curve analysis and the DeLong test were employed to compare the diagnostic performance of radiomic models from different phases and models combining selected clinical features. The chi-square test and corrected chi-square test were used to validate the results, with a *p*-value less than 0.05 indicating significant differences between groups.

Results

Analysis of clinical characteristics

A total of 121 patients diagnosed with IPMNs were included in the study, ranging in age from 36 to 77 years, with a mean age of 60.2 years. The cohort consisted of 61 males (50.41%) and 60 females (49.59%). Among them, 61 patients were determined to have benign IPMNs, including 30 with mild dysplasia, 21 with mild to moderate dysplasia, and 10 with moderate dysplasia. The malignant IPMN group comprised 60 patients, including 10 with moderate to severe dysplasia, 9 with severe dysplasia, and 41 with IPMNs associated with invasive carcinoma. Both qualitative and quantitative data from the 121 patients were statistically analysed, and the results are detailed in Table 1.

Selection and extraction of radiomic features

Using RIAS software, we extracted 1427 dimensional features from the ROIs in the cross-validation set (Fig. 1). Following intragroup validation, 954 arterial-phase features and 843 venous-phase features were retained. Through dimensionality reduction, 8 and 9 dimensional arterial-phase features and venous-phase features were retained respectively. Figure 2 and Table 2 illustrate the visualized radiomic features after dimensionality

Clinical features		Benign	Malignant	P-value
Sex				0.87
	Male	32	29	
	Female	29	31	
Age		59	61	0.54
BMI (kg/m ²)		22.55 ± 2.84	22.20 ± 2.93	0.199
Type				< 0.05
	MD-IPMNs	12	39	
	BD-IPMNs	44	9	
	MX-IPMNs	5	12	
Position				0.18
	Head	28	37	
	Neck, body and tail	27	23	
	All	6		
Size				< 0.05
	< 3 cm	46	26	
	> 3 cm	15	44	
Cystic solid				< 0.05
	Cystic	52	28	
	Solid	4	18	
	Both	5	14	
Single or multiple				0.304
	Single	50	50	
	Multiple	11	10	
Lymph node involvement				0.904
	Not involved	36	38	
	Involved	25	22	
Pancreatic duct dilation				0.627
	No	23	14	
	Yes	38	46	
Neutrophil/lymphocyte ratio		2.25 ± 1.17	3.13 ± 2.43	0.96
Tumour markers	CA199 (U/ml)	9.48	41.33	< 0.05
	CA125 (U/ml)	8.92	13.13	< 0.05
	CEA (ng/ml)	2.23	2.85	0.649
	CA724 (U/ml)	6.28	6.53	0.207
	CA242 (U/ml)	8.33	30.5	0.915
Total bilirubin (μmol/l)		11.9	15.1	< 0.05
Alkaline phosphatase (U/ml)		75.5	99.4	< 0.05
γ-ggtt (U/l)		21.7	27.8	< 0.05
Diabetes				< 0.05
	No	53	43	
	Yes	8	17	
Chronic pancreatitis				0.71
	No	53	55	
	Yes	8	5	
Symptom				0.24
	No	24	19	
	Yes	37	41	

Table 1. Clinical data and statistical analysis results of 121 patients.

reduction in arterial-phases; Fig. 3 and Table 3 illustrate the visualized radiomic features after dimensionality reduction in venous-phases.

Establishment of an arterial phase radiomics feature model and evaluation of predictive performance

We divided the entire sample dataset into training and testing sets at a ratio of 7:3. Employing dimensionality reduction techniques, we extracted meaningful feature parameters from radiomic features and constructed an

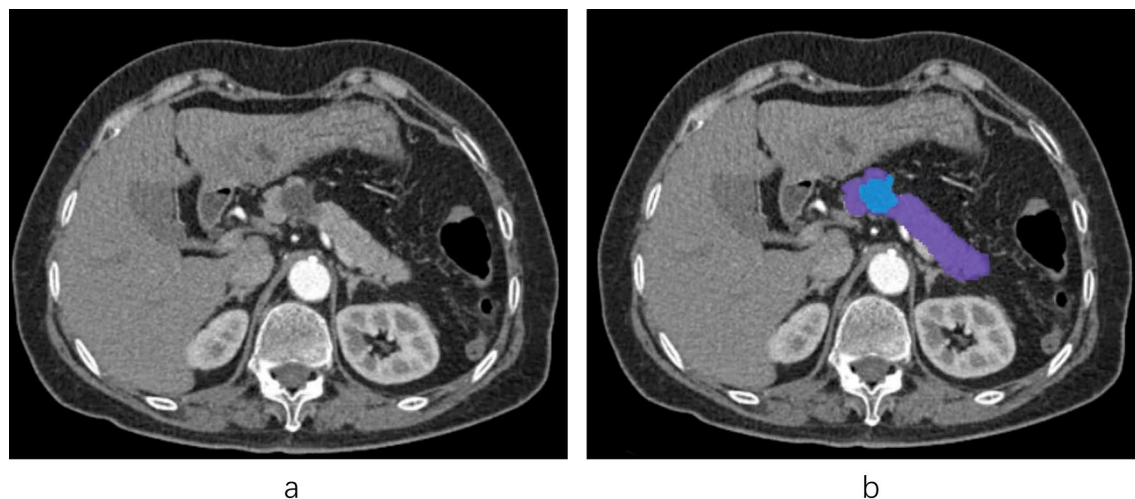


Figure 1. The segmentation of pancreatic lesion. (a) The arterial-phase enhance CT of pancreas (b) Blue part is the lesion region. Purple part is the remaining pancreatic region. The RIAS software extracts radiomic features by labeling the purple and blue areas.

