Abdominal Radiology (2022) 47:4139–4150
https://doi.org/10.1007/s00261-022-03663-6

PANCREAS

Classification of pancreatic cystic neoplasms using radiomic feature analysis is equivalent to an experienced academic radiologist: a step toward computer-augmented diagnostics for radiologists

Linda C. Chu[1] · Seyoun Park[1] · Sahar Soleimani[1] · Daniel F. Fouladi[1] · Shahab Shayesteh[1] · Jin He[2] · Ammar A. Javed[3] · Christopher L. Wolfgang[3] · Bert Vogelstein[4] · Kenneth W. Kinzler[4] · Ralph H. Hruban[5] · Elham Afghani[6] · Anne Marie Lennon[6] · Elliot K. Fishman[1] · Satomi Kawamoto[1]

Received: 14 April 2022 / Revised: 21 August 2022 / Accepted: 23 August 2022 / Published online: 13 September 2022

Abstract

Purpose A wide array of benign and malignant lesions of the pancreas can be cystic and these cystic lesions can have overlapping imaging appearances. The purpose of this study is to compare the diagnostic accuracy of a radiomics-based pancreatic cyst classifier to an experienced academic radiologist.

Methods In this IRB-approved retrospective single-institution study, patients with surgically resected pancreatic cysts who underwent preoperative abdominal CT from 2003 to 2016 were identified. Pancreatic cyst(s) and background pancreas were manually segmented, and 488 radiomics features were extracted. Random forest classification based on radiomics features, age, and gender was evaluated with fourfold cross-validation. An academic radiologist blinded to the final pathologic diagnosis reviewed each case and provided the most likely diagnosis.

Results 214 patients were included (64 intraductal papillary mucinous neoplasms, 33 mucinous cystic neoplasms, 60 serous cystadenomas, 24 solid pseudopapillary neoplasms, and 33 cystic neuroendocrine tumors). The radiomics-based machine learning approach showed AUC of 0.940 in pancreatic cyst classification, compared with AUC of 0.895 for the radiologist.

Conclusion Radiomics-based machine learning achieved equivalent performance as an experienced academic radiologist in the classification of pancreatic cysts. The high diagnostic accuracy can potentially maximize the efficiency of healthcare utilization by maximizing detection of high-risk lesions.

✉ Linda C. Chu
lchu1@jhmi.edu

1 Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, Baltimore, MD, USA

2 Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD, USA

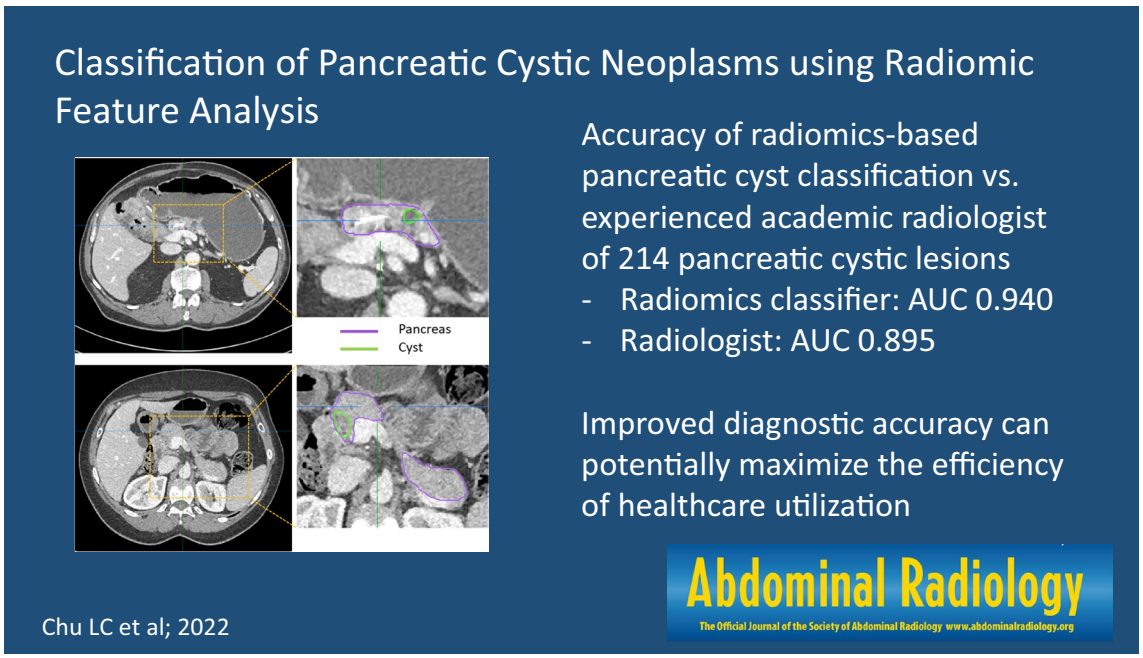
3 Department of Surgery, New York University Grossman School of Medicine, New York, NY, USA

4 Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD, USA

5 Department of Pathology, Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins University School of Medicine, Baltimore, MD, USA

6 Division of Gastroenterology and Hepatology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Graphical abstract



Keywords Pancreatic cysts · Pancreatic neoplasms · Intraductal pancreatic mucinous neoplasm (IPMN) · Radiomics · Machine learning · CT

Introduction

Pancreatic cystic lesions are frequently identified incidentally due to increased utilization of cross-sectional abdominal imaging and improvement in scanner technology. The prevalence of incidental pancreatic cysts ranges from 2% on CT to 45% on MRI, with pooled estimated prevalence of 8% [1–3]. These pancreatic cysts are formed by a diverse group of lesions, ranging from benign neoplasms to neoplasms with the potential for aggressive clinical behavior. The mucin-producing cysts (e.g., intraductal papillary mucinous neoplasms (IPMNs) and mucinous cystic neoplasms (MCNs)) are premalignant lesions that can transform into invasive pancreatic ductal adenocarcinoma (PDAC). Accurate diagnosis of pancreatic cystic lesions is important in determining appropriate treatment and surgical candidacy [4, 5]. However, these pancreatic lesions that form cysts can share overlapping clinical and radiological features making it difficult to differentiate among benign and potentially malignant lesions during preoperative evaluation. As evidenced in surgical databases, 17–25% of patients who undergo surgical resection for a presumed mucin-producing cystic lesion are found to have a benign cyst [6, 7]. Pancreatic resections are some of the most complex abdominal operations that are associated with considerable morbidity, with approximately 40% of patients experiencing postoperative complications [8]. Patients with a pancreatic cyst that does not meet criteria for surgical resection at the time of diagnosis are often followed clinically for 10 years or even longer based on current guidelines [9], which can be a significant burden to patients and an increased cost to the healthcare system.

