

Quantitative Radiomic Features From Computed Tomography Can Predict Pancreatic Cancer up to 36 Months Before Diagnosis

Wansu Chen, PhD¹, Yichen Zhou, MS¹, Vahid Asadpour, PhD¹, Rex A. Parker, MD², Eric J. Puttock, PhD¹, Eva Lustigova, MPH¹ and Bechien U. Wu, MD, MPH³

INTRODUCTION: Pancreatic cancer is the third leading cause of cancer deaths among men and women in the United States. We aimed to detect early changes on computed tomography (CT) images associated with pancreatic ductal adenocarcinoma (PDAC) based on quantitative imaging features (QIFs) for patients with and without chronic pancreatitis (CP).

METHODS: Adults 18 years and older diagnosed with PDAC in 2008–2018 were identified. Their CT scans 3 months–3 years before the diagnosis date were matched to up to 2 scans of controls. The pancreas was automatically segmented using a previously developed algorithm. One hundred eleven QIFs were extracted. The data set was randomly split for training/validation. Neighborhood and principal component analyses were applied to select the most important features. A conditional support vector machine was used to develop prediction algorithms separately for patients with and without CP. The computer labels were compared with manually reviewed CT images 2–3 years before the index date in 19 cases and 19 controls.

RESULTS: Two hundred twenty-seven of 554 scans of non-CP cancer cases/controls and 70 of 140 scans of CP cancer cases/controls were included (average age 71 and 68 years, 51% and 44% females for non-CP patients and patients with CP, respectively). The QIF-based algorithms varied based on CP status. For non-CP patients, accuracy measures were 94%–95% and area under the curve (AUC) measures were 0.98–0.99. Sensitivity, specificity, positive predictive value, and negative predictive value were in the ranges of 88%–91%, 96%–98%, 91%–95%, and 94%–96%, respectively. QIFs on CT examinations within 2–3 years before the index date also had very high predictive accuracy (accuracy 95%–98%; AUC 0.99–1.00). The QIF-based algorithm outperformed manual rereview of images for determination of PDAC risk. For patients with CP, the algorithms predicted PDAC perfectly (accuracy 100% and AUC 1.00).

DISCUSSION: QIFs can accurately predict PDAC for both non-CP patients and patients with CP on CT imaging and represent promising biomarkers for early detection of pancreatic cancer.

SUPPLEMENTARY MATERIAL accompanies this paper at <http://links.lww.com/CTG/A892>

Clinical and Translational Gastroenterology 2023;14:e00548. <https://doi.org/10.14309/ctg.0000000000000548>

INTRODUCTION

Pancreatic cancer is the third leading cause of cancer deaths in the United States with 48,220 estimated deaths in 2021 (1). The 5-year survival in 2012–2017 was only 10.8% (1). Pancreatic ductal adenocarcinoma (PDAC) is the most common form of pancreatic cancer accounting for up to 90% of all cases. Early detection of PDAC is difficult owing to lack of specific

symptoms or established screening. As a result, nearly 50% of cases have distant metastases at the time of diagnosis. Given the challenges with early detection in PDAC, the Scientific Framework for PDAC issued by the National Cancer Institute called for an evaluation of longitudinal screening protocols including imaging biomarkers for patients at high risk of PDAC (2).

¹Kaiser Permanente Southern California Research and Evaluation, Pasadena, California, USA; ²Department of Radiology, Los Angeles Medical Center, Southern California Permanente Medical Group, Los Angeles, California, USA; ³Center for Pancreatic Care, Department of Gastroenterology, Los Angeles Medical Center, Southern California Permanente Medical Group, Los Angeles, California, USA. **Correspondence:** Wansu Chen, PhD. E-mail: Wansu.Chen@KP.org.

Received March 18, 2022; accepted October 18, 2022; published online November 26, 2022

© 2022 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of The American College of Gastroenterology

A key step for early detection is the ability to identify precursors of PDAC on conventional cross-sectional imaging. Abnormalities of the pancreas such as main duct dilatation may be early indicators of PDAC and can be detected on computed tomography (CT) with a high degree of reproducibility (3–5). However, these findings are often nonspecific, and improved methods are needed to identify accurate and reliable early indicators of pancreatic cancer on cross-sectional imaging.

Automated radiomic analysis of quantitative imaging features (QIFs) (6) abstracted directly from cross-sectional imaging offers a promising opportunity to objectively identify precursor findings related to pancreatic cancer. Radiomics analysis has been used to predict the survival of patients with PDAC (7–15); differentiate functional abdominal pain, recurrent acute pancreatitis, and chronic pancreatitis (16); and distinguish autoimmune pancreatitis from PDAC (17). One study attempted to classify PDAC cases from normal pancreas based on QIFs of CT images after cancer diagnosis (18); however, the ability to identify precursor lesions before cancer development remains a key step to early detection. A review including 70 studies concluded that “radiomics of the pancreas holds promise as a quantitative imaging biomarker of both focal pancreatic lesions and diffuse changes of the pancreas” (19). However, studies using QIFs to predict PDAC are limited.

Patients with chronic pancreatitis (CP) represent a select group at increased risk of PDAC with a nearly 8-fold risk of developing pancreatic cancer in the 5 years after diagnosis, compared with individuals without CP (20). Because imaging abnormalities such as calcifications, glandular atrophy, and pancreatic duct dilatation are common among patients with CP, early diagnosis of pancreatic cancer can be particularly challenging in this patient population (21).

The objective of this study was to determine the ability of radiomics-based direct image analysis to identify changes in the pancreas on cross-sectional imaging associated with the subsequent development of pancreatic cancer. Specifically, we sought to develop automated computer algorithms for patients with and without CP to identify important QIFs and subsequently assess performance for prediction of pancreatic PDAC overall and within time-specific intervals before development of cancer.

METHODS

Study design and setting

We conducted a nested case-control study, including cross-sectional abdominal CT images from health plan enrollees of Kaiser Permanente Southern California (KPSC). KPSC is an integrated healthcare system that provides comprehensive healthcare services for more than 4.8 million enrollees across 15 medical centers and 250+ medical offices throughout the Southern California region. KPSC health plan enrollees are broadly representative of the Southern California population at large in diversity in race/ethnicity, socioeconomic status, and other demographics (22). The study protocol was approved by the KPSC’s Institutional Review Board.

Eligible study subjects and CT scans

We identified adults 18 years and older diagnosed with PDAC in 2008–2018 (index date, t_0) based on the KPSC Cancer Registry by using the *International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM)* code C25.x and histology codes listed in Supplementary Document 1 (see

Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>). The KPSC Cancer Registry is a prospectively maintained registry and part of the Surveillance, Epidemiology, and End Results reporting program. Patients with a history of pancreatectomy or those who did not have at least 12 months of health plan enrolment before t_0 were excluded (a gap of ≤ 45 days was allowed). CT scans 3 months–3 years before t_0 of the eligible PDAC cases were obtained and were matched to up to 2 scans of controls by age, sex, race/ethnicity, CT contrast status, and year of scan (± 2 years). Controls met the same eligibility criteria as above but were pancreatic cancer-free up to t_0 of the matched cases. All the CT scans were manually reviewed to remove those that did not capture the entire pancreas or had formatting errors. The resolution of scans was 512×512 pixels with slice thickness in the range of 2.5–5 mm. We separately formed (i) a group of cases/matched controls without a history of acute pancreatitis (*ICD-9: 577.0; ICD-10: K85.x*) or CP (*ICD-9: 577.1; ICD-10: K86.0, K86.1*) and (ii) a group of cases/matched controls with a history of CP, regardless of a history of acute pancreatitis. The same algorithm for training, validation, and testing processes described below applied to each group separately.