Radiomics features	Coefficients
log-sigma-3-0-mm-3D_gldm_LargeDependenceLowGrayLevelEmphasis	− 0.02341
log-sigma-3-0-mm-3D_gldm_LowGrayLevelRunEmphasis	− 0.06002
logarithm_gldm_Idn	0.12689
logarithm_gldm_SmallDependenceLowGrayLevelEmphasis	− 0.03907
square_glszm_LargeAreaEmphasis	0.03838
wavelet-HHL_gldm_Imc2	0.05013
wavelet-HHH_gldm_Imc2	0.01859
wavelet-HHH_glszm_SizeZoneNonUniformityNormalized	0.12258

Table 2. The remaining arterial-phase radiomic features after LASSO algorithm.

SVM model for training and testing purposes. The model underwent fivefold cross-validation for validation. The final arterial-phase SVM model performed on the test set demonstrated the following metrics: sensitivity 0.55; specificity 0.37; precision 0.45; recall 0.47; F1 score 0.37; and accuracy 0.41; Fig. 4 shows the results of the ROC curves analyses for the arterial-phase training set, validation set, cross-validation set, and test set.

Establishment of a venous phase radiomics feature model and evaluation of predictive performance

We divided the entire sample dataset into training and testing sets at a 7:3 ratio. Employing dimensionality reduction techniques, we extracted valuable features from radiomic characteristics and built an SVM model for both training and testing. Fivefold cross-validation was implemented for validation purposes. The venous-phase SVM model performed on the test set exhibited notable performance metrics: sensitivity, 0.8; specificity, 0.74; precision, 0.72; recall, 0.74; F1 score, 0.74; and accuracy, 0.74. The results of the ROC curve analyses for the arterial-phase training set, validation set, cross-validation set, and test set are shown in Fig. 5.

Establishment and evaluation of a combined model with partial qualitative and quantitative clinical data and radiomic features

We integrated clinical features with $p < 0.05$ along with radiomic features from the venous-phase ROIs to construct a novel SVM model. Following fivefold cross-validation, the model demonstrated superior predictive performance on the test set, with a sensitivity of 0.904, a specificity of 0.74, a precision of 0.76, a recall of 0.88, an F1 score of 0.74, and an accuracy of 0.8 (Fig. 6). Compared with the radiomics feature model alone, the model in which partial qualitative and quantitative clinical data were included resulted in a significant improvement in diagnostic performance (DeLong test, P value > 0.05).

Discussion

We established a cohort of 121 patients with IPMNs and utilized radiomic features to distinguish between benign and malignant IPMNs. To address this challenging issue within current clinical and radiological practice, we aimed to develop a model with clinical practice value, that could lead to improved recommendations for patients