Recently, radiomics-based approaches have been explored to differentiate pancreatic cysts. Radiomics converts imaging data into high-dimensional mineable quantitative features [10] and has shown remarkable progress correlating image features and clinical features with patient outcomes [11]. Previous studies have mostly focused on risk stratification of IPMNs by identifying radiomic signatures that are predictive of the grade of dysplasia [12–18]. Other studies aimed to discriminate between pancreatic serous cystadenoma (SCA), a benign neoplasm, from mucin-producing cystic neoplasms [19–23], mucin-producing cystic neoplasms from non-mucinous cysts [24], or among the 3–4 classes of pancreatic cystic lesions [25, 26].

To our knowledge, no study has examined a head-to-head comparison of the relative accuracy between a radiomic analysis and an expert radiologist in determining cyst type with 5 cyst classes. This is an important void in the literature as it is yet unclear if a radiomic approach using available methodology has a potential advantage over the

current standard of care. The purposes of this study are to compare the diagnostic accuracy of a radiomics-based pancreatic cyst classifier to an experienced academic radiologist and to explore the added value of radiomics-based classifier.

Materials and methods

Patients and CT acquisition

This retrospective study was HIPPA compliant and was approved by our institutional review board. A total of 214 patients (69 male, 145 female, average age: 54.8 ± 17.0 years) who underwent surgical resection for a pancreatic cyst(s) from 2003 to 2016 were randomly selected from our radiology and pathology databases, with enrichment of the rarer types of pancreatic cysts (MCNs, SCAs, SPNs, and cystic PanNETs). The selected cases include 64 patients with IPMNs, 33 MCNs, 60 SCAs, 24 SPNs, and 33 patients with cystic PanNETs (Table 1). The SCAs were resected based on clinical symptoms and/or diagnostic uncertainty in the preoperative setting. Among the 214 patients, 115 have been previously reported [25] in a study of the application of random forest and neural networks to the classification of IPMNs, MCNs, SCAs, and SPNs. In cases with multiple cysts within the pancreas, the pathologic diagnosis was labeled as the pathology of the dominant cyst.

One hundred and fifty-nine patients with cystic lesions were scanned with dual-source MDCT scanner (Somatom Definition, Definition Flash, or Force, Siemens Healthineers), 35 patients were scanned on a 64-slice MDCT scanner (Somatom Sensation 64, Siemens Healthineers), and 20 patients were scanned on a 16-slice MDCT scanner (Somatom Sensation 16, Siemens Healthineers). Patients were injected with between 100 and 120 mL of iohexol 350 (Omnipaque, GE Healthcare) at an injection rate of 4–5 mL/sec. Contrast dose weight based at a dose of approximately 1.5 mL/kg, up to dose of 120 mL. Scan protocols were customized for each patient to minimize dose but were on the order of 120 kVp, effective mAs of 270, and pitch of 0.6–0.8. The collimation was 128 × 0.6 mm or 192 × 0.6 mm or the dual-source scanner, 64 × 0.6 mm for the 64-slice scanner. Arterial phase imaging was performed with fixed delay or bolus triggering, usually between 30 and 35-s post-injection, and venous phase imaging was performed at 60–70 s. The venous phase images were used for the analysis in this study. All images were reconstructed with 0.5-mm increment and 0.75-mm slice thickness.

Image Segmentation

Preoperative CTs were reviewed by an abdominal radiologist with > 7 years of experience to document the size and location of pancreatic cysts and presence of calcifications or pancreatic duct dilatation (> 3 mm in diameter). The 214 CT exams were randomly divided between two trained researchers (3 years of experience) for image segmentation. The entire three-dimensional (3D) volume of the cystic lesion(s) and pancreas were manually segmented (Fig. 1) based on venous phase images using the Medical Imaging Interaction Toolkit (MITK) and a commercial annotation software (VelocityTM, Varian Medical Systems Inc.) [27].

Table 1 Demographic and image characteristics of the pancreatic cyst cases

	All (n = 214)	IPMN (n = 64)	MCN (n = 33)	SCA (n = 60)	SPN (n = 24)	Cystic PanNET (n = 33)	p-value
Age, mean ± SD	54.8 ± 17.0	68.2 ± 11.1	46.4 ± 12.9	56.7 ± 12.7	28.8 ± 11.4	52.8 ± 13.4	< 0.001
Sex, *N* (%) Male Female	69 (32.2) 145 (67.8)	31 (48.4) 33 (51.6)	1 (3.0) 32 (97.0)	15 (25.0) 45 (75.0)	2 (8.3) 22 (91.7)	20 (60.6) 13 (39.4)	0.0003
Cyst size (cm), mean ± SD	4.28 ± 3.25	3.24 ± 2.02	5.65 ± 5.00	4.53 ± 2.61	5.59 ± 3.05	3.53 ± 3.40	0.001
Cyst attenuation (HU)	38.67 ± 27.66	24.53 ± 12.92	21.69 ± 9.19	41.22 ± 26.47	63.41 ± 21.32	64.29 ± 38.15	< 0.001
Whole pancreas attenuation (HU)	72.11 ± 28.52	69.22 ± 25.27	66.60 ± 36.77	72.71 ± 28.98	79.23 ± 23.20	78.71 ± 27.68	0.28
Location Head/uncinate (%)	77 (36.0)	40 (62.5)	0 (0)	20 (33.3)	10 (41.7)	7 (21.2)	0.004
Neck (%)	14 (6.5)	2 (3.1)	0 (0)	8 (13.3)	1 (4.2)	3 (9.1)	
Body (%)	50 (23.4)	10 (15.6)	12 (36.4)	16 (26.7)	6 (25.0)	6 (18.2)	
Tail (%)	73 (24.1)	12 (18.8)	21 (63.6)	16 (26.7)	7 (29.2)	17 (51.5)	
Calcification, present—*N* (%)	54 (25.2)	8 (12.5)	10 (30.3)	21 (35.0)	11 (45.8)	4 (12.1)	0.044
Pancreatic duct dilatation (> 3 mm)—*N* (%)	48 (22.4)	34 (53.1)	2 (6.1)	11 (18.3)	0 (0)	1 (3.0)	< 0.001

IPMN intraductal papillary mucinous neoplasm, *MCN* mucinous cystic neoplasm, *SCA* serous cystadenoma, *SPN* solid pseudopapillary neoplasm, *PanNET* pancreatic neuroendocrine tumor, *HU* Hounsfield unit

