UCLA
UCLA Previously Published Works

Title

Identification of Individuals at Increased Risk for Pancreatic Cancer in a Community-Based Cohort of Patients With Suspected Chronic Pancreatitis.

Permalink

Journal

Authors

Jeon, Christie
Chen, Qiaoling
Yu, Wei
et al.

Publication Date

DOI

Identification of Individuals at Increased Risk for Pancreatic Cancer in a Community-Based Cohort of Patients With Suspected Chronic Pancreatitis

Christie Y. Jeon, ScD[1,2], Qiaoling Chen, MS[3], Wei Yu, MS[3], Elizabeth Y. Dong, MD[4], Joanie Chung, MPH[3], Stephen J. Pandol, MD[1,5], Dhiraj Yadav, MD, MPH[6], Darwin L. Conwell, MD, MSc[7] and Bechien U. Wu, MD, MPH[4]

OBJECTIVES:	**We lack reliable methods for identifying patients with chronic pancreatitis (CP) at increased risk for pancreatic cancer. We aimed to identify radiographic parameters associated with pancreatic cancer in this population.**
METHODS:	**We conducted a retrospective cohort study of patients with suspected CP within an integrated healthcare system in Southern California in 2006–2015. Patients were identified by a diagnostic code and confirmed by imaging findings (parenchymal calcification, ductal stones, glandular atrophy, pseudocyst, main duct dilatation, duct irregularity, abnormal side branch, or stricture) defined by the natural language processing of radiographic reports. We used Cox regression to determine the relationship of smoking, alcohol use, acute pancreatitis, diabetes, body mass index, and imaging features with the risk of incident pancreatic cancer at least 1 year after abnormal pancreas imaging.**
RESULTS:	**We identified 1,766 patients with a diagnostic code and an imaging feature for CP with a median follow-up of 4.5 years. There were 46 incident pancreatic cancer cases. Factors that predicted incident pancreatic cancer after 1-year of follow-up included obesity (hazard ratio 2.7, 95% confidence interval: 1.2–6.1) and duct dilatation (hazard ratio 10.5, 95% confidence limit: 4.0–27). Five-year incidence of pancreatic cancer in this population with duct dilatation was 6.3%.**
DISCUSSION:	**High incidence of pancreatic cancer in suspected patients with CP with pancreatic duct dilatation warrants regular surveillance for pancreatic cancer.**

INTRODUCTION

Chronic pancreatitis (CP) is an established risk factor for pancreatic cancer (1–3), a leading cause of cancer-related mortality among patients with CP (4). However, routine screening for pancreatic cancer in this population is not recommended based on the relatively low absolute risk of cancer and potential harms associated with screening (5). Therefore, improved methods for early cancer detection in patients with CP are a key step to improve the outcome in this patient population.

Retrospective analyses of the real-world clinical data have identified an excess risk of pancreatic cancer soon after a diagnosis of CP (6), raising concern that early cancers may be misdiagnosed as CP. Apart from the risk of missing an occult neoplasm, previous literature has also consistently identified higher long-term risk of pancreatic cancer among patients with CP than in the general population (1–3). To date, it remains unclear which patients with CP are at increased risk for developing pancreatic cancer in the future. Addressing this gap in knowledge is an important step to develop effective strategies for the early detection of cancer in these patients.

The objective of this study was to assess clinical and radiographic features associated with increased risk of pancreatic cancer after imaging for CP diagnosis. Considering that the cost effectiveness of future diagnostic or screening recommendations depends on the probability of discovering pancreatic cancer, we also aimed to estimate the age- and sex-specific incidence rates of

[1]Samuel Oschin Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA; [2]Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, California, USA; [3]Research and Evaluation, Kaiser Permanente Southern California, Pasadena, California, USA; [4]Center for Pancreatic Care, Kaiser Permanente Los Angeles Medical Center, Los Angeles, California, USA; [5]Division of Digestive and Liver Diseases, Cedars-Sinai Medical Center, Los Angeles, California, USA; [6]Division of Gastroenterology, Hepatology and Nutrition, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA; [7]Division of Gastroenterology, Ohio State University Wexner Medical Center, Columbus, Ohio, USA. **Correspondence:** Christie Y. Jeon, ScD. E-mail: christie.jeon@cshs.org.
Received March 20, 2019; accepted February 3, 2020; published online April 4, 2020

pancreatic cancer in the CP population and the relative incidence as compared to the general population.