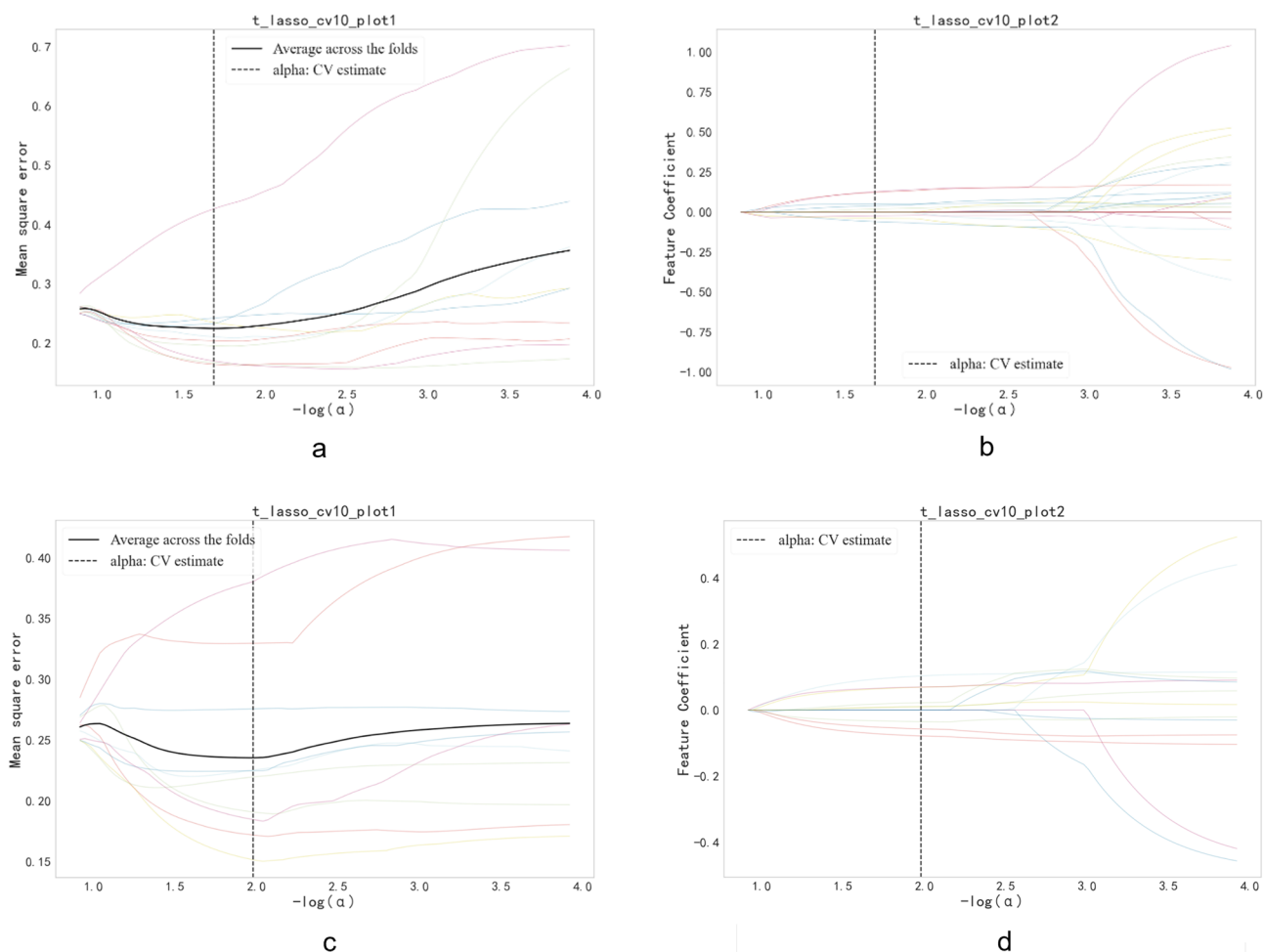


Figure 2. Parameter selection and image omics feature screening of LASSO model. (a) Parameter optimization of the arterial-phase LASSO regression model. (b) Selected feature parameters after fivefold cross-validation of the arterial-phase LASSO model. (c) Parameter optimization of the venous-phase LASSO regression model. (d) Selected feature parameters after fivefold cross-validation of the venous-phase LASSO model.

Radiomics features	Coefficients
original_shape_Compactness2	− 0.03476
log-sigma-3-0-mm-3D_gldm_LargeDependenceLowGrayLevelEmphasis	− 0.07797
logarithm_glcmm_Idn	0.069528
squareroot_glszm_SizeZoneNonUniformity	0.01097
wavelet-LLH_glcmm_Imc2	− 0.05657
wavelet-LLH_glszm_LargeAreaLowGrayLevelEmphasis	0.070131
wavelet-HHL_ngtdm_Busyness	0.01039
wavelet-HHH_glcmm_Imc2	0.103743
wavelet-LLL_glcmm_Imc2	0.022522

Table 3. The remaining venous-phase radiomic features after LASSO algorithm.

and their families and more effective management of their care. Once a patient is diagnosed with an IPMN, risk stratification can be conducted, thereby aiding in improving clinical decision-making. In patients for whom the risk of developing malignant tumours is low, surgery can be avoided, opting instead for close patient monitoring. For high-risk patients, tumour resection significantly improves patient outcomes.

In this retrospective study, the patient cohort was divided into benign and malignant groups, and a multivariate model was established, using logistic regression to assess the ability of the model to predict benign outcomes. This study diverges from traditional methods that rely solely on either venous- or arterial-phase