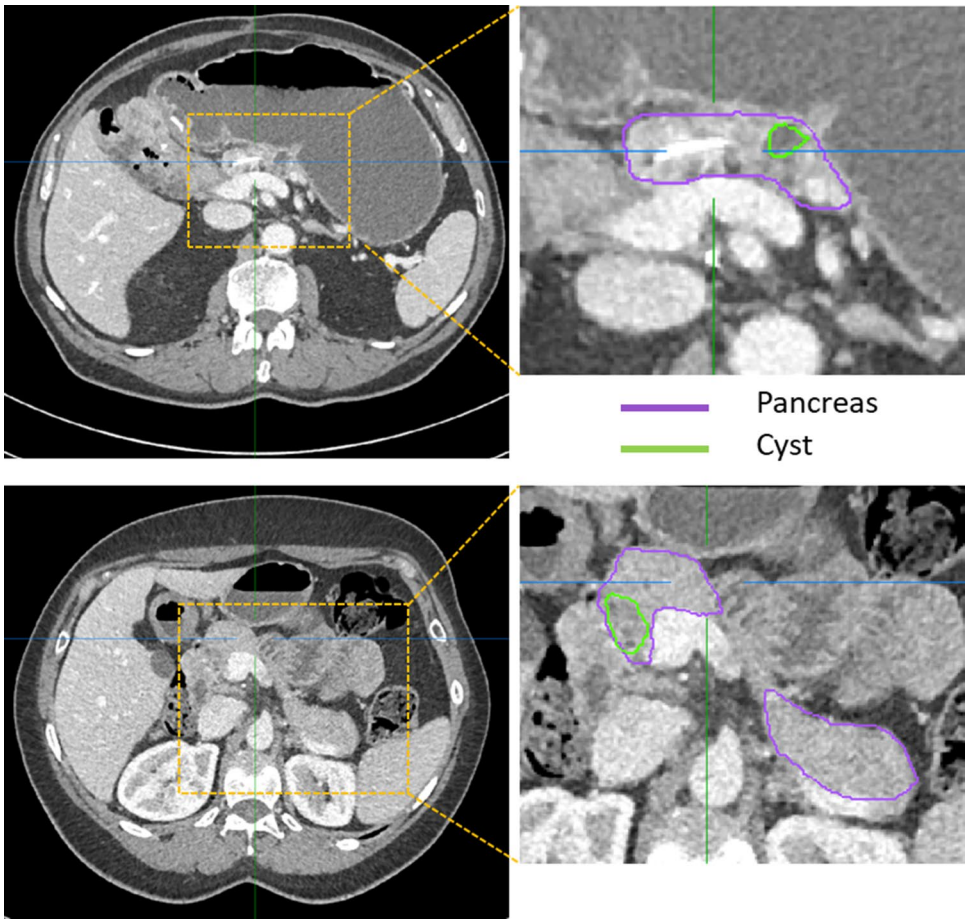


Fig. 1 Two example cases of manual segmentations of pancreas and cystic lesion. The boundary of cystic lesion is outlined in green and the boundary of the background pancreas is outlined in purple

The boundaries were verified by three abdominal radiologists with 7–30 years of experience. The researchers and the radiologists had face-to-face sessions to review each case to correct any errors in segmentation.

Computation of radiomics features can be affected by the segmentation accuracy, and the inter-observer variability in segmentation accuracy was analyzed. 20 cases among 214 cases were randomly selected and segmented by two image labelers independently. The inter-observer variation between two image labelers was evaluated by two performance parameters, Dice-Sørenson similarity coefficient (DSC) and Jaccard index (JI) to measure the similarity of two regions. 1 indicates perfect overlap and 0 indicates absence of overlap for both measures.

Image analysis and machine learning

A total of 488 radiomics features [10, 28] from the segmented volume were extracted to define cystic lesion and pancreas phenotypes based on venous phase images (Fig. 2). Radiomics features used in this study included 14 first-order statistics of the volumetric CT intensities, 8 shape features of the target structure, 33 texture features from a gray-level co-occurrence matrix and a gray-level run-length matrix, 376 texture features from the 8 filtered volumes by wavelets [28], and an additional 47 texture features form the filtered volume by Laplacian of Gaussian (LoG). Ten image features were extracted from the whole pancreatic region. Table 2 represents the whole feature set used for cyst classification in this study. Two demographic features, age and gender, were also incorporated into the final model.

To effectively test the limited number of cases of less common cyst types, such as SPNs, fourfold cross-validation was performed in this study. Each type of cystic lesion was randomly divided to four groups and each group from all 5 cyst types composed a fold. A random forest machine learning algorithm was used for cyst type classification. There were a total of a hundred thousand trees built. To test each fold and each decision node was divided until a unique case remained.

Radiologist interpretation

One academic abdominal radiologist (> 25 years of experience) who was blinded to the pathologic diagnosis reviewed the venous phase images for each case and provided their