METHODS

Study design and setting

We conducted a retrospective cohort study of patients diagnosed with CP within Kaiser Permanente Southern California (KPSC) between January 2006 and December 2015. Kaiser Permanente is a community-based integrated healthcare system comprising 7 distinct regions, of which Southern California is one of the largest. The present study was approved by the Institutional Review Board of KPSC and the Consortium for the Study of Chronic Pancreatitis, Diabetes, and Pancreatic Cancer.

Patient population

Patients with CP were identified through diagnosis codes (International Classification of Disease [ICD]-9 577.1) with additional confirmation by the presence of at least 1 radiographic feature suggestive of CP on computed tomography, transabdominal ultrasound, or magnetic resonance imaging (MRI). When patients had multiple imaging procedures performed, we selected the earliest image demonstrating an abnormal finding as the date of cohort entry. Among images conducted on patients with suspected CP, the most common radiographic studies were computed tomography (52%), followed by transabdominal ultrasound (23%) and MRI (9%). Patients with a history of pancreatic cancer before the first radiographic image with a positive feature for CP, those with imaging performed before 2006, those with less than 1 year of continuous membership before the imaging procedure, and those with less than 1 year of follow-up after an abnormal pancreas radiograph were excluded. Occult, prevalent pancreatic cancer cases were excluded, along with persons with follow-up less than 1 year.

CP imaging feature set (natural language processing)

One of the key challenges in characterizing imaging features among patients with CP is the lack of structured codes or established template for defining specific radiographic features. To address this limitation, we developed a natural language processing (NLP) algorithm to identify the imaging features reported on the free text of findings and impressions focused on the pancreas in cross-sectional imaging studies. We developed search strings for the following radiographic features of CP based on the established parameters from clinical guidelines (7): atrophy, calcification, pseudocyst, main duct dilatation, pancreatic duct irregularity, intraductal stone/calculus, intraductal filling defect, abnormal side-branch, and pancreatic duct stricture.

We used an iterative process for each of the factors combining chart validation with algorithm refinement until >98% accuracy was achieved compared with a manual chart review for 300 patients with CP. An in-house NLP software platform was used to develop a series of imaging algorithms. The software package was developed based on an open-source Application Programming Interface including Natural Language Toolkit, NegEx/pyConText, and Stanford CoreNLP (8–10). The NLP algorithm for the CP image was developed in sequential steps. First, the imaging report was preprocessed for section and sentence boundary detection. Second, we used text patterns to map and classify analogous terms. Third, the distance between key words, such as "pancreatic duct" and "irregular," were measured and algorithms were developed to apply maximum distance between words. Fourth, the absence of features was detected through the

negation algorithm pyContText/NegEx. Fifth, features described outside of certain temporal boundaries (e.g., historical finding) were detected by pyConText and excluded. Finally, the NLP note-level results were combined into patient-level results.

Validation of CP

Relying on the ICD codes alone to identify CP has suboptimal positive predictive value (11), we therefore required an abnormal pancreas imaging result in addition to ICD codes to identify patients with CP. Although persons with calcification or calculi most definitively have CP, other features are less definitive for CP. We therefore performed a manual review of electronic health records and imaging reports of 10% of patients with CP without calculi or calcification. The reviewer was blinded to the NLP-discovered features and the eventual pancreatic cancer status of the patient.

Outcome assessment

Pancreatic cancer cases included patients with at least 2 outpatient or inpatient visits with the diagnosis codes (ICD-9 157.*), patients registered in the internal KPSC cancer registry as having a malignant neoplasm in the pancreas and those with pancreatic cancer as cause of death in the Death Index. All cancer cases were confirmed through the manual chart review, and only primary pancreatic cancer cases were included. The date of diagnosis was defined as earliest of the following: date of the healthcare visit with pancreatic cancer–specific ICD-9 code, the date of pancreatic cancer diagnosis in the tumor registry, or the date of death due to pancreatic cancer. Time-to-event was defined as days from 12 months of the first positive pancreas imaging to pancreatic cancer diagnosis. Patients were censored by death due to causes other than pancreatic cancer, discontinuation of KPSC membership, or December 31, 2016, whichever came first.