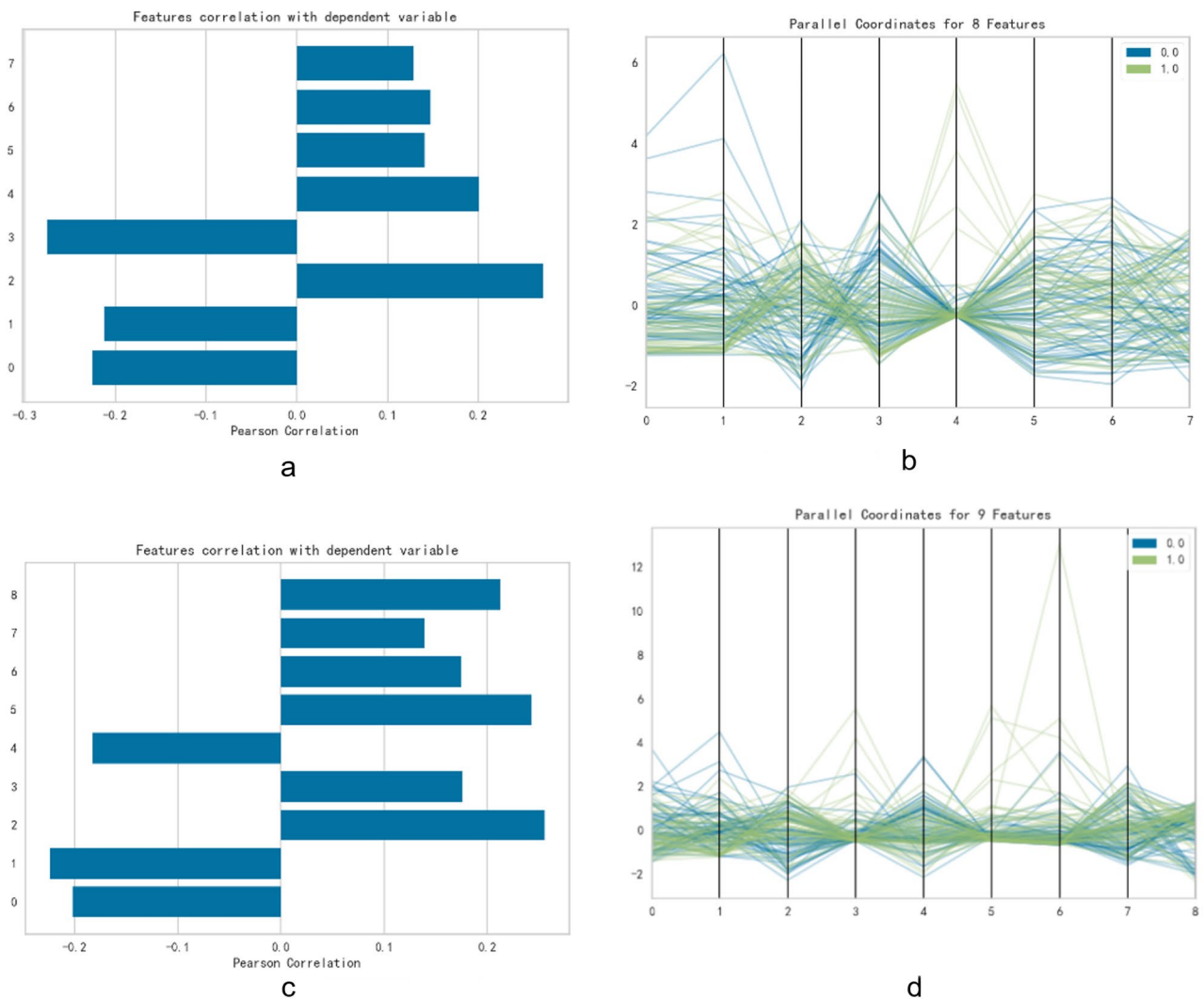


Figure 3. Feature visualization after dimensionality reduction (a,b) is the feature visualization after dimensionality reduction in arterial-phase, (c,d) is the venous-phase.

CT scans by employing multiphase enhanced CT coupled with radiomic analysis. Interestingly, we observed that each comparison period radiomic signature contained distinctive dimensional features. Specifically, the univariate features from multiple venous phases yielded superior prediction results. Moreover, the multivariate analysis revealed that the arterial-phase model included more dimensional features that may not be discernible through visual inspection and experience alone. Our findings indicate that, compared with a pure venous-phase imaging assessment, more influential features can be extracted with multiphase radiomic assessment of IPMNs.

Among the various types of IPMNs, main-duct IPMNs (MD-IPMNs) represent the most common form of malignant transformation, with up to 65% exhibiting malignancy¹³. Diabetes is associated with an increased risk of primary involvement and high-grade dysplasia or invasive cancer¹⁴. Mixed-type IPMNs (MT-IPMNs) share a similar pathophysiology to MD-IPMNs and are therefore considered to have a higher potential for malignancy compared to branch-duct IPMNs (BD-IPMNs)¹⁵. A summary of over 1500 postoperative IPMN patients revealed that jaundice and solid tumour components were significant risk factors for malignancy¹⁶. CA19-9 levels > 37 U/ml, tumour diameter > 3 cm, and elevated alkaline phosphatase are concerning indicators¹⁷. Studies have indicated that combining CA19-9 with CA125 provides a more accurate prediction of IPMN malignancy¹⁸. Furthermore, proteomic analysis of IPMN liquid tissue lysate demonstrated overexpression of tissue transglutaminase in patients with malignant disease¹⁹. Combining radiomics features with the nine selected independent risk factors significantly increased the accuracy of the model diagnosis. The AUC of our model for predicting benign and malignant IPMNs reached 0.904, surpassing the diagnostic accuracy of the International Consensus Fukuoka Guidelines.

IPMNs represent a predominant type of pancreatic cystic neoplasm (PCN), with an incidence rate reaching 70%²⁰. The increasing incidence reflects advancements in imaging technologies and improved understanding and detection of pancreatic lesions. Compared with research on PDAC grading, research on IPMN grading is considerably less extensive; furthermore, IPMN grading methods often lack validation and have insufficient diagnostic performance for practical application. Serum CA19-9 levels are associated with malignant IPMNs but