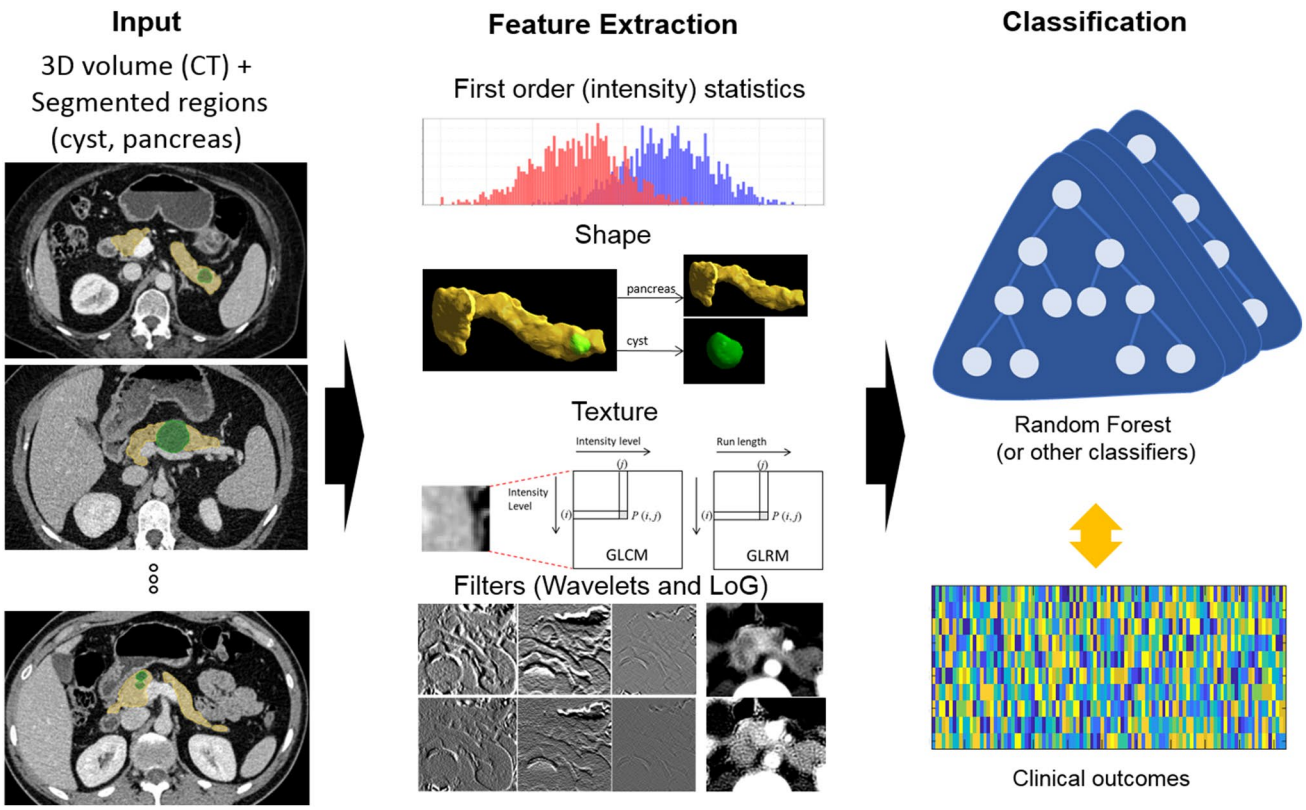


Fig. 2 The radiomics-based classification process. Pancreatic cysts and background pancreas are manually segmented from abdominal CT (left column). Feature extraction process extracts first-order signal intensity statistics, shape features based on 3D surface mask, texture features among adjacent voxels, and filtered features using wavelet or Laplacian of Gaussian filters (middle column). Features are analyzed with machine learning techniques such as random forest classifier to predict clinical outcomes (right column)

most likely diagnosis. The radiologist was provided the patient age and gender for each case, but was otherwise blinded as to the final pathology.

Statistical analysis

The diagnostic performance of the radiomics model was compared to the radiologist. The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and area under the curve (AUC) of the receiver operating characteristics (ROC) curves were calculated for each cyst type.

Results

The mean maximal 2D diameter of the cystic lesions was 4.28 ± 3.25 cm. The site of cystic lesions were the head and/or uncinate process ($n = 77$), neck ($n = 14$), body ($n = 50$), and tail ($n = 73$) of the pancreas. The demographic and cystic lesion features stratified by cyst type are summarized in Table 1. As expected, there were significant differences in the age ($p < 0.001$) and gender ($p = 0.0003$) for different cyst types. MCNs (97.0% female), SCAs (75.0% female), and SPNs (91.7% female) were more common in women, cystic PanNETs were more common in men (39.4% female), and IPMNs (51.6% female) did not have a stronger gender predilection in our sample. Patients with SPN (28.8 ± 11.4 years) were significantly younger than patients with other cyst types. MCNs were seen exclusively in the body and tail. MCNs (5.65 ± 5.00 cm) and SPNs (5.59 ± 3.05 cm) were larger than IPMNs (3.24 ± 2.02 cm), SCAs (4.53 ± 2.61 cm), and cystic PanNETs (3.53 ± 3.40 cm) ($p = 0.001$). SPNs (63.41 ± 21.32 HU) and cystic PanNETs (64.29 ± 38.15 HU) showed the highest attenuation due to the presence of solid components and IPMNs (24.53 ± 12.92 HU) and MCNs (21.69 ± 9.19 HU) showed the lowest attenuation. Pancreatic duct dilatation was most frequently associated with IPMNs (53.1%) ($p < 0.001$).

The inter-observer similarity for pancreas segmentation was $90.44 \pm 3.68\%$ and $82.74 \pm 6.10\%$ in terms of DSC and JI, respectively. The similarity of cyst segmentation was $89.04 \pm 11.88\%$ and $81.14 \pm 15.93\%$ with DSC and JI, respectively. The boxplot of inter-observer variation is represented in Fig. 3. The inter-observer variation showed high

Table 2 The demographic and image features computed from the segmented pancreatic cystic lesion and the whole pancreas

A. Demographic features (2)			
Age	Gender		
B. Radiomics features for cystic lesion phenotype (478)			
B.1. First-order statistics (14)	B.2. Shape features (8)	B.3. Texture features (33)	B.4. Features from Filtered Images (423)
Energy	Compactness1	Gray-level co-occurrence matrix based features (22)	Wavelet: LLL (47)
Entropy	Compactness2	Autocorrelation	First-order statistics (14)
Kurtosis	Maximum 3D diameter	Cluster prominence	Texture features (33)
Maximum	Spherical disproportion	Cluster shade	Wavelet: LLH (47)
Mean	Sphericity	Cluster tendency	First-order statistics (14)
Mean absolute deviation	Surface area (mm^2)	Contrast	Texture features (33)
Median	Surface-to-volume ratio	Correlation	Wavelet: LHL (47)
Minimum	Volume (cc)	Difference entropy	First-order statistics (14)
Range		Dissimilarity	Texture Features (33)
RMS		Energy	Wavelet: LHH (47)
Skewness		Entropy	First-order statistics (14)
Standard deviation		Fractal dimension	Texture Features (33)
Uniformity		Homogeneity1	Wavelet: HLL (47)
Variance		Homogeneity2	First-order statistics (14)
		IMC	Texture Features (33)
		IDMN	Wavelet: HLH (47)
		IDN	First-order statistics (14)
		Inverse variance	Texture Features (33)
		Maximum probability	Wavelet: HHL (47)
		Sum average	First-order statistics (14)
		Sum entropy	Texture Features (33)
		Sum variance	Wavelet: HHH (47)
		Variance	First-order statistics (14)
		Gray-level run-length matrix based features (11)	Texture Features (33)
		Short run emphasis	Laplacian of Gaussian (47)
		Long run emphasis	First-order statistics (14)
		Gray-level nonuniformity	Texture Features (33)
		Run-length nonuniformity	
		Run percentage	
		LGLRE	
		HGLRE	
		SRLGLE	
		SRHGLE	
		LRLGLE	
		LRHGLE	
C. Additional image features for pancreas phenotype (10)			
Mean intensity	Standard deviation of intensity	Volume	Length
Tumor volume portion	Entropy	Uniformity	Fractal dimension
Spherical distortion	Texture cluster shade		

IMC informational measure of correlation, *IDMN* inverse different moment normalized, *IDN* inverse difference normalized, *LGLRE* low gray-level run emphasis, *HGLRE* high gray-level run emphasis, *SRLGLE* short run low gray-level emphasis, *SRHGLE* short run high gray-level emphasis, *LRLGLE* long run low gray-level emphasis, *LRHGLE* long run high gray-level emphasis