Covariate definitions

Age was determined at the time of the first physician encounter with a diagnosis of CP. We also extracted data on sex and race, classified as non-Hispanic white, non-Hispanic black, Asian, Hispanic and others. Current consumption of alcohol (yes vs no) and smoking status (current, former, or never) were determined by self-report based on the most updated health status data available on or before cohort entry. Body mass index (BMI) was assessed at the time of the abnormal imaging result. History of acute pancreatitis and diabetes *before cohort entry* was assessed by ICD-9 codes 577.0, 250.xx, respectively.

Data analysis

We first constructed a cumulative incidence plot of pancreatic cancer using the Kaplan–Meier survival method. The demographic and clinical characteristics of the CP population were described by the pancreatic cancer status. We then used Cox regression with Firth penalized partial likelihood approach to evaluate the risk factors associated with the development of pancreatic cancer, given the limited number of events. Covariates in the model included patient-related factors (age, sex, race/ethnicity, alcohol consumption, smoking, diabetes, BMI, and previous history of acute pancreatitis) and radiographic features (atrophy, calcification, pseudocyst, main duct dilatation, pancreatic duct irregularity, intraductal stone/calculus, intraductal filling defect, abnormal side-branch, and pancreatic duct stricture). We additionally performed a competing risk analysis by Fine and Gray regression with death

due to causes other than pancreatic cancer as a competing risk to confirm that covariate-pancreatic cancer outcomes were not influenced by mortality unrelated to pancreatic cancer.

In addition, to estimate the age- and sex-specific incidence rates of pancreatic cancer, we determined the number of incident pancreatic cancer by age groups (≤40, 41–50, 51–60, 61–70, 71–80, and ≥81) and sex in the overall cohort and divided by the person-years observed. The age- and sex-specific incidence rates of pancreatic cancer in the non-CP population were also estimated for comparison and the age- and sex-standardized incidence rate ratios of pancreatic cancer were computed. Statistical analysis was performed with SAS statistical software, version 9.4 (Cary, NC), and all reported P values are 2-sided with alpha 0.05.

RESULTS

After applying our inclusion and exclusion criteria, we identified a total of 1,766 patients diagnosed with CP who had at least 1 abnormal pancreatic finding on radiographic imaging and who had survived at least 1 year without pancreatic cancer (Figure 1). Of 86 patients with coded diagnosis of CP without calcification or calculi at first positive imaging who were subject to the manual chart review, 36 (42%) had or developed definitive CP based on the endoscopic ultrasound findings or based on the Cambridge 3 or 4 classification on MRI/computed tomography imaging. Another 43 (50%) had or developed acute pancreatitis without definitive CP, of whom 20 (23%) were documented to have severe acute pancreatitis. Seven (8%) had no clear evidence of CP or acute pancreatitis, whereas one of these patients had a pancreatic pseudocyst and another had elevated CA19-9.

A descriptive summary of the study cohort is presented in Table 1. The median duration of follow-up after positive pancreas imaging was 4.5 years (interquartile range: 2.5, 7.0 years). The mean age at cohort entry was 60.1 years (SD 16.0), and 46% of the cohort were women. Thirty percent of the population reported regularly consuming alcohol at cohort entry. Twenty percent of the patients in the study cohort were current smokers, with an additional 33% having been a smoker in the past. Overall, 48% of patients had a history of acute pancreatitis and 44% of patients had a pre-existing diagnosis of diabetes before being diagnosed with CP. A few patients (n = 10, 0.6%) had a diagnosis of intraductal papillary mucinous neoplasm (IPMN) at the time of imaging procedure. The median time to incident pancreatic cancer was 2.4 years.

Of the CP imaging features, pseudocyst was the most common abnormality (58%) followed closely by parenchymal calcification (48%). Ductal dilation was present in 43% of the total study population and present in most patients who went on to be diagnosed pancreatic cancer (89%).

A total of 46 patients developed pancreatic cancer during the study period, 2 of whom had IPMN at baseline. An incidence plot for pancreatic cancer since 1 year after abnormal pancreas imaging in suspected patients with CP is presented in Figure 2. Incidence of pancreatic cancer increased steadily, and estimated cumulative incidence of pancreatic cancer was 2% after 2 years of follow-up and was 3.2% after 5 years of follow-up.