Characteristics of study subjects

Patient demographics, behavioral characteristics, and clinical characteristics on t_0 or in the 12 months before t_0 were extracted. Scan indications and the associated diagnosis for the visit were also captured. The tumor size was determined by manual review of radiology notes within the window of 3 months at the time of cancer diagnosis. The definitions of the clinical features are described in Supplementary Document 2 (see Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>).

Pancreas segmentation

The method to automatically extract the volumetric shape of the pancreas was previously described (23). When the same method was applied to the images of this study, we started from the estimated parameters previously derived and adjusted them to fit the data of this study (see the method of enhancement in Supplementary Document 3, Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>). To evaluate the performance of the adjusted algorithm, we calculated dice similarity coefficient based on 9 randomly selected scans of PDAC cases without CP by comparing the automated segmentation and that of manually delineated by the study radiologist (R.A.P.).

Extraction of QIFs

First, we normalized the center and width of the intensity window of all the scans. One hundred eleven previously validated QIFs (24) (see Supplementary Document 4, Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>) were extracted from the segmented areas of the pancreas using MATLAB software (25). Finally, the QIFs were standardized such that they all had a mean value of zero and standard deviation of 1 (26).

Algorithm development and validation

Preparation of training and validation data sets. The non-CP image set was randomly split for training (50%, DS1) and validation (50%, DS2). DS2 was further divided into 5 subsets based on the temporal distance between the scan date and t_0 (or t_0 of the matched case): 3–6, 6–12, 12–18, 18–24, and 24–36 months. By design, there is no overlap between training and validation data

sets. The CP image set was randomly split for training (50%, DS3) and validation (50%, DS4).

QIF selection method. Two competing methods of feature selection were applied and compared based on DS1. The neighborhood component analysis (NCA) algorithm is a nonparametric method aiming for maximum prediction accuracy (27,28). Using a regularization parameter (λ), the process of NCA learns feature weights by minimizing the expected leave-one-out classification accuracy. To avoid overfitting, we tuned the regularization parameter λ based on 5-fold cross-validation. The λ corresponding to the minimum classification loss was selected as the best. QIF with a weight greater than 2% of the maximum weight under the best λ was deemed important (28) but could be eliminated if its contribution to area under the curve (AUC) was less than 0.001 in a backward elimination process. In addition, we applied a second approach for feature selection: principal component analysis (PCA), to transform all 111 QIFs into linear combinations that are orthogonal to one another (29). Principal components with eigenvalues ≥ 1 were considered important. The analyses were performed by the *fscna* function in MATLAB (30) and the *prcomp* function in R (31).

Algorithm development. A conditional support vector machine (SVM) (32) was applied to develop prediction algorithms based on DS1. The reason for using conditional SVM is to account for the paired structure of the matched data. More specifically, we centered the within-pair data by its mean (32). For example, for a specific feature, the values were 0.4 and 0.6, respectively; from the images of a case and one of his/her controls, the centered values became -0.1 and 0.1 . SVM is a high-performing nonlinear classifier to map input data into higher dimensional space with the purpose of better ability of data separation (33). Moreover, SVM can ignore outliers. The kernel functions used in this study included Gaussian, linear, and sigmoid (34,35). The 2 hyperparameters were tuned based on 5-fold cross-validation (see Supplementary Document 5, Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>).

Algorithm validation. The performance of the prediction algorithms was evaluated using sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and AUC based on DS2. Accuracy was defined as the total number of correctly predicted individuals divided by the total number of patients. The validation was performed using the entire validation data set and repeated in each of the 5 validation subsets.

Sensitivity analysis. To understand whether the algorithms have the potential to specifically predict early-stage cancer, we validated the algorithm developed based on principal components analysis in the non-CP data set limited to cases diagnosed with stage I PDAC and their matched controls. The same performance measures mentioned above were estimated. The analysis was not performed in CP cases and controls because of the small sample size.

Exploratory analyses. Because the selection of QIFs by NCA may vary largely from one data set to another, we conducted exploratory analyses based on non-CP cases and controls to understand how the instability of QIF selection may affect the performance of final prediction algorithms (see Supplementary Document 6, Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>).

Clinical evaluation

A blind manual rereview of scans of 100% of non-CP cases and 50% of non-CP controls in the validation data set 24–36 months before t_0 was performed by the study radiologist (R.A.P.) to determine the risk of PDAC. Patients were classified as low risk if they only had diffuse atrophy, smaller cyst, simple cyst, loss of normal lobulation, or mild diffuse duct dilatation/prominence or did not have any obvious morphological features. Patients with complex cyst, cyst larger than 2 cm, or loss of normal lobulation of contour or having 2 or more low-risk features were considered having medium risk. Those with solid mass, focal abnormal enhancement, focal duct stricture, or focal/segmental atrophy were deemed high risk. The reviewer was not informed about the computer labels or case/control status at the time of review. The manual risk assessment was compared with the assigned computer labels (high risk “50%+” vs low risk “<50%”) based on the prediction algorithm developed with the Gaussian kernel function and PCA.

Statistical analysis

Characteristics of cases and controls were compared by using a conditional logistic regression model to account for the matched design. The *P* values were based on the likelihood ratio test taking the nature of matching into consideration. AUC measures were estimated by the MLR package in R (36), and their 95% confidence intervals (CIs) were estimated (37). All the analyses were performed using SAS (38), except for the MATLAB (39) or R packages (40) mentioned previously.

RESULTS

Characteristics of the study cohort

Patients without CP. This study included 277 scans of PDAC cases and 554 matching scans of controls (Figure 1). 34.7% cases without CP were American Joint Committee on Cancer stages I–II, 7.6% stage III, 36.8% stage IV, and 21% stage unknown at the time of diagnosis (Table 1). The median tumor size was 3.3 cm (interquartile range 2.4–4.2) in the 148 cases with known information. On an average, the patients were aged 70.8 years and 50.5% were women. 47.3% were non-Hispanic Whites, 26.7% were Hispanic, 18.4% were African American, and 7.6% were Asian/Pacific Islanders (Table 1). Ever tobacco use was frequent in both cases (57.1%) and controls (51.7%) ($P = 0.14$). Family history of pancreatic cancer and diabetes were more common in cases than in controls (Table 1). Body mass index and weight change in 1 year were comparable between cases and controls. Among specified scan indications and associated diagnosis for the visit, abdominal pain, other pain, and GI problems appeared most frequently (see Supplementary Document 7, Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>).

Patients with CP. Seventy scans of PDAC cases and 140 matching scans of controls were included (Figure 1). 35.7% cases with CP were stages I–II, 1.4% stage III, 28.6% stage IV, and 34.3% stage unknown at the time of diagnosis (Table 1). Higher frequencies of Medicare and private pay insurance coverage, ever tobacco usage, and alcohol abuse were observed in controls than in cases, although the differences were not statistically significant at the 95% level. Scan indications and associated diagnosis can be found in Supplementary Document 7 (see Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>).