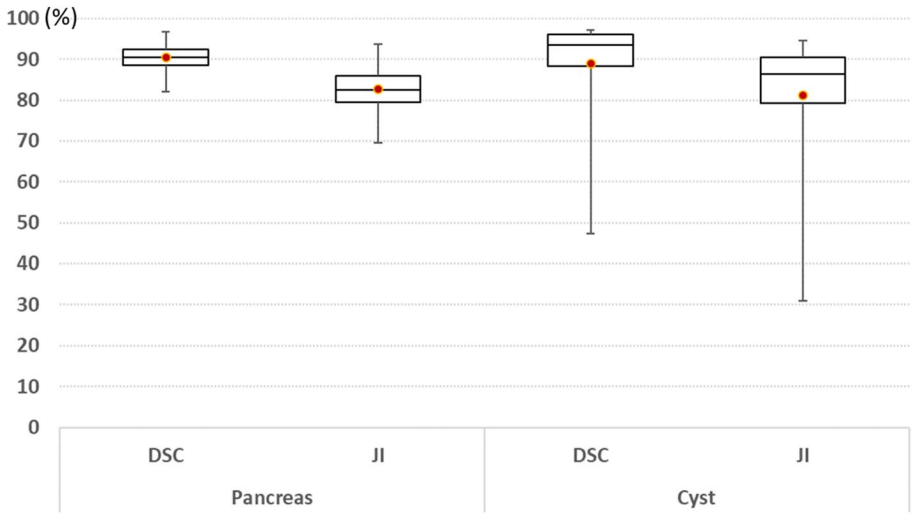


Fig. 3 The boxplot of the inter-observer variations of manual segmentation. DSC and JI percentage values are represented for the whole pancreas and pancreatic cyst lesions. The box represents the first quartile, median, and the third quartile from the lower border, middle, and the upper boarder, respectively, and the lower and the upper whiskers show the minimum and the maximum values. The dot point represents the average value of each performance measure

similarity among the labelers as the segmentation was performed under controlled protocols [27].

Among the whole 490 features (488 radiomics features plus age and gender), thirty features were found to reduce redundancy by the minimum-redundancy maximum-relevancy feature selection [29] based on mutual information, which showed the best classification performance, with AUC of 0.940 (Table 3). Age and gender were included in the model due to the known gender and gender associations for pancreatic cysts. These demographic features would be available to the radiologist at the time of exam, and this would simulate the real-world application. Age, median and mean intensities of the original images and wavelets, and fractal dimension were highly ranked for the classifications. Gender was ranked as 29th feature for the classification. The list of selected features is described in Table 4. The distribution of four representative features among the 30 selected features for each type is shown in Fig. 4.

The radiologist's interpretation of the 214 cases showed AUC of 0.895 for overall cyst classification with AUC of 0.889 for IPMNs, AUC of 0.842 for MCNs, AUC of 0.769 for SCAs, AUC of 0.865 for SPNs, and AUC of 0.908 for cystic PanNETs. The radiomics-based machine learning approach showed AUC of 0.940 for overall cyst classification and AUC of 0.942 for IPMNs, AUC of 0.883 for MCNs, AUC of 0.851 for SCAs, AUC of 0.828 for SPNs, and AUC of 0.905 for cystic PanNETs. The confusion matrix

Table 3 Confusion matrix of radiologist's and radiomics-based classification

Ground truth											
	IPMN	MCN	SCA	SPN	Cystic PNET	Sensitivity	Specificity	PPV	NPV	Accuracy	AUC
Radiologist prediction											
IPMN	58	3	13	0	3	90.6	87.3	75.3	95.6	88.3	0.889
MCN	1	24	7	0	0	72.7	95.6	75.0	95.1	92.1	0.842
SCA	4	2	37	4	2	61.7	92.2	75.5	86.1	83.6	0.769
SPN	0	2	2	18	0	75.0	97.9	81.8	96.9	95.3	0.865
Cystic PNET	1	2	1	2	28	84.9	96.7	82.4	97.2	94.9	0.908
Overall accuracy						87.3	90.6	95.6	75.3	88.3	0.895
Radiomics prediction											
IPMN	60	2	3	2	1	93.8	94.7	88.2	97.3	94.4	0.942
MCN	0	26	3	0	1	78.8	97.8	86.7	96.2	94.9	0.883
SCA	3	3	48	6	3	80.0	90.3	76.2	92.1	87.4	0.851
SPN	0	0	2	16	0	66.7	98.9	88.9	85.9	95.3	0.828
Cystic PNET	1	2	3	0	28	84.9	96.1	80.0	97.2	94.4	0.905
Overall accuracy						94.7	93.8	97.3	88.2	94.4	0.940

The radiologist and radiomics-based classification are shown in rows. The ground truth pathologic diagnosis is shown in columns

PPV positive predictive value, *NPV* negative predictive value, *AUC* area under the receiving operating characteristics curve

Table 4 Radiomics features highly dependent on cyst types

No	Category	Features	No	Category	Features
1	A	Age	16	B.4 Wavelet HHH	LTLHLR
2	B.1	Median of cyst	17	B.4 LoG	Energy
3	B.1	Mean of cyst	18	B.4 Wavelet LLL	Entropy
4	B.1	RMS	19	B.4 Wavelet LLL	Run percentage of wavelet
5	B.4 Wavelet LLL	Median intensity	20	B.4 LoG	Sum entropy
6	B.4 Wavelet LLL	Mean intensity	21	B.4 LoG	Entropy
7	B.4 Wavelet LLL	RMS	22	B.2	Compactness (type1)
8	B.4 LoG	Median intensity	23	B.2	Compactness (type 2)
9	B.4 LoG	Mean intensity	24	B.4 Wavelet HHL	Gray-level nonuniformity
10	B.4 LoG	Fractal dimension	25	C	Tumor volume portion
11	B.4 Wavelet LLL	Short run emphasis	26	B.4 Wavelet LLL	Difference entropy
12	B.4 LoG	Uniformity	27	B.4 Wavelet HHL	Run-length uniformity
13	B.4 Wavelet HHH	Long run emphasis	28	B.2	Surface-to-volume ratio
14	B.1	Maximum intensity	29	A	Gender
15	B.3	Fractal dimension	30	B.4 Wavelet LLL	Entropy

Top 30 features are listed with related categories (Table 2)