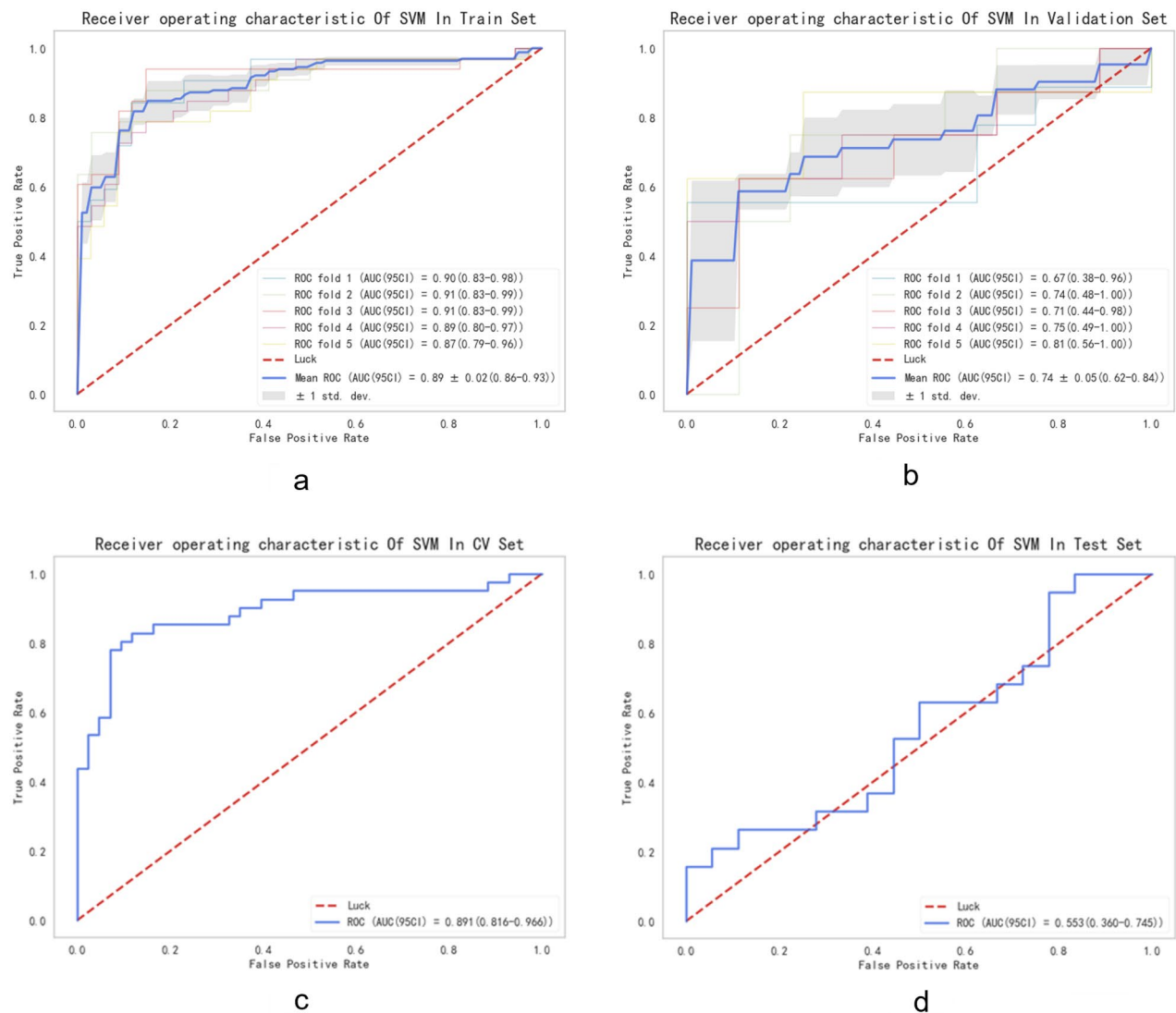


Figure 4. ROC curve of arterial-phase imaging model. **(a)** Train set, **(b)** validation set, **(c)** cross-validation set, **(d)** test set.

demonstrate limited diagnostic capability²¹, with reported AUC values ranging from 0.62 to 0.78. Serum CEA and CA19-9 concentrations combined with anti-p53 antibody positivity can distinguish high-grade dysplasia and malignant IPMNs from low-grade IPMNs, albeit with a sensitivity of only 38.4% and specificity of 81.6%²². Recent studies suggest that combining DNA methylation biomarkers can further increase diagnostic accuracy, whereas blood protein and microRNA biomarkers can facilitate the precise diagnosis and risk stratification of IPMN patients^{23,24}.

In a previous imaging study, Hanania et al.²⁵ identified grayscale co-occurrence matrix biomarkers on arterial phase 2D images, achieving 82% diagnostic accuracy in predicting malignant or high-grade diseases. Chakraborty et al.²⁶ utilized 2D radiomic texture features from venous-phase images, and reported a diagnostic accuracy of 74% in identifying high-risk IPMNs. Flammia et al.²⁷ identified guideline-relevant risk features and worrisome features via diffusion-weighted imaging (DWI)/apparent diffusion coefficient (ADC) maps of