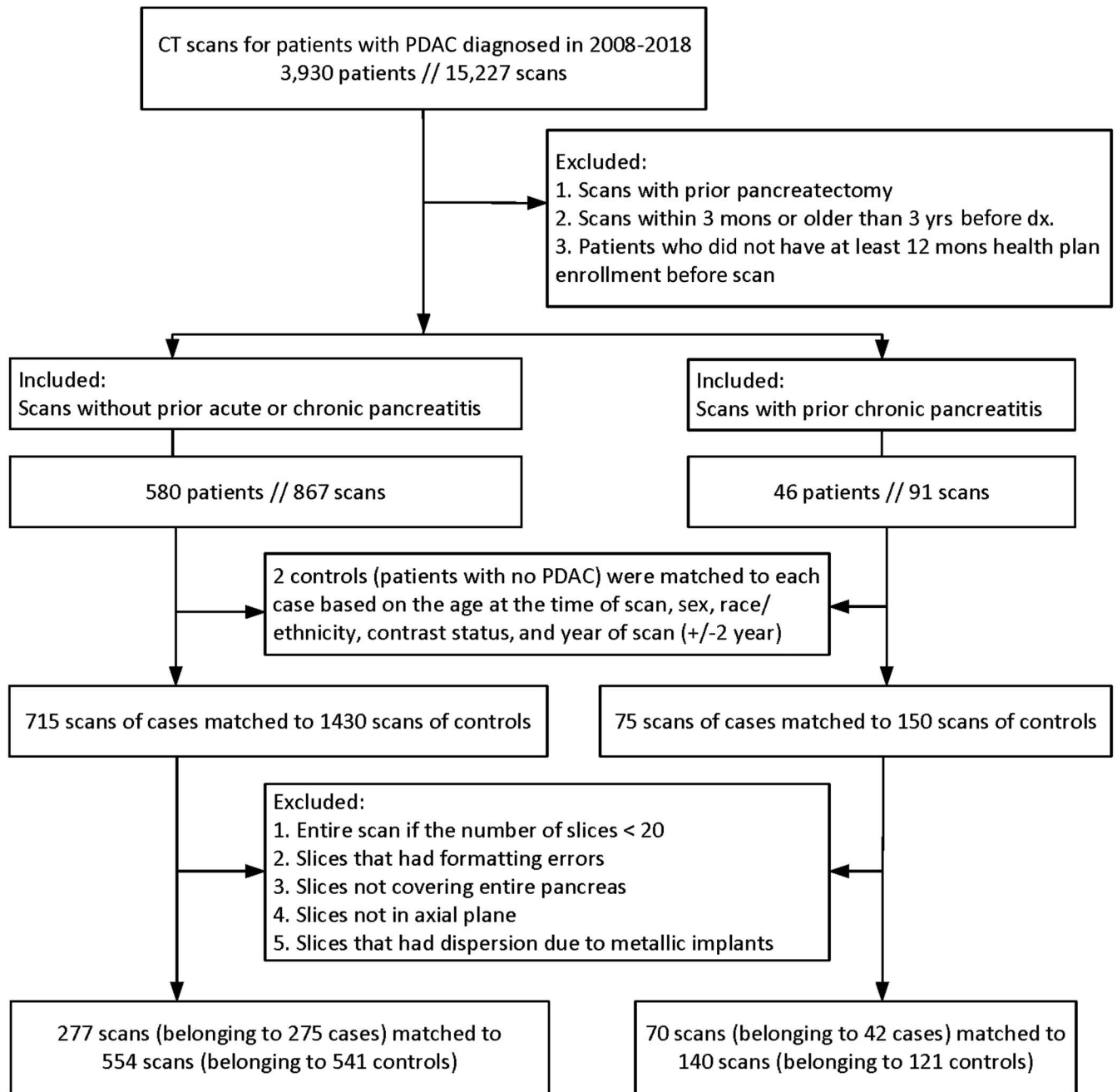


Figure 1. Consort diagram. CT, computed tomography; PDAC, pancreatic ductal adenocarcinoma.

Pancreas segmentation

The average of dice similarity coefficient from the automated pancreas segmentation algorithm was 83.25 (SD 5.19). Two example images with automated pancreas segmentation and manual group reference segmentation are presented in Figure 2.

Model training and validation

Patients without CP. For NCA, 5 QIFs were selected (see Supplementary Document 8, Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>). For PCA, 19 principal components were formed (see Supplementary Document 9, Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>).

Tables 2a and 2b summarize the performance measures of the prediction algorithms based on the QIF selected by NCA and the principal components formed by PCA, respectively.

Performance based on entire scan window (3 months–3 years before t_0). Accuracy/AUC based on DS2 were 94%/0.98 and 95%/0.99, respectively, for prediction algorithms developed based on NCA and PCA, regardless of kernel functions applied. Sensitivity, specificity, PPV, and NPV based on DS2 were in the ranges of 88%–91%, 96%–98%, 91%–95%, and 94%–96%, respectively, varying slightly by QIF selection method (NCA vs PCA) and kernel function used to develop the prediction algorithms. Algorithms developed based on PCA had better performance than

Table 1. Characteristics of study subjects at baseline by CP diagnosis and PDAC case/control status, n (%) unless otherwise stated