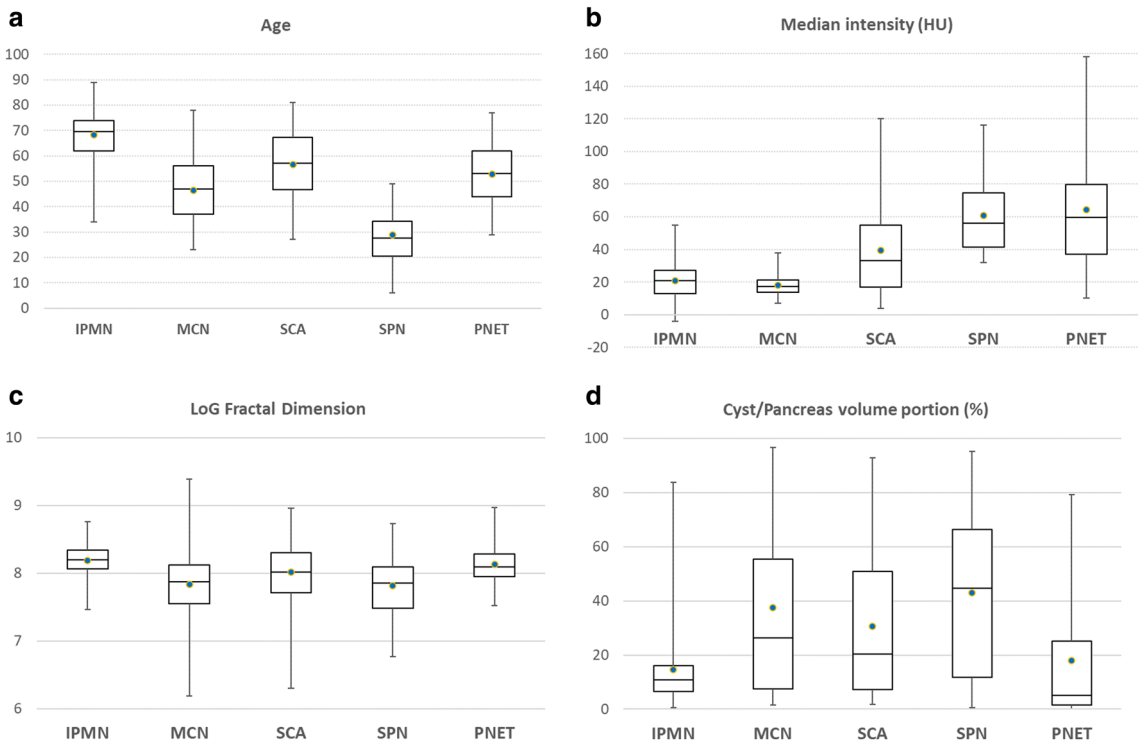


Fig. 4 The box plots of feature distributions highly correlated to differentiate cystic types, including age (**a**), median intensity (HU) (**b**), Laplacian of Gaussian fractal dimension (**c**), and cyst to pancreas volume portion (**d**)

is summarized in Table 3 and the ROC curves for radiologist and radiomics classification are shown in Fig. 5

The radiologist and radiomics feature-based classification showed similar distribution in the confusion matrix. Although the AUC for overall cyst classification of the radiomics-based classification was slightly higher than the radiologist classification (AUC of 0.940 vs. 0.895), it failed to reach statistical significance due to the small sample size. Examples with discordant predictions between radiologist and radiomics-based model are shown in Figs. 6 and 7.

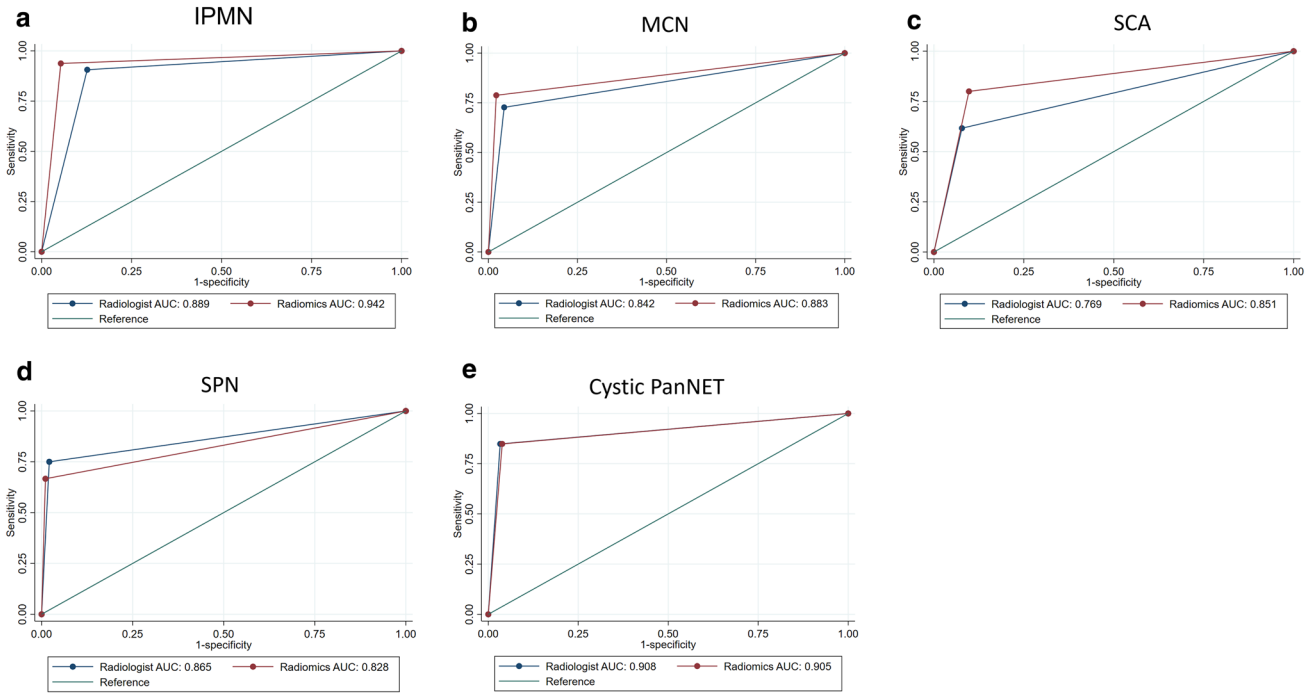


Fig. 5 The receiving operating characteristics curves between radiologist and radiomics classifications of pancreatic cysts for (**a**) intraductal papillary mucinous neoplasm, (**b**) mucinous cystic neoplasm, (**c**) serous cystadenoma, (**d**) solid pseudopapillary neoplasm, and (**e**) cystic pancreatic neuroendocrine neoplasm

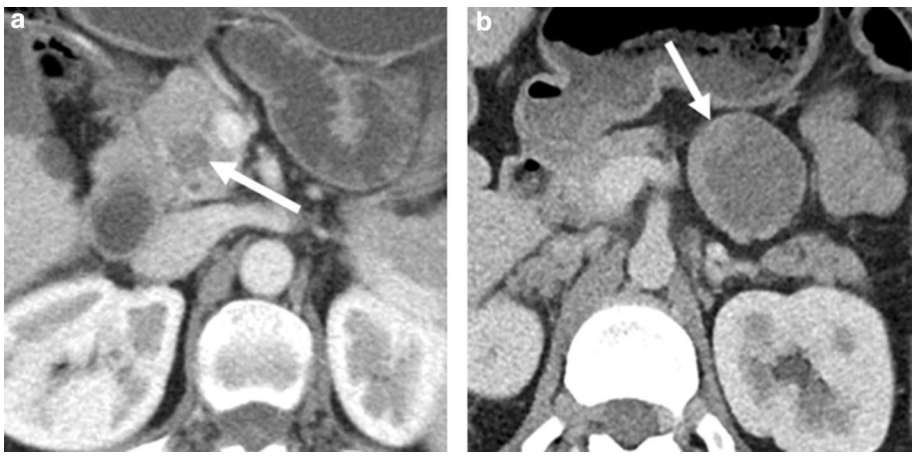


Fig. 6 Examples that radiologist's prediction was incorrect but the radiomics classification was correct. (**a**) Axial IV contrast-enhanced CT image showed a well-circumscribed lobulated cystic lesion in the pancreatic head (arrow). The radiologist's prediction of an intraductal papillary mucinous neoplasm was incorrect. Radiomics classification of serous cystadenoma was correct. (**b**) Axial IV contrast-enhanced CT image showed a well-circumscribed thick-walled exophytic cystic lesion arising from the body of pancreas (arrow). The radiologist's prediction of a cystic pancreatic neuroendocrine tumor was incorrect. Radiomics classification of a solid pseudopapillary neoplasm was correct