Table 2 presents the results of Cox proportional hazards regression for incident pancreatic cancer. Although age was associated with increased risk of pancreatic cancer, other patient-related factors such as pre-existing diabetes, smoking, or alcohol history were not associated with increased risk of pancreatic cancer after adjusting for findings on radiographic imaging. Being obese (hazard ratio [HR] 2.7, 95% confidence limit [CL]: 1.2–6.1) and duct dilatation (HR 10.5, 95% CL: 4.0–26.1) were independently associated with increased risk of pancreatic cancer. Of note, being underweight was also marginally associated with pancreatic cancer (HR 3.1, 95% CL: 0.9–10.8), demonstrating that the association of BMI with pancreatic cancer is not monotonic. A competing risk analysis by Fine and Gray regression confirmed that our findings on the factors associated with pancreatic cancer were not influenced by death due to causes other than pancreatic cancer.

Duct dilatation was a salient feature of CP that was associated with the future risk of pancreatic cancer. The 2- and 5-year risks of pancreatic cancer in patients with CP with duct dilatation who survived 1 year without pancreatic cancer were 3.7% and 6.3%, respectively. By contrast, the 2- and 5-year risks of pancreatic cancer in CP patients without duct dilation were significantly lower: 0.6% and 0.6%, respectively ($P < 0.001$) (Figure 3). Although incidence also varied by the presence of calcification (2-year risk of 1.7% and 5-year risk of 3.2% in persons with calcification and 2-year risk of 2.3% and 5-year risk of 3.5% in persons without calcification), the differences were not statistically significant ($P = 0.31$).

The incidence rates of pancreatic cancer per 100,000 person years by sex and age group are presented in Supplementary Digital Content 1 (see Table S1, http://links.lww.com/CTG/A222). The standardized rate ratio of pancreatic cancer is presented in Supplementary Digital Content 1 (see Table S2, http://links.lww.com/CTG/A222). The incidence rate of pancreatic cancer in patients with CP was 12 times higher (incidence rate ratio 12, 95% CL: 8.8–16) than in the referent population, without adjustment for BMI.
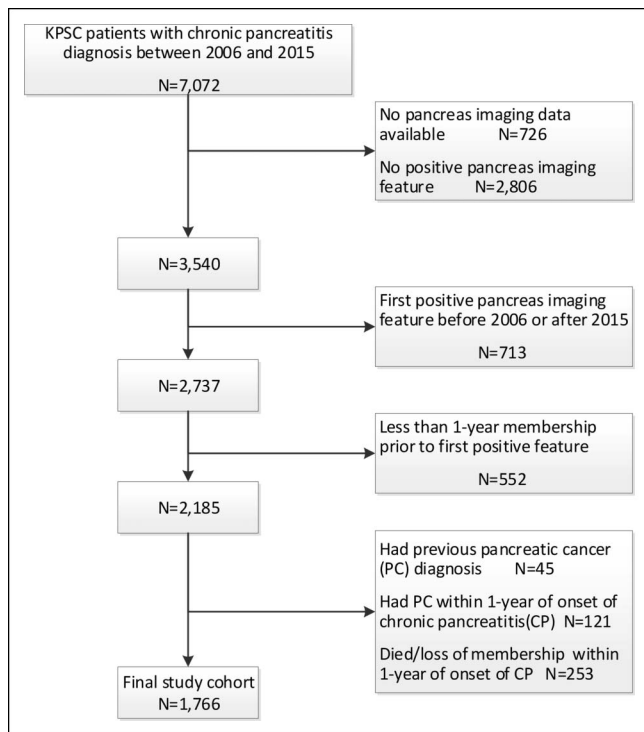
Figure 1. Selection of patients with suspected chronic pancreatitis with abnormal radiographic pancreas imaging. KPSC, Kaiser Permanente Southern California.

PANCREAS

Table 1. Baseline characteristics of Kaiser Permanente Southern California patients with suspected CP between 2006 and 2015