Patient characteristics	Patients without CP			Patients with CP		
	Cases (N = 277)	Controls (N = 554)	P value	Cases (N = 70)	Controls (N = 140)	P value
Age, mean (SD), years	70.8 (10.33)	70.8 (10.32)	NA	68.4 (9.20)	68.3 (9.19)	NA
Female	140 (50.5)	280 (50.5)	NA	31 (44.3)	62 (44.3)	NA
Race/ethnicity			NA			NA
Non-Hispanic White	131 (47.3)	262 (47.3)		33 (47.1)	66 (47.1)	
African American	51 (18.4)	102 (18.4)		19 (27.1)	38 (27.1)	
Hispanic	74 (26.7)	148 (26.7)		16 (22.9)	32 (22.9)	
Asian and Pacific Islanders	21 (7.6)	42 (7.6)		2 (2.9)	4 (2.9)	
Insurances (mutually inclusive)						
Commercial	88 (31.8)	179 (32.3)	0.94	31 (44.3)	48 (34.3)	0.27
Medi-CAL	9 (3.2)	19 (3.4)	0.89	0 (.)	9 (6.4)	1.00
Medicare	185 (66.8)	370 (66.8)	1.00	39 (55.7)	89 (64.6)	0.12
Private pay	104 (37.5)	238 (43)	0.12	16 (22.9)	41 (39.3)	0.48
Years of health plan enrolment, mean (SD)	27.5 (16.27)	25.5 (13.58)	0.051	19.7 (11.70)	20.5 (13.72)	0.61
Tobacco use			0.14			0.67
Ever	157 (57.1)	286 (51.7)		40 (57.1)	96 (68.6)	
Never	118 (42.9)	267 (48.3)		30 (42.9)	43 (30.7)	
Unknown	2 (.)	1 (.)		0 (.)	1 (0.7)	
Diagnosis of alcohol abuse in the past year	13 (4.7)	28 (5.1)	0.91	18 (25.7)	29 (41.4)	0.25
Diagnosis of alcohol abuse any time in the past	30 (10.8)	51 (9.2)	0.61	38 (27.1)	51 (36.4)	0.14
Family history of pancreatic cancer	18 (6.5)	5 (0.9)	<0.001	3 (4.3)	7 (5.0)	0.62
BMI			0.88			0.42
Underweight (<18.5)	7 (2.5)	22 (4)		4 (5.7)	13 (9.3)	
Normal weight (18.5–24.9)	79 (28.5)	166 (30)		29 (41.4)	54 (38.6)	
Overweight (25–29.9)	100 (36.1)	187 (33.8)		21 (30.0)	30 (21.4)	
Obese (30+)	86 (31)	169 (30.5)		16 (22.9)	39 (27.9)	
Unknown	5 (1.8)	10 (1.8)		0 (.)	4 (2.9)	
Weight change in 1 yr(kg)			0.07			0.10
Median (IQR)	−2.7 (−5.4 to 0.9)	−2.1 (−4.5 to 1.5)		−2.3 (−7.3 to 0.9)	−1.4 (−6.5 to 1.8)	
≤−6 kg	50 (18.1)	89 (16.1)		18 (25.7)	35 (25.0)	
>−6 and ≤−4 kg	23 (8.3)	38 (6.9)		5 (7.1)	11 (7.9)	
>−4 and ≤−2 kg	34 (12.3)	71 (12.8)		10 (14.3)	13 (9.3)	
>−2 and <2 kg	87 (31.4)	167 (30.1)		22 (31.4)	39 (27.9)	
≥2 and <4 kg	22 (7.9)	44 (7.9)		5 (7.1)	9 (6.4)	
≥4 kg	14 (5.1)	60 (10.8)		3 (4.3)	18 (12.9)	
Unknown	47 (17)	85 (15.3)		7 (10.0)	15 (10.7)	
Diabetes	122 (44)	191 (34.5)	0.004	34 (48.6)	74 (52.9)	0.47
Biliary tract disease	58 (20.9)	114 (20.6)	0.62	31 (44.3)	83 (59.3)	0.47
Depression	87 (31.4)	181 (32.7)	0.78	22 (31.4)	73 (52.1)	0.27
Deep vein thrombosis	9 (3.2)	37 (6.7)	0.08	17 (24.3)	22 (15.7)	0.30
Gallstone disorders	47 (17)	98 (17.7)	0.89	20 (28.6)	68 (48.6)	0.68
Hereditary cancer syndromes	37 (13.4)	89 (16.1)	0.35	21 (30.0)	32 (22.9)	0.81
Peptic ulcer	5 (1.8)	16 (2.9)	0.25	8 (11.4)	37 (26.4)	0.30

Table 1. (continued)

Patient characteristics	Patients without CP			Patients with CP		
	Cases (N = 277)	Controls (N = 554)	P value	Cases (N = 70)	Controls (N = 140)	P value
AJCC stage at diagnosis			NA			NA
I	36 (13.0)			6 (8.6)		
II	60 (21.7)			19 (27.1)		
III	21 (7.6)			1 (1.4)		
IV	102 (36.8)			20 (28.6)		
Unknown ^a	58 (21.0)			24 (34.3)		
T Stage at diagnosis			NA			NA
T0	1 (0.4)			0 (.)		
T1	15 (5.4)			0 (.)		
T2	67 (24.2)			11 (15.7)		
T3	63 (22.7)			21 (30.0)		
T4	38 (13.7)			8 (11.4)		
TX	60 (21.7)			16 (22.9)		
Unknown ^a	33 (11.9)			14 (20.0)		

^aUnknown includes not recorded or unavailable (e.g., diagnosed outside of Kaiser Permanente Southern California). Cancer stage is not recorded for cases diagnosed in 2018.

AJCC, American Joint Committee on Cancer; BMI, body mass index; CP, chronic pancreatitis; IQR, interquartile range; NA, not available; PDAC, pancreatic ductal adenocarcinoma.

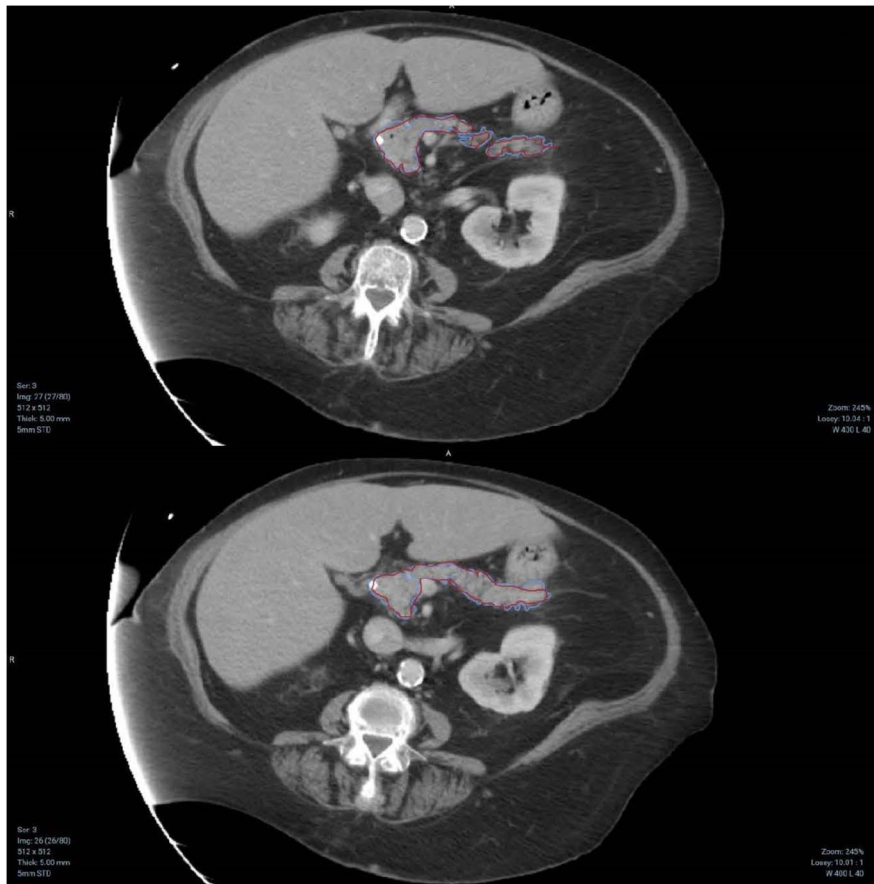


Figure 2. Two example images with automated pancreas segmentation (blue) and manual group truth segmentation (red).