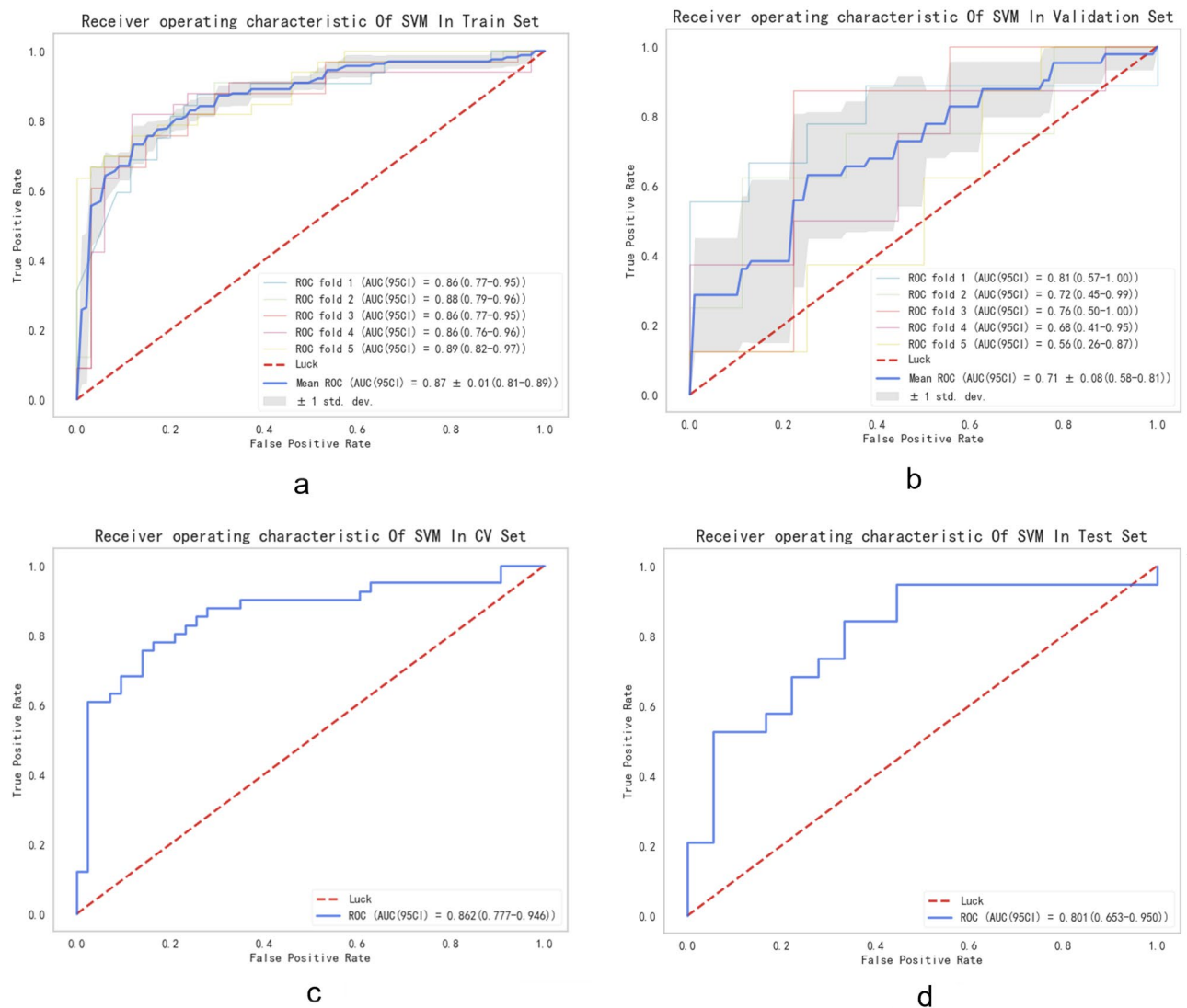


Figure 5. ROC curve of venous-phase imaging model. (a) Train set, (b) validation set, (c) cross-validation set, (d) test set.

contrast-enhanced MR images in 50 IPMN patients. Additionally, Cui and others, through in-depth analysis, constructed a joint nomogram model that combines radiomic features of MR images with clinical indicators (CA19-9 levels and MPD size) to more accurately identify branching types of tumours in IPMN patients²⁸.

Compared with these previous studies, our research boasts as a relatively large sample size, contributing to a more stable model. While prior efforts focused that primarily on 2D analysis at the level of the largest tumour cross-section, our approach incorporates a 3D total lesion analysis. This methodology allows the extraction of more comprehensive dimensional features, significantly augmenting the depth and breadth of the analysis, and thus provides a more nuanced understanding of the disease. However, it is essential to note that the data in our study were sourced from a single centre, and studies on data from multiple centres are needed to further validate the results.