Discussion

Improvement in radiological techniques and increased utilization of cross-sectional imaging have led to an increase in the diagnosis of pancreatic cystic neoplasms. These neoplasms include a diverse cohort, with some lesions being benign and others having malignant potential. Given the considerable morbidity associated with pancreatic resections, accurate preoperative assessment of the lesion is vital in determining surgical candidacy. These cystic lesions present a diagnostic challenge with prior work

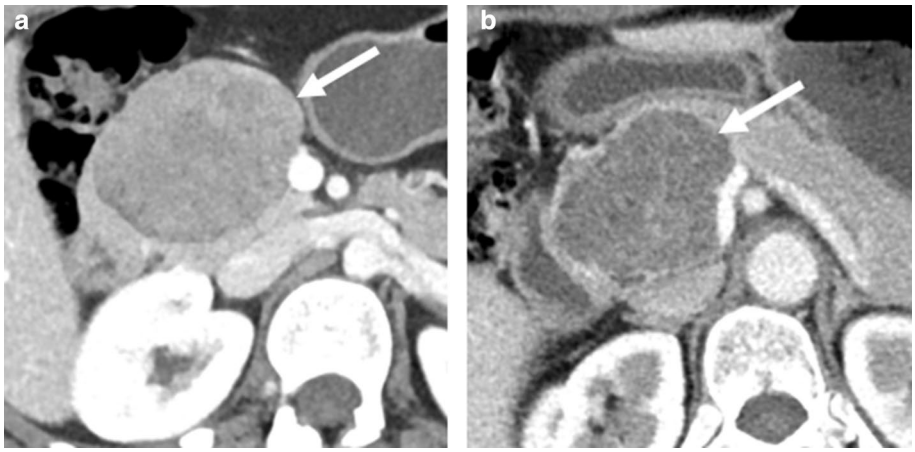


Fig. 7 Examples that the radiomics classification was incorrect and the radiologist's prediction was correct. (**a**) Axial IV contrast-enhanced CT image showed a well-circumscribed solid and cystic lesion in the pancreatic head (arrow). The radiomics classification of a serous cystadenoma was incorrect. Radiologist's prediction of a solid pseudopapillary neoplasm was correct. (**b**) Axial IV contrast-enhanced CT image showed a well-circumscribed cystic lesion with internal septations in the head of pancreas (arrow). The radiomics classification of an intraductal papillary mucinous neoplasm was incorrect. The radiologist's prediction of a serous cystadenoma was correct

demonstrating that a considerable proportion of patients undergoing resection are found on final histopathological analysis to have a benign lesion that, in retrospect, did not warrant surgical resection. While radiological assessment remains the gold standard for evaluation of these cysts, there is a need for a more powerful tool to accurately classify cyst types. In the current study we developed a radiomics feature-based classification system that was able to accurately classify cystic lesions and outperformed clinical judgment.

In this study, the performance of the radiomics feature-based classification achieved AUC of 0.940 in distinguishing among five types of pancreatic cystic neoplasms. The performance was similar to previous studies with multi-class pancreatic cyst classifications that included three or four cyst types, with accuracy of 79.6–83.6% [25, 26]. Previous studies on radiomics-based pancreatic cyst classification [19–26] did not include a direct comparison with a radiologist, therefore, it was difficult to assess if the radiomics-based classification reported provided any added value relative to the standard of care. The current study showed that the radiomics-based pancreatic cyst classification achieved equivalent performance as an academic radiologist with more than 25 years of experience. These results indicate that radiomics-based classification could be valuable in improving the current standard of care. Given that this model incorporates both clinical data and radiomic features, we believe that it is more widely applicable and comprehensive in assessment of pancreatic cysts. The radiomics-based classification showed AUC of 0.851 in the diagnosis of SCAs, which corroborated previous studies that showed AUC of 0.75–0.989 in differentiating SCAs from mucin-producing cysts [19–23]. The ability to confidently and accurately diagnosis SCAs, a "leave-alone" benign lesion, has the potential to eliminate unnecessary imaging surveillance and unnecessary surgery, which can reduce patient morbidity and healthcare costs. These radiomics-based classification systems may achieve superior performance to clinical and/or guideline-based features [14, 15]. This refined risk assessment can help with initial triage and tailor the surveillance duration and intensity to maximize the chance of cancer detection while minimizing costs. These cost savings can potentially offset costs associated with algorithm development and implementation.

This study has a few limitations. First, it was a single-center retrospective study. Fourfold cross-validation was used to assess radiomics-based model performance due to the small sample size relative to the number of cyst types. All these cases underwent surgical resection, which may bias the dataset toward atypical appearance of benign lesions (i.e., SCAs). The dataset was enriched with rarer pancreatic cyst types relative to IPMNs to evaluate the ability of the radiologist and radiomics model to discriminate among these rarer cyst types, which may limit the generalizability to the general population, in whom IPMNs are significantly more common. This study was performed on CT scanners from a single vendor. It is unclear whether variations related to scan acquisition (e.g., protocols, vendors) may affect the performance of the radiomics classification model. We only analyzed portal venous phase images in the current study, and the addition of arterial phase images may improve the accuracy of pancreatic cyst classification. MRI is frequently used in the evaluation of pancreatic cysts and can improve diagnostic confidence in the assessment of pancreatic cysts. We chose to apply

the radiomics model to CT due to greater heterogeneity with MRI (e.g., vendor, imaging sequences) compared to CT and additional normalization is needed to transform arbitrary gray intensity values from MRI. Therefore, most of the existing publication on pancreas AI have focused on CT rather than MRI. Future research is needed to validate these results with larger external datasets from different institutions and to translate results across imaging modalities. Secondly, the performance of the radiomics-based model was compared to the performance of a single-academic radiologist. The experienced academic radiologist in this study may be more accurate at pancreatic cyst classification than an average radiologist in the community, which may underestimate the incremental value of the radiomics-based model. Future reader studies should also recruit multiple readers with a wide range of experience to measure the real-world impact of these radiomics tools. Thirdly, the current radiomics model only used CT-based features plus patient age and demographics. Other important clinical features such as symptoms, family history, laboratory values, and cyst fluid molecular markers [7] were not included in the current model, which should be incorporated into future models. Our prior experience has demonstrated that the predictive power offered by multiple features is often additive and can result in a stronger model [7].