Table 2. Performance of conditional SVM classifier with various kernel functions based on (a) the 5 features selected by NCA and (b) the 19 principal components formed by PCA for patients without chronic pancreatitis

(a) Data set	Kernel	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	AUC
Training: DS1 (3 mo–3 yr) N = 414	Gaussian	97.1 (94.3–99.9)	98.6 (97.1–100.0)	97.1 (94.3–99.9)	98.6 (97.1–100.0)	98.1 (96.7–99.4)	0.998 (0.993–1.000)
	Linear	94.9 (91.3–98.6)	97.1 (95.1–99.1)	94.2 (90.4–98.1)	97.5 (95.6–99.3)	96.4 (94.6–98.2)	0.994 (0.985–1.000)
	Sigmoid	91.3 (86.6–96.0)	96.7 (94.6–98.8)	93.3 (89.1–97.5)	95.7 (93.3–98.1)	94.9 (92.8–97.0)	0.983 (0.967–0.998)
Validation: DS2 (3 mo–3 yr) N = 417	Gaussian	89.2 (84.1–94.4)	95.7 (93.3–98.1)	91.2 (86.4–95.9)	94.7 (92.0–97.3)	93.5 (91.2–95.9)	0.977 (0.960–0.995)
	Linear	88.5 (83.2–93.8)	96.4 (94.2–98.6)	92.5 (88.0–97.0)	94.4 (91.7–97.0)	93.8 (91.4–96.1)	0.984 (0.969–0.999)
	Sigmoid	87.8 (82.3–93.2)	97.5 (95.6–99.3)	94.6 (90.7–98.5)	94.1 (91.4–96.8)	94.2 (92.0–96.5)	0.978 (0.960–0.995)
Validation (3–6 mo) N = 87	Gaussian	86.2 (73.7–98.8)	87.9 (79.5–96.3)	78.1 (63.8–92.4)	92.7 (85.9–99.6)	87.4 (80.4–94.3)	0.955 (0.900–1.000)
	Linear	89.7 (78.6–100.0)	93.1 (86.6–99.6)	86.7 (74.5–98.8)	94.7 (88.9–100.0)	92.0 (86.2–97.7)	0.980 (0.943–1.000)
	Sigmoid	82.8 (69.0–96.5)	94.8 (89.1–100.0)	88.9 (77.0–100.0)	91.7 (84.7–98.7)	90.8 (84.7–96.9)	0.943 (0.882–1.000)
Validation (6–12 mo) N = 102	Gaussian	91.2 (81.6–100.0)	100.0 (100.0–100.0)	100.0 (100.0–100.0)	95.8 (91.1–100.0)	97.1 (93.8–100.0)	0.996 (0.980–1.000)
	Linear	88.2 (77.4–99.1)	100.0 (100.0–100.0)	100.0 (100.0–100.0)	94.4 (89.2–99.7)	96.1 (92.3–99.8)	0.996 (0.980–1.000)
	Sigmoid	91.2 (81.6–100.0)	98.5 (95.7–100.0)	96.9 (90.8–100.0)	95.7 (91.0–100.0)	96.1 (92.3–99.8)	0.996 (0.980–1.000)
Validation (12–18 mo) N = 81	Gaussian	92.6 (82.7–100.0)	96.3 (91.3–100.0)	92.6 (82.7–100.0)	96.3 (91.3–100.0)	95.1 (90.3–99.8)	0.967 (0.919–1.000)
	Linear	81.5 (66.8–96.1)	94.4 (88.3–100.0)	88.0 (75.3–100.0)	91.1 (83.6–98.5)	90.1 (83.6–96.6)	0.973 (0.928–1.000)
	Sigmoid	88.9 (77.0–100.0)	96.3 (91.3–100.0)	92.3 (82.1–100.0)	94.5 (88.5–100.0)	93.8 (88.6–99.1)	0.976 (0.935–1.000)
Validation (18–24 mo) N = 90	Gaussian	83.3 (70.0–96.7)	98.3 (95.1–100.0)	96.2 (88.8–100.0)	92.2 (85.6–98.8)	93.3 (88.2–98.5)	0.981 (0.946–1.000)
	Linear	90.0 (79.3–100.0)	96.7 (92.1–100.0)	93.1 (83.9–100.0)	95.1 (89.7–100.0)	94.4 (89.7–99.2)	0.974 (0.934–1.000)
	Sigmoid	83.3 (70.0–96.7)	98.3 (95.1–100.0)	96.2 (88.8–100.0)	92.2 (85.6–98.8)	93.3 (88.2–98.5)	0.979 (0.943–1.000)
Validation (24–36 mo) N = 57	Gaussian	94.7 (84.7–100.0)	94.7 (87.6–100.0)	90.0 (76.9–100.0)	97.3 (92.1–100.0)	94.7 (88.9–100.0)	0.990 (0.959–1.000)
	Linear	94.7 (84.7–100.0)	97.4 (92.3–100.0)	94.7 (84.7–100.0)	97.4 (92.3–100.0)	96.5 (91.7–100.0)	0.997 (0.980–1.000)
	Sigmoid	94.7 (84.7–100.0)	100 (100.0–100.0)	100.0 (100.0–100.0)	97.4 (92.5–100.0)	98.2 (94.8–100.0)	0.997 (0.980–1.000)
Training: DS1 (3 mo–3 yr) N = 414	Gaussian	96.4 (93.3–99.5)	98.9 (97.7–100.0)	97.8 (95.3–100.0)	98.2 (96.6–99.8)	98.1 (96.7–99.4)	0.996 (0.988–1.000)
	Linear	96.4 (93.3–99.5)	98.2 (96.6–99.8)	96.4 (93.3–99.5)	98.2 (96.6–99.8)	97.6 (96.1–99.1)	0.996 (0.988–1.000)
	Sigmoid	94.9 (91.3–98.6)	98.2 (96.6–99.8)	96.3 (93.2–99.5)	97.5 (95.6–99.3)	97.1 (95.5–98.7)	0.995 (0.988–1.000)
Validation: DS2 (3 mo–3 yr) N = 417	Gaussian	89.9 (84.9–94.9)	97.5 (95.6–99.3)	94.7 (90.9–98.5)	95.1 (92.6–97.6)	95.0 (92.9–97.1)	0.988 (0.975–1.000)
	Linear	91.4 (86.7–96.0)	97.5 (95.6–99.3)	94.8 (91.0–98.5)	95.8 (93.4–98.1)	95.4 (93.4–97.4)	0.987 (0.973–1.000)
	Sigmoid	89.2 (84.1–94.4)	97.8 (96.1–99.5)	95.4 (91.8–99.0)	94.8 (92.2–97.3)	95.0 (92.9–97.1)	0.986 (0.972–1.000)
Validation (3–6 mo) N = 87	Gaussian	93.1 (83.9–100.0)	94.8 (89.1–100.0)	90.0 (79.3–100.0)	96.5 (91.7–100.0)	94.3 (89.4–99.1)	0.974 (0.932–1.000)
	Linear	93.1 (83.9–100.0)	93.1 (86.6–99.6)	87.1 (75.3–98.9)	96.4 (91.6–100.0)	93.1 (87.8–98.4)	0.976 (0.935–1.000)
	Sigmoid	93.1 (83.9–100.0)	94.8 (89.1–100.0)	90.0 (79.3–100.0)	96.5 (91.7–100.0)	94.3 (89.4–99.1)	0.973 (0.930–1.000)
Validation (6–12 mo) N = 102	Gaussian	94.1 (86.2–100.0)	97.1 (93.0–100.0)	94.1 (86.2–100.0)	97.1 (93.0–100.0)	96.1 (92.3–99.8)	0.994 (0.974–1.000)
	Linear	97.1 (91.4–100.0)	98.5 (95.7–100.0)	97.1 (91.4–100.0)	98.5 (95.7–100.0)	98.0 (95.3–100.0)	0.993 (0.972–1.000)
	Sigmoid	97.1 (91.4–100.0)	98.5 (95.7–100.0)	97.1 (91.4–100.0)	98.5 (95.7–100.0)	98.0 (95.3–100.0)	0.992 (0.971–1.000)
Validation (12–18 mo) N = 81	Gaussian	85.2 (71.8–98.6)	98.1 (94.6–100.0)	95.8 (87.8–100.0)	93.0 (86.4–99.6)	93.8 (88.6–99.1)	0.979 (0.941–1.000)
	Linear	81.5 (66.8–96.1)	98.1 (94.6–100.0)	95.7 (87.3–100.0)	91.4 (84.2–98.6)	92.6 (86.9–98.3)	0.980 (0.942–1.000)
	Sigmoid	81.5 (66.8–96.1)	98.1 (94.6–100.0)	95.7 (87.3–100.0)	91.4 (84.2–98.6)	92.6 (86.9–98.3)	0.979 (0.941–1.000)
Validation (18–24 mo) N = 90	Gaussian	83.3 (70.0–96.7)	98.3 (95.1–100.0)	96.2 (88.8–100.0)	92.2 (85.6–98.8)	93.3 (88.2–98.5)	0.992 (0.968–1.000)
	Linear	90.0 (79.3–100.0)	98.3 (95.1–100.0)	96.4 (89.6–100.0)	95.2 (89.8–100.0)	95.6 (91.3–99.8)	0.988 (0.961–1.000)
	Sigmoid	80.0 (65.7–94.3)	98.3 (95.1–100.0)	96.0 (88.3–100.0)	90.8 (83.7–97.8)	92.2 (86.7–97.8)	0.989 (0.963–1.000)
Validation (24–36 mo) N = 57	Gaussian	94.7 (84.7–100.0)	100.0 (100.0–100.0)	100.0 (100.0–100.0)	97.4 (92.5–100.0)	98.2 (94.8–100.0)	0.999 (0.987–1.000)
	Linear	94.7 (84.7–100.0)	100.0 (100.0–100.0)	100.0 (100.0–100.0)	97.4 (92.5–100.0)	98.2 (94.8–100.0)	0.999 (0.987–1.000)
	Sigmoid	94.7 (84.7–100.0)	100.0 (100.0–100.0)	100.0 (100.0–100.0)	97.4 (92.5–100.0)	98.2 (94.8–100.0)	0.999 (0.987–1.000)