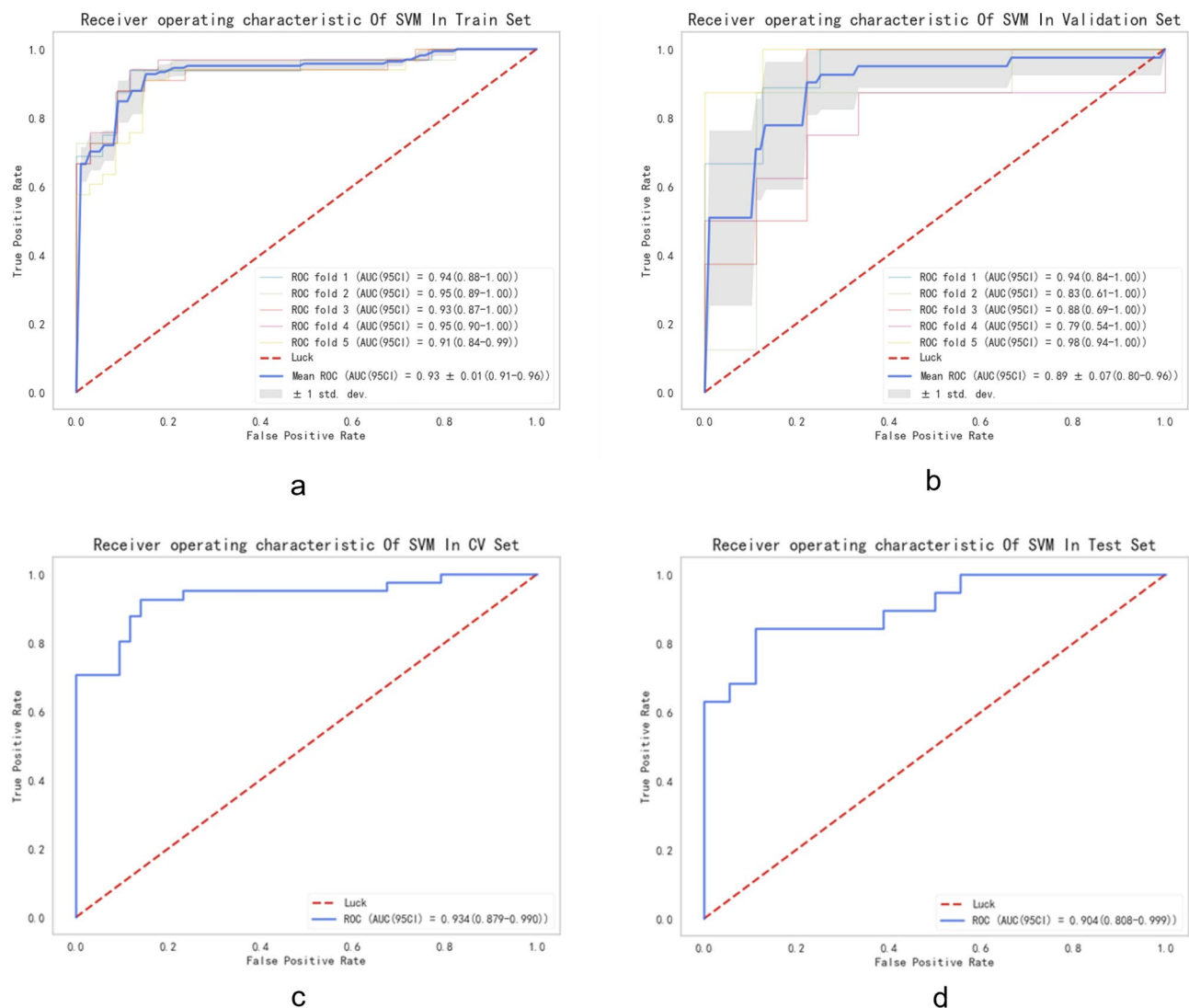


Figure 6. Clinical features combined with radiomics model. (a) Train set, (b) validation set, (c) cross-validation set, (d) test set.

Data availability

The data underlying this article cannot be shared publicly due to the privacy of individuals that participated in the study. The data will be shared on reasonable request to the corresponding author.

Received: 19 April 2024; Accepted: 19 July 2024

Published online: 26 July 2024

References

1. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. *CA Cancer J. Clin.* **72**, 7–33. <https://doi.org/10.3322/caac.21708> (2022).
2. Xia, C. *et al.* Cancer statistics in china and united states, 2022: Profiles, trends, and determinants. *Chin. Med. J.* **135**, 584–590. <https://doi.org/10.1097/CM9.0000000000002108> (2022).
3. Siegel, R. L., Giaquinto, A. N. & Jemal, A. Cancer statistics, 2024. *CA Cancer J. Clin.* **74**, 12–49. <https://doi.org/10.3322/caac.21820> (2024).
4. Nakano, Y. *et al.* Clinical predictive factors of long-term survival after curative resection of pancreatic cancer: A retrospective study. *Cancer Med.* **6**, 2278–2286. <https://doi.org/10.1002/cam4.1178> (2017).
5. Rawla, P., Sunkara, T. & Gaduputi, V. Epidemiology of pancreatic cancer global trends, etiology and risk factors. *World J. Oncol.* **10**, 10–27. <https://doi.org/10.14740/wjon1166> (2019).
6. Wood, L. *et al.* Systematic review of challenging issues in pathology of intraductal papillary mucinous neoplasms. *Pancreatology* **23**, 878–891. <https://doi.org/10.1016/j.pan.2023.08.002> (2023).
7. Rezaee, N. *et al.* Intraductal papillary mucinous neoplasm (IPMN) with high-grade dysplasia is a risk factor for the subsequent development of pancreatic ductal adenocarcinoma. *HPB* **18**, 236–246. <https://doi.org/10.1016/j.hpb.2015.10.010> (2016).
8. Capretti, G. *et al.* Invasive IPMN relapse later and more often in lungs in comparison to pancreatic ductal adenocarcinoma. *Pancreatology* **22**, 782–788. <https://doi.org/10.1016/j.pan.2022.05.006> (2022).