Conclusion

This study showed that a radiomics-based model can achieve equivalent performance as an experienced academic radiologist in the classification of a wide array of pancreatic cysts with variable malignant potential. This model has the potential to refine pancreatic cyst management by improving diagnostic accuracy of cystic lesions, which can minimize healthcare utilization while maximizing detection of malignant lesions. This study confirms the ability of a radiomic-based model to accurately classify pancreatic cystic neoplasms. Further validation and clinical integration of this model could help optimize management of pancreatic cysts by maximizing the rate of detection of malignant lesions while reducing healthcare utilization.

Author contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by LC, SP, SS, DF, SS, and SK. The first draft of the manuscript was written by SP and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This research was supported by the Lustgarten Foundation.

Declarations

Conflict of interest The authors declare no conflicts of interest.

References

1. Laffan, T.A., et al., *Prevalence of unsuspected pancreatic cysts on MDCT.* AJR Am J Roentgenol, 2008. **191**(3): p. 802-7.
2. Zanini, N., et al., *Estimation of the prevalence of asymptomatic pancreatic cysts in the population of San Marino.* Pancreatology, 2015. **15**(4): p. 417-22.
3. Zerboni, G., et al., *Systematic review and meta-analysis: Prevalence of incidentally detected pancreatic cystic lesions in asymptomatic individuals.* Pancreatology, 2019. **19**(1): p. 2-9.
4. Scheiman, J.M., J.H. Hwang, and P. Moayyedi, *American gastroenterological association technical review on the diagnosis and management of asymptomatic neoplastic pancreatic cysts.* Gastroenterology, 2015. **148**(4): p. 824–48 e22.
5. Elta, G.H., et al., *ACG Clinical Guideline: Diagnosis and Management of Pancreatic Cysts.* Am J Gastroenterol, 2018. **113**(4): p. 464-479.
6. Valsangkar, N.P., et al., *851 resected cystic tumors of the pancreas: a 33-year experience at the Massachusetts General Hospital.* Surgery, 2012. **152**(3 Suppl 1): p. S4-12.
7. Springer, S., et al., *A multimodality test to guide the management of patients with a pancreatic cyst.* Sci Transl Med, 2019. **11**(501): p. eaav4772.
8. Javed, A.A., et al., *Pancreatic Fistula and Delayed Gastric Emptying After Pancreatectomy: Where do We Stand?* Indian J Surg, 2015. **77**(5): p. 409-25.
9. Megibow, A.J., et al., *Management of Incidental Pancreatic Cysts: A White Paper of the ACR Incidental Findings Committee.* J Am Coll Radiol, 2017. **14**(7): p. 911-923.
10. Aerts, H.J., et al., *Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach.* Nat Commun, 2014. **5**: p. 4006.
11. Gillies, R.J., P.E. Kinahan, and H. Hricak, *Radiomics: Images Are More than Pictures, They Are Data.* Radiology, 2016. **278**(2): p. 563-77.
12. Hanania, A.N., et al., *Quantitative imaging to evaluate malignant potential of IPMNs.* Oncotarget, 2016. **7**(52): p. 85776-85784.
13. Permuth, J.B., et al., *Combining radiomic features with a miRNA classifier may improve prediction of malignant pathology for pancreatic intraductal papillary mucinous neoplasms.* Oncotarget, 2016. **7**(52): p. 85785-85797.
14. Attiyeh, M.A., et al., *Preoperative risk prediction for intraductal papillary mucinous neoplasms by quantitative CT image analysis.* HPB (Oxford), 2019. **21**(2): p. 212-218.
15. Chakraborty, J., et al., *CT radiomics to predict high-risk intraductal papillary mucinous neoplasms of the pancreas.* Med Phys, 2018. **45**(11): p. 5019-5029.
16. Polk, S.L., et al., *Multiphase computed tomography radiomics of pancreatic intraductal papillary mucinous neoplasms to predict malignancy.* World J Gastroenterol, 2020. **26**(24): p. 3458-3471.
17. Jeon, S.K., et al., *Assessment of malignant potential in intraductal papillary mucinous neoplasms of the pancreas using MR findings and texture analysis.* Eur Radiol, 2021. **31**(5): p. 3394-3404.
18. Tobaly, D., et al., *CT-Based Radiomics Analysis to Predict Malignancy in Patients with Intraductal Papillary Mucinous Neoplasm (IPMN) of the Pancreas.* Cancers (Basel), 2020. **12**(11).
19. Yang, J., et al., *Discrimination of Pancreatic Serous Cystadenomas From Mucinous Cystadenomas With CT Textural Features: Based on Machine Learning.* Front Oncol, 2019. **9**: p. 494.

20. Wei, R., et al., *Computer-Aided Diagnosis of Pancreas Serous Cystic Neoplasms: A Radiomics Method on Preoperative MDCT Images.* Technol Cancer Res Treat, 2019. **18**: p. 1533033818824339.
21. Xie, H., et al., *Preoperative differentiation of pancreatic mucinous cystic neoplasm from macrocystic serous cystic adenoma using radiomics: Preliminary findings and comparison with radiological model.* Eur J Radiol, 2020. **122**: p. 108747.
22. Chen, S., et al., *Preoperative differentiation of serous cystic neoplasms from mucin-producing pancreatic cystic neoplasms using a CT-based radiomics nomogram.* Abdom Radiol (NY), 2021.
23. Yang, R., et al., *CT classification model of pancreatic serous cystic neoplasms and mucinous cystic neoplasms based on a deep neural network.* Abdom Radiol (NY), 2022. **47**(1): p. 232-241.
24. Awe, A.M., et al., *Machine learning principles applied to CT radiomics to predict mucinous pancreatic cysts.* Abdom Radiol (NY), 2022. **47**(1): p. 221-231.
25. Dmitriev, K., et al., *Classification of Pancreatic Cysts in Computed Tomography Images Using a Random Forest and Convolutional Neural Network Ensemble.* Med Image Comput Comput Assist Interv, 2017. **10435**: p. 150-158.
26. Shen, X., et al., *A Contrast-Enhanced Computed Tomography Based Radiomics Approach for Preoperative Differentiation of Pancreatic Cystic Neoplasm Subtypes: A Feasibility Study.* Front Oncol, 2020. **10**: p. 248.
27. Park, S., et al., *Annotated normal CT data of the abdomen for deep learning: Challenges and strategies for implementation.* Diagn Interv Imaging, 2020. **101**(1): p. 35-44.
28. Chu, L.C., et al., *Utility of CT Radiomics Features in Differentiation of Pancreatic Ductal Adenocarcinoma From Normal Pancreatic Tissue.* AJR Am J Roentgenol, 2019. **213**(2): p. 349-357.
29. Peng, H., F. Long, and C. Ding, *Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy.* IEEE Trans Pattern Anal Mach Intell, 2005. **27**(8): p. 1226-38.