Percent and 95% CI.

AUC, area under the curve; CI, confidence interval; NCA, neighborhood component analysis; NPV, negative predictive value; PCA, principal component analysis; PPV, positive predictive value; SVM, support vector machine.

those of NCA for all measures, regardless of kernel functions applied.

Performance based on 5 individual subtype windows. The performance within each of the time windows remained high for both

NCA and PCA methods. For the PCA method, the lowest sensitivity, specificity, PPV, NPV, accuracy, and AUC were 82%, 93%, 87%, 91%, 93%, and 0.97, respectively. QIFs within 2–3 years before t_0 also had very high predictive power (accuracy 95%–98%; AUC 0.99–1.00).

Table 3. A blind^a manual review of prediagnostic CT images of patients without chronic pancreatitis in 24–36 months before PDAC diagnosis and the CT images of the matched controls by case-control status and computer labels (high risk vs low risk)^b

Manually assigned ^c risk of PDAC	Cases (n = 19)			Controls (n = 19)		
	Labeled by computer as having high risk (n = 18)	Labeled by computer as having low risk (n = 1)	Total	Labeled by computer as having high risk (n = 0)	Labeled by computer as having low risk (n = 19)	Total
Low	14	0	14	0	16	16
Median	3	1	4	0	2	2
High	1	0	1	0	0	0
Nondiagnostic	0	0	0	0	1	1

CT, computed tomography; PDAC, pancreatic ductal adenocarcinoma.

^aReviewer was not informed about the computer labels or PDAC case/control status at the time of review.

^bComputer label was assigned according to the prediction algorithm developed with Gaussian kernel function and principal component analysis.

^cPatients were classified as having low risk if they only had diffuse atrophy, smaller cyst, simple cyst, loss of normal lobulation, or mild diffuse duct dilatation/prominence or did not have any obvious morphological features. Patients with complex cyst, cyst larger than 2 cm, or loss of normal lobulation of contour or having 2 or more low-risk features were considered having medium risk. Those with solid mass, focal abnormal enhancement, focal duct stricture, or focal/segmental atrophy were deemed high risk.

Patients with CP. For NCA, 2 QIFs were selected (see Supplementary Document 8, Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>). For PCA, 12 principal components were formed (see Supplementary Document 9, Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>). Sensitivity, specificity, PPV, NPV, and accuracy measures were all 100% (95% confidence interval [CI] 100%–100%) and AUC 1.00 (95% CI 1.00–1.00) for both NCA and PCA, regardless of kernel functions used to develop the prediction algorithms.

Sensitivity analysis

When the developed algorithms were applied to patients with stage I cancer in the validation data set (DS2), the sensitivity (0.89, 95% CI 0.74–1.00), specificity (0.97, 0.92–1.00), PPV (0.94, 0.83–1.00), NPV (0.95, 0.87–1.00), accuracy (0.94, 0.88–1.00), and AUC (0.98, 0.92–1.00 for Gaussian and linear kernel functions and 0.97, 0.91–1.00 for sigmoid kernel function) were all comparable with the performance measures reported in the previous section for all patients in the validation data sets (Table 2b). As expected, the 95% confidence intervals of the performance measures were larger for patients with stage I cancer compared with those reported in Table 2b (for all patients) because of the smaller sample size of the sensitivity analysis.

Exploratory analyses

Six and 14 QIFs were selected from EDS1 and EDS2, respectively (see Supplementary Document 6, Tables S1 and S3, Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>). Five of the 6 QIFs selected from EDS1 were in the list of QIFs selected from EDS2. The performance measures between the 2 sets of selected QIFs were similar (see Supplementary Document 6, Tables S2 and S4, Supplementary Digital Content 1, <http://links.lww.com/CTG/A892>). The performance measures between the 2 sets of data were comparable.

Clinical evaluation

In the 19 non-CP PDAC cases with prediagnostic images in 24–36 months, 14 patients, 4 patients, and 1 patient were classified as having low, medium, and high risk based on the manual review of the CT images, respectively (Table 3). Similarly, in the 19 matched controls, an overwhelming majority were believed to have low risk

based on the manual review of the CT images (Table 3). However, the computer algorithms only mistakenly predicted 1 patient in cases and correctly classified all the controls.

DISCUSSION

In this study, we developed and validated automated computer algorithms to predict PDAC solely based on QIFs of prediagnostic CT scans for patients with CP and non-CP patients separately. Our results showed that QIFs of prediagnostic scans can accurately predict PDAC in 3–36 months before diagnosis. Interestingly, when the validation was stratified by timing of scan in relation to diagnosis, the performance seemed to be maintained in all the periods examined before cancer diagnosis. The QIF-based algorithm had excellent ability to predict PDAC even for scans from 24 to 36 months before cancer diagnosis. A manual assessment of the scans in 24–36 months before PDAC diagnosis revealed that an overwhelming majority had low risk of PDAC, yet the computer algorithm was able to predict the outcome correctly, except for 1 case who had mild ductal prominence and loss of lobulation in the pancreas head.