	No pancreatic cancer (N = 1,720)	Incident pancreatic cancer (N = 46)	Total (N = 1,766)
Patient baseline demographics			
Age at CP diagnosis, mean (SD)	60.0 (16.05)	64.5 (12.06)	60.1 (15.97)
≤40	197 (11.5%)	0 (0%)	197 (11.2%)
41–50	251 (14.6%)	7 (15.2%)	258 (14.6%)
51–60	402 (23.4%)	9 (19.6%)	411 (23.3%)
61–70	396 (23%)	16 (34.8%)	412 (23.3%)
71–80	302 (17.6%)	10 (21.7%)	312 (17.7%)
≥81	172 (10%)	4 (8.7%)	176 (10%)
Gender			
Female	787 (45.8%)	19 (41.3%)	806 (45.6%)
Male	933 (54.2%)	27 (58.7%)	960 (54.4%)
Race/ethnicity			
Asian	907 (52.7%)	27 (58.7%)	934 (52.9%)
Black	262 (15.2%)	10 (21.7%)	272 (15.4%)
Hispanic	420 (24.4%)	8 (17.4%)	428 (24.2%)
White	116 (6.7%)	1 (2.2%)	117 (6.6%)
Others/unknown	15 (0.9%)	0 (0%)	15 (0.8%)
Membership duration: prepositive imaging (yr)			
Mean (SD)	15.4 (12.41)	14.1 (11.44)	15.4 (12.38)
Median	12.1	11.5	12.0
Interquartile range	5.2–22.8	4.5–22.1	5.1–22.8
Range	1.0–111.1	1.5–46.2	1.0–111.1
Patient baseline clinical characteristics			
Alcohol			
No	901 (45.7%)	60 (43.8%)	967 (45.2%)
Yes	584 (29.6%)	37 (27%)	633 (29.6%)
Unknown	488 (24.7%)	40 (29.2%)	540 (25.2%)
Smoking status			
Non-smoker	624 (36.3%)	14 (30.4%)	638 (36.1%)
Former smoker	568 (33%)	16 (34.8%)	584 (33.1%)
Current smoker	342 (19.9%)	8 (17.4%)	350 (19.8%)
Unknown	186 (10.8%)	8 (17.4%)	194 (11%)
BMI			
Underweight (<18.5 kg/m^2)	82 (4.8%)	3 (6.5%)	85 (4.8%)
Normal (18.5–24.9 kg/m^2)	631 (36.7%)	11 (23.9%)	642 (36.3%)
Overweight (25–29.9 kg/m^2)	564 (32.8%)	15 (32.6%)	579 (32.8%)
Obese (≥30 kg/m^2)	443 (25.8%)	17 (37%)	460 (26.1%)
Acute pancreatitis	834 (48.5%)	18 (39.1%)	852 (48.2%)
Diabetes	748 (43.5%)	24 (52.2%)	772 (43.7%)
Radiographic reports: Positive imaging features			
Atrophy	596 (34.7%)	23 (50%)	619 (35.1%)

Table 1. (*continued*)

	No pancreatic cancer (N = 1,720)	Incident pancreatic cancer (N = 46)	Total (N = 1,766)
Calcification	832 (48.4%)	19 (41.3%)	851 (48.2%)
Pseudocyst	995 (57.8%)	29 (63%)	1,024 (58%)
Duct dilatation	722 (42%)	41 (89.1%)	763 (43.2%)
Ductal irregularity	123 (7.2%)	6 (13%)	129 (7.3%)
Intraductal calculi	184 (10.7%)	4 (8.7%)	188 (10.6%)
Intraductal filling defect	68 (4%)	2 (4.3%)	70 (4%)
Abnormal side-branch	113 (6.6%)	8 (17.4%)	121 (6.9%)
Stricture	89 (5.2%)	8 (17.4%)	97 (5.5%)
Follow-up			
Length of follow-up (yr)			
Mean (SD)	5.0 (2.75)	3.3 (2.35)	4.9 (2.75)
Median	4.6	2.4	4.5
Interquartile range	2.5–7.1	1.6–4.6	2.5–7.0
Range	1.0–10.9	1.0–9.9	1.0–10.9
Follow-up event			
Death	334 (19.4%)	8 (17.4%)	334 (18.9%)
Loss of membership	252 (14.7%)	0 (0%)	252 (14.3%)
End of study	1,134 (65.9%)	0 (0%)	1,134 (64.2%)
Pancreatic cancer	0 (0%)	46 (100%)	46 (2.6%)

BMI, body mass index; CP, chronic pancreatitis.