9. Kaiser, J. *et al.* IPMN-associated pancreatic cancer: Survival, prognostic staging and impact of adjuvant chemotherapy. *Eur. J. Surg. Oncol.* **48**, 1309–1320. <https://doi.org/10.1016/j.ejso.2021.12.009> (2022).
10. Exarchakou, A. *et al.* Pancreatic cancer incidence and survival and the role of specialist centres in resection rates in England, 2000 to 2014: A population-based study. *Pancreatol.* **20**, 454–461. <https://doi.org/10.1016/j.pan.2020.01.012> (2020).
11. Gillies, R., Anderson, A., Gatenby, R. & Morse, D. The biology underlying molecular imaging in oncology: From genome to anatomy and back again. *Clin. Radiol.* **65**, 517–521. <https://doi.org/10.1016/j.crad.2010.04.005> (2010).
12. Lambin, P. *et al.* Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer* **48**, 441–446. <https://doi.org/10.1016/j.ejca.2011.11.036> (2012).
13. Gaiser, R. A. *et al.* Enrichment of oral microbiota in early cystic precursors to invasive pancreatic cancer. *Gut* **68**, 2186–2194. <https://doi.org/10.1136/gutjnl-2018-317458> (2019) (Epub 2019 Mar 14).
14. Pergolini, I. *et al.* Diabetes and weight loss are associated with malignancies in patients with intraductal papillary mucinous neoplasms. *Clin. Gastroenterol. Hepatol.* **19**, 171–179. <https://doi.org/10.1016/j.cgh.2020.04.090> (2021).
15. Gardner, T. B., Park, W. G. & Allen, P. J. Diagnosis and management of pancreatic cysts. *Gastroenterology* <https://doi.org/10.1053/j.gastro.2024.02.041> (2024) (Epub ahead of print).
16. European Study Group on Cystic Tumours of the Pancreas. European evidence-based guidelines on pancreatic cystic neoplasms. *Gut* **67**, 789–804. <https://doi.org/10.1136/gutjnl-2018-316027> (2018) (Epub 2018 Mar 24).
17. Zelga, P. *et al.* Number of worrisome features and risk of malignancy in intraductal papillary mucinous neoplasm. *J. Am. Coll. Surg.* **234**, 1021–1030. <https://doi.org/10.1097/XCS.000000000000176> (2022) (Epub 2022 Mar 22).
18. Qian, Y. *et al.* Carbohydrate antigen 125 supplements carbohydrate antigen 19–9 for the prediction of invasive intraductal papillary mucinous neoplasms of the pancreas. *World J. Surg. Oncol.* **20**, 310. <https://doi.org/10.1186/s12957-022-02720-0> (2022).
19. Cheung, W. *et al.* Application of a global proteomic approach to archival precursor lesions: Deleted in malignant brain tumors 1 and tissue transglutaminase 2 are upregulated in pancreatic cancer precursors. *Pancreatol.* **8**, 608–616. <https://doi.org/10.1159/000161012> (2008) (Epub 2008 Oct 13).
20. Cipriani, D. *et al.* An elevated ca 19–9 is associated with invasive cancer and worse survival in IPMN. *Pancreatol.* **20**, 729–735. <https://doi.org/10.1016/j.pan.2020.04.002> (2020).
21. Tanaka, M. *et al.* Revisions of international consensus Fukuoka guidelines for the management of IPMN of the pancreas. *Pancreatol.* **17**, 738–753. <https://doi.org/10.1016/j.pan.2017.07.007> (2017).
22. Hata, T. *et al.* Serum anti-p53 antibody can serve as a predictive marker for histological grade of intraductal papillary mucinous neoplasms of the pancreas. *Pancreas* **49**, 768–773. <https://doi.org/10.1097/MPA.0000000000001570> (2020).
23. Chhoda, A. *et al.* Utility of promoter hypermethylation in malignant risk stratification of intraductal papillary mucinous neoplasms. *Clin. Epigenetics* **15**, 28. <https://doi.org/10.1186/s13148-023-01429-5> (2023).
24. Zhang, C. *et al.* Blood-based diagnosis and risk stratification of patients with pancreatic intraductal papillary mucinous neoplasm (IPMN). *Clin. Cancer Res.* **29**, 1535–1545. <https://doi.org/10.1158/1078-0432.CCR-22-2531> (2023).
25. Hanania, A. *et al.* Quantitative imaging to evaluate malignant potential of IPMNs. *Oncotarget* **7**, 85776–85784. <https://doi.org/10.18632/oncotarget.11769> (2016).
26. Chakraborty, J. *et al.* Ct radiomics to predict high-risk intraductal papillary mucinous neoplasms of the pancreas. *Med. Phys.* **45**, 5019–5029. <https://doi.org/10.1002/mp.13159> (2018).
27. Flammia, F. *et al.* Branch duct-intraductal papillary mucinous neoplasms BD-IPMNs): An MRI-based radiomic model to determine the malignant degeneration potential. *Radiol. Med.* **128**, 383–392. <https://doi.org/10.1007/s11547-023-01609-6> (2023).
28. Cui, S. *et al.* Radiomic nomogram based on MRI to predict grade of branching type intraductal papillary mucinous neoplasms of the pancreas: A multicenter study. *Cancer Imaging* **21**, 26. <https://doi.org/10.1186/s40644-021-00395-6> (2021).

Author contributions

All authors made substantial contributions to the study conception and design, and to acquisition, analysis and interpretation of data. All authors gave approval of the final version to be submitted. Fengxiang Lou undertook the data organization and the drafting of the main content of the article, while Mingyang LI handled the data processing. Other authors were responsible for data collection, screening, and statistical analysis. Feng Wei was in charge of designing and guiding the paper. All authors declare no conflicts of interest.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to F.W.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2024