A major barrier for early detection in pancreatic cancer is the inability to reliably identify precursor lesions based on conventional imaging. Pancreatic intraepithelial neoplasia (PanIN) III or high-grade dysplasia are histologic findings not typically visible on cross-sectional imaging (41). Previous studies have identified specific abnormalities that can be identified in up to 50% of cancer cases before diagnosis of PDAC including main duct dilatation, atrophy, and duct stricture (4,42,43). However, such findings lack sensitivity and in many cases, are subject to interpreter variability. Therefore, a systematic approach of direct imaging analysis that applies objective assessment of imaging-based parameters provides a promising opportunity to identify imaging-based signatures of early pancreatic cancer. The performance of the algorithms was high in all the time windows being studied.

Radiomics has been applied in the diagnosis, prediction, and prognosis of other cancer types (e.g., lung cancer, lung nodule and breast cancer) (44). While radiomics has been applied for prognosis in PDAC (7–12), very little progress has been made in early detection. Using 225 training and 125 validation images, Chu et al. differentiated CT scans of patients with PDAC and healthy controls with a sensitivity of 100% and specificity of 98.5% (18). However, in the study by Chu, the images being analyzed were

obtained after diagnosis (mean tumor size 4.1 cm). In addition, the pancreas boundaries were manually segmented (18), and thus, utility of the algorithm in clinical operation is limited. In the current study, prediagnostic scans were applied, and segmentation of the pancreas was computerized.

A key distinction related to the assessment of prediagnostic imaging is whether the automated QIF algorithms were classifying cancer risk based on features readily visible on the images or incorporating additional elements from the data contained within the images. What we observed in this study is that relying on human assessment of the prediagnostic CT images obtained 2–3 years before cancer diagnosis by an expert body-image radiologist with 15 years of experience failed to distinguish patients with cancer and healthy patients at the level of performance obtained by the radiomics-based algorithm, suggesting that the QIF provide additional information beyond identification of established radiographic findings.

This study has several strengths. First, algorithms were developed separately for patients with and without CP to allow for more accurate prediction, given the expected morphological differences between patients with CP and non-CP patients. Second, we implemented a pancreas segmentation algorithm and series of quantitative imaging features that were previously validated. Therefore, the developed classifiers can potentially be widely implemented across healthcare systems. Third, we relied on prediagnostic images 3 months–3 years before cancer diagnosis. Thus, the developed algorithms have the potential to predict pancreatic cancer in a time frame that would allow for sufficient lead time for intervention to affect the disease course. Fourth, we applied 2 machine learning approaches to select the most influential radiomic features. Although both methods worked well, it seems that algorithms based on principal components formed by PCA are more accurate. This is likely to be the result of the more inclusive nature of PCA compared with NCA. Finally, the comparison between blinded expert human review and automated QIF-based algorithm helped demonstrate the potential added value of this approach beyond identification of established abnormalities of the pancreatic parenchymal or duct system.

This study had several limitations. First, although the images being used in the analyses were at least 3 months before cancer diagnosis, a small number of scans may contain visible tumors, which may deform the shape of the pancreas and thus negatively affect the performance of the final algorithms. A few (<5) images were manually removed because of visible tumors. However, additional images with visible tumors may still have been present. Second, the number of CP cases is much smaller compared with that of non-CP cases. The high performance in this group of patients could be an overestimation of the true performance because of the small sample size. However, our approach demonstrated the possibility of successful classification in this group of patients. Validation of the developed algorithms in a larger group of patients with CP is warranted in the future. Third, we did not consider higher order statistics QIF (e.g., Gabor wavelet transformation) in this study. Previously, studies have shown that mapping QIFs into a higher order feature space can further improve accuracy (13,15). Fourth, cases and controls were not matched by the indications for the scans. Fifth, the manual assessment for PDAC risk was only performed for scans in 2–3 years before the index date and inter-rater reliability was not assessed. Finally, this study lacks an external validation, in which CT images from another health care may be used to test the robustness of our algorithms.

In summary, the radiomics-based automated algorithms provide a method to detect PDAC as early as 2–3 years before cancer

diagnosis. The algorithm has the potential to be used for future early detection protocols for pancreatic cancer. A prospective validation study at multiple institutions in the future could provide evidence on generalizability. Future studies could expand these analyses by developing and evaluating radiomics-based algorithms to detect early signals by localization and histopathology of pancreatic cancer (45). Future research is required to understand feasibility, challenges, and cost-effectiveness of such an implementation.

CONFLICTS OF INTEREST

Guarantor of the article: Wansu Chen, PhD.

Specific author contributions: W.C.: conceptualization, methodology, software, validation, investigation, resources, writing—original draft, writing—review and editing, visualization, and supervision. Y.Z.: methodology, software, validation, formal analysis, investigation, data curation, writing—review and editing, and visualization. V.A.: methodology, software, formal analysis, investigation, data curation, and writing—review and editing. R.A.P.: conceptualization, methodology, validation, investigation, data curation, and writing—review and editing. E.L.: conceptualization, validation, investigation, writing—review and editing, and supervision. E.P.: validation, investigation, and writing—review and editing. B.U.W.: conceptualization, methodology, validation, resources, writing—review and editing, supervision, and funding acquisition.

Financial support: Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Number R01CA230442. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Potential competing interests: None to report.

ACKNOWLEDGEMENTS

The authors thank Sole Cardoso for the assistance with formatting the manuscript and Botao Zhou for the additional analyses. The images were kindly provided by the Medical Imaging Technology and Informatics group of Kaiser Permanente Southern California Permanente Medical Group.

Study Highlights

WHAT IS KNOWN

- ✓ Pancreatic cancer is the third leading cause of cancer deaths among men and women in the United States.
- ✓ Early detection of pancreatic ductal adenocarcinoma (PDAC) is difficult owing to lack of specific symptoms or established screening.

WHAT IS NEW HERE

- ✓ Quantitative imaging features (QIFs) of prediagnostic computed tomography (CT) scans can accurately predict PDAC in 3–36 months before diagnosis (accuracy 94%–95% and area under the curve [AUC] 0.98–0.99 for patients without chronic pancreatitis and accuracy 100% and AUC 1.00 for patients with chronic pancreatitis).
- ✓ QIFs on CT examinations within 2–3 years before cancer diagnosis also had very high predictive accuracy (accuracy 95%–98%; AUC 0.99–1.00).
- ✓ The QIF-based algorithm outperformed manual rereview of images for the determination of PDAC risk.