DISCUSSION

A major challenge in improving survival for patients that develop pancreatic cancer has been the lack of an effective strategy for early detection. In this community-based retrospective cohort study, patients with suspected CP were at substantially increased risk of developing pancreatic cancer compared with the age- and
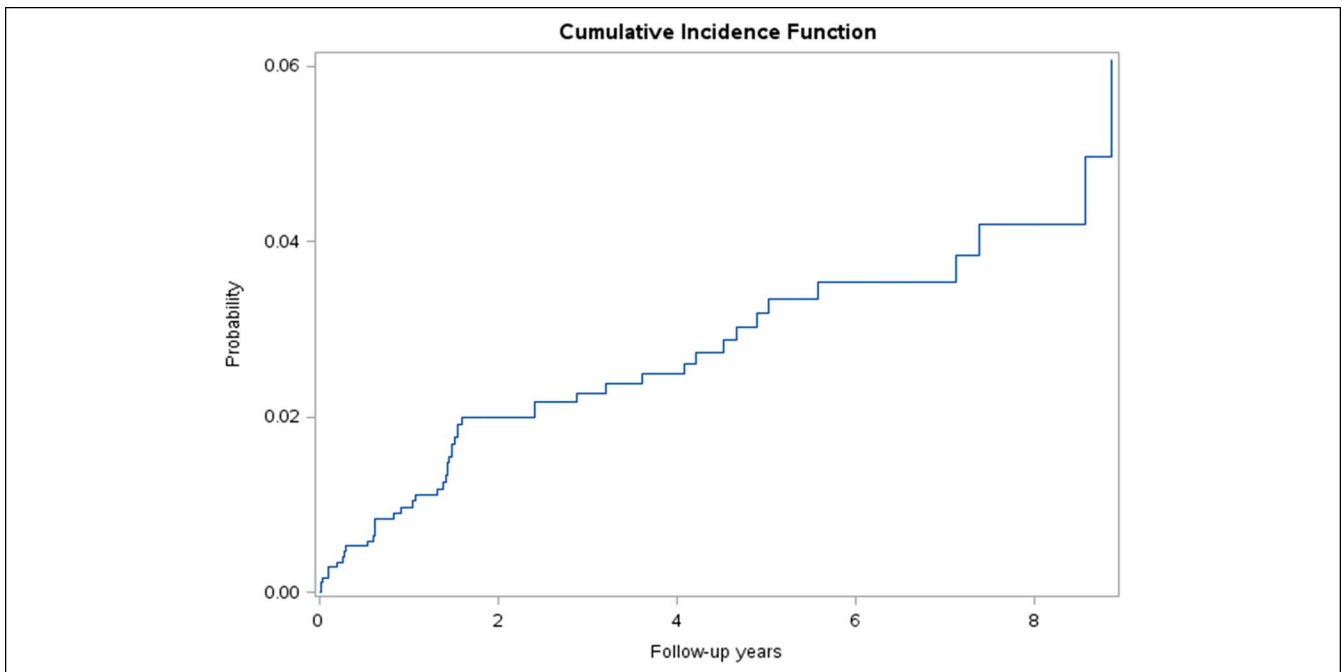
Figure 2. Cumulative incidence curve of pancreatic cancer ≥1 year after abnormal pancreas imaging in patients with suspected chronic pancreatitis.

PANCREAS

Table 2. Adjusted HR for incident pancreatic cancer ≥1 year after abnormal pancreas imaging in suspected patients with chronic pancreatitis

	HR	95% CLs	*P* value
Age	1.03	1.00–1.05	0.02
Gender (male vs female)	1.27	0.67–2.41	0.46
Race/ethnicity (ref = White)			0.85
Asian	0.57	0.10–3.16	0.52
Black	1.19	0.55–2.54	0.66
Hispanic	0.74	0.32–1.71	0.48
Others/unknown	1.53	0.08–31.32	0.78
Alcohol (ref = no)			0.31
Unknown	1.69	0.69–4.15	0.25
Yes	1.74	0.81–3.75	0.16
Smoking (ref = non-smoker)			0.99
Current smoker	0.96	0.38–2.42	0.92
Former smoker	1.04	0.50–2.19	0.91
Unknown	1.11	0.38–3.23	0.85
BMI (ref = normal weight)			0.07
Underweight	3.06	0.87–10.79	0.08
Overweight	1.58	0.70–3.55	0.27
Obese	2.72	1.21–6.09	0.02
Acute pancreatitis (yes vs no)	0.92	0.49–1.73	0.80
Diabetes (yes vs no)	1.44	0.77–2.70	0.25
Radiographic reports (yes vs no)			
Atrophy	1.18	0.61–2.26	0.63
Calcification	0.63	0.32–1.23	0.18
Cyst	1.24	0.65–2.35	0.52
Ductal dilatation	10.46	4.03–27.12	<0.001
Ductal irregularity	0.91	0.36–2.27	0.83
Calculi	0.59	0.20–1.75	0.34
Filling defect	1.02	0.25–4.14	0.98
Side branch	1.00	0.42–2.41	0.99
Stricture	1.92	0.81–4.54	0.14

BMI, body mass index; CL, confidence limit; HR, hazard ratio.

sex-matched reference populations. Patients with specific duct abnormalities who did not experience pancreatic cancer within 1 year were still at a >10-fold increased risk of developing subsequent malignancy as compared to patients with CP without duct dilatation. Moreover, we found that obese patients with CP experienced a 2.7-fold increased risk of pancreatic cancer as compared to patients with CP of normal weight. Periodic surveillance of these patients may help lead to early detection and treatment of pancreatic cancer.

We observed a 12-fold increased risk of pancreatic cancer in suspected patients with CP (see Table S2, Supplementary Digital Content 1, http://links.lww.com/CTG/A222) as compared to the referent population. Consistent with this, a recent meta-analysis of case-control and cohort studies estimated a relative risk of

pancreatic cancer of 16.16 (95% confidence interval: 12.59–20.73) due to CP, with a 2-year lag period after CP diagnosis (2). Previous studies that have examined the risk of pancreatic cancer with no specific lag period showed a marked variation in relative risk estimates ranging from 1.5 to 28 based on the proximity in timing to the diagnosis of CP. The present study population was defined in a different manner compared with previous cohort studies that have defined CP based on either the ICD diagnosis codes only from national health registries (4,12,13) or clinical evaluation from the local disease registries (2,14–17). This distinction is important because our definitions incorporated both physician diagnosis of CP and pancreatic abnormality on imaging. By selecting patients with least 1 positive CP radiographic feature, we may have enriched the CP population to those with progression of pancreatic disease, potentially attributable to pancreatic cancer.

The 5-year incidence of pancreatic cancer (3.2%) in our study population of suspected patients with CP was even higher than the 10-year incidence of pancreatic cancer (1.8%) reported for the CP cohort populations investigated in the past (14). That the previous study estimated the risk of pancreatic cancer in persons with at least 2 years of follow-up, among younger individuals (median age: 40–50 years) over a recruitment period (1946–1989) when there were greater competing risks by death due to other causes, collectively contribute to the differences in pancreatic cancer risk estimates.

Duct dilatation was the most salient radiographic feature of CP that pointed to greatly elevated risk of pancreatic cancer, estimated at 6.3% in 5 years. The finding that pancreatic duct dilatation may be an early harbinger of pancreatic cancer has been previously demonstrated in a study of 27 patients with CP and 21 patients with pancreatic cancer (18). Conversely, another study of 427 patients undergoing endoscopic ultrasound, of whom 42 developed pancreatic cancer, did not substantiate this and rather attributed periductal hypoechoic sign as an indicator of pancreatic cancer (19). Our study of 1,766 suspected patients with CP, in whom 46 were diagnosed with pancreatic cancer, is the largest study, to date, establishing the relationship between CP radiographic features and the diagnosis of pancreatic cancer. Furthermore, we used the NLP-based methods for radiographic reports, thus developing a scalable algorithm for identifying CP features that foretells the future risk of pancreatic cancer.

Obesity was a risk factor for future development of pancreatic cancer in this cohort of suspected patients with CP. This is consistent with the trends in the general population, where obesity is associated with a 20%–50% increased risk of pancreatic cancer (20–22). Of many correlates of BMI, central obesity has shown a strong and consistent association with pancreatic cancer (20,23). Central obesity may promote tumor growth by way of causing inflammation in the pancreas. *In vivo* experiments also demonstrate that obese mice are prone to higher ductal cell replication (24) and development of pancreatic cancer (25).

There were several limitations to the present study. First, our definition of CP used a combination of diagnosis codes and imaging findings. Although this approach represents a clear advantage over reliance on the diagnosis codes alone, which has limited specificity (26), there remains a possibility of misclassification. Our imaging-based approach would likely fail to include patients with small-duct or "minimal change" CP (27) that may have been detected on endosonography or based on the results of pancreas function testing. Given that nearly a quarter of our study population underwent transabdominal ultrasound and