REFERENCES

- National Institutes of Health. NIH Surveillance, Epidemiology, and End Results Program. US Department of Health and Human Services; National Institutes of Health; National Cancer Institute: Bethesda, MD, 2021.
- National Institutes of Health. Scientific Framework for Pancreatic Ductal Adenocarcinoma (PDAC). National Cancer Institute: Bethesda, MD, 2014, pp 1–29.
- Tanaka S, Nakao M, Ioka T, et al. Slight dilatation of the main pancreatic duct and presence of pancreatic cysts as predictive signs of pancreatic cancer: A prospective study. *Radiology* 2010;254(3):965–72.
- Gangi S, Fletcher JG, Nathan MA, et al. Time interval between abnormalities seen on CT and the clinical diagnosis of pancreatic cancer: Retrospective review of CT scans obtained before diagnosis. *AJR Am J Roentgenol* 2004;182(4):897–903.
- Wu BU, Sampath K, Berberian CE, et al. Prediction of malignancy in cystic neoplasms of the pancreas: A population-based cohort study. *Am J Gastroenterol* 2014;109(1):121–9; quiz 130.
- Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology* 2016;278(2):563–77.
- Eilaghi A, Baig S, Zhang Y, et al. CT texture features are associated with overall survival in pancreatic ductal adenocarcinoma - a quantitative analysis. *BMC Med Imaging* 2017;17(1):38.
- 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.
- Attiyeh MA, Chakraborty J, Doussot A, et al. Survival prediction in pancreatic ductal adenocarcinoma by quantitative computed tomography image analysis. *Ann Surg Oncol* 2018;25(4):1034–42.
- Chakraborty J, Langdon-Embry L, Cunan KM, et al. Preliminary study of tumor heterogeneity in imaging predicts two year survival in pancreatic cancer patients. *PLoS One* 2017;12(12):e0188022.
- Yun G, Kim YH, Lee YJ, et al. Tumor heterogeneity of pancreas head cancer assessed by CT texture analysis: Association with survival outcomes after curative resection. *Sci Rep* 2018;8(1):7226.
- Sandrasegaran K, Lin Y, Asare-Sawiri M, et al. CT texture analysis of pancreatic cancer. *Eur Radiol* 2019;29(3):1067–73.
- Zhang Y, Lobo-Mueller EM, Karanicolas P, et al. Improving prognostic performance in resectable pancreatic ductal adenocarcinoma using radiomics and deep learning features fusion in CT images. *Sci Rep* 2021; 11(1):1378.
- Zhang Y, Lobo-Mueller EM, Karanicolas P, et al. CNN-based survival model for pancreatic ductal adenocarcinoma in medical imaging. *BMC Med Imaging* 2020;20(1):11.
- Khalvati F, Zhang Y, Baig S, et al. Prognostic value of CT radiomic features in resectable pancreatic ductal adenocarcinoma. *Sci Rep* 2019;9(1):5449.
- Mashayekhi R, Parekh VS, Faghih M, et al. Radiomic features of the pancreas on CT imaging accurately differentiate functional abdominal pain, recurrent acute pancreatitis, and chronic pancreatitis. *Eur J Radiol* 2020;123:108778.
- Park S, Chu LC, Hruban RH, et al. Differentiating autoimmune pancreatitis from pancreatic ductal adenocarcinoma with CT radiomics features. *Diagn Interv Imaging* 2020;101(9):555–64.
- Chu LC, Park S, Kawamoto S, et al. Utility of CT radiomics features in differentiation of pancreatic ductal adenocarcinoma from normal pancreatic tissue. *AJR Am J Roentgenol* 2019;213(2):349–57.
- Abunahel BM, Pontre B, Kumar H, et al. Pancreas image mining: A systematic review of radiomics. *Eur Radiol* 2021;31(5):3447–67.
- Kirkegård J, Mortensen FV, Cronin-Fenton D. Chronic pancreatitis and pancreatic cancer risk: A systematic review and meta-analysis. *Am J Gastroenterol* 2017;112(9):1366–72.
- Chen W, Chen Q, Parker RA, et al. Risk prediction of pancreatic cancer in patients with abnormal morphologic findings related to chronic pancreatitis: A machine learning approach. *Gastro Hep Adv* 2022;1(6): 1014–26.
- Koebnick C, Langer-Gould AM, Gould MK, et al. Sociodemographic characteristics of members of a large, integrated health care system: Comparison with US Census Bureau data. *Perm J* 2012;16(3):37–41.
- Asadpour V, Parker RA, Mayock RP, et al. Pancreatic cancer tumor analysis in CT images using patch-based multi-resolution convolutional neural network. *Biomed Signal Process Control* 2021;68:102652.
- 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(2): 328–38.
- MathWorks. MATLAB parallel computing toolbox software. MathWorks: Natick, MA, 2020.
- Huo Y, Tang Y, Chen Y, et al. Stochastic tissue window normalization of deep learning on computed tomography. *J Med Imaging (Bellingham)* 2019;6(4):044005.
- Goldberger J, Roweis S, Hinton G, et al. Neighbourhood components analysis. Department of Computer Science, University of Toronto: Toronto, Canada, 2004.
- Yang W, Wang K, Zuo W. Neighborhood component feature selection for high-dimensional data. *J Comput* 2012;7(1):161–8.
- Rao CR. The use and interpretation of principal component analysis in applied research. *Indian J Stat Ser A* 1964;26(4):329–58.
- MathWorks. Statistics and Machine Learning Toolbox. MathWorks: Natick, MA, 2021. (<https://www.mathworks.com/help/stats/neighborhood-component-analysis.html>). Accessed December 5, 2022.
- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria, 2017. (<https://www.R-project.org/>). Accessed August 11, 2021.
- Stanfill B, Reehl S, Bramer L, et al. Extending classification algorithms to case-control studies. *Biomed Eng Comput Biol* 2019;10: 1179597219858954.
- Gunn SR. Support vector machines for classification and regression. 1998.
- Souza CR. Kernel functions for machine learning applications. 2010. (<http://crsouza.com/2010/03/17/kernel-functions-for-machine-learning-applications/>). Accessed June 9, 2021.
- Hofmann T, Scholkopf B, Smola AJ. Kernel methods in machine learning. *Ann Stat* 2008;36(3):1171–220.
- Bischi B, Lang M, Kotthoff L, et al. ROC analysis and performance curves. (https://mlr.mlr-org.com/articles/tutorial/roc_analysis.html). Accessed September 7, 2022.
- Confidence intervals for the area under an ROC curve. In: PASS sample size software. (https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/PASS/Confidence_Intervals_for_the_Area_Under_an_ROC_Curve.pdf). Accessed September 7, 2022.
- SAS version 9.4 for unix. SAS Institute: Cary, NC, 2021.
- MATLAB and statistics toolbox release 2017b. MathWorks: Natick, MA, 2017.
- R Core Team. R software: Version 3.6.2. R Foundation for Statistical Computing: Vienna, Austria, 2021.
- Basturk O, Hong SM, Wood LD, et al. A revised classification system and recommendations from the baltimore consensus meeting for neoplastic precursor lesions in the pancreas. *Am J Surg Pathol* 2015; 39(12):1730–41.
- Singh DP, Sheedy S, Goenka AH, et al. Computerized tomography scan in pre-diagnostic pancreatic ductal adenocarcinoma: Stages of progression and potential benefits of early intervention: A retrospective study. *Pancreatol* 2020;20(7):1495–501.
- Chen W, Butler RK, Zhou Y, et al. Prediction of pancreatic cancer based on imaging features in patients with duct abnormalities. *Pancreas* 2020; 49(3):413–9.
- Gillies RJ, Schabath MB. Radiomics improves cancer screening and early detection. *Cancer Epidemiol Biomarkers Prev* 2020;29(12):2556–67.
- Haeberle L, Esposito I. Pathology of pancreatic cancer. *Transl Gastroenterol Hepatol* 2019;4:50.

Open Access This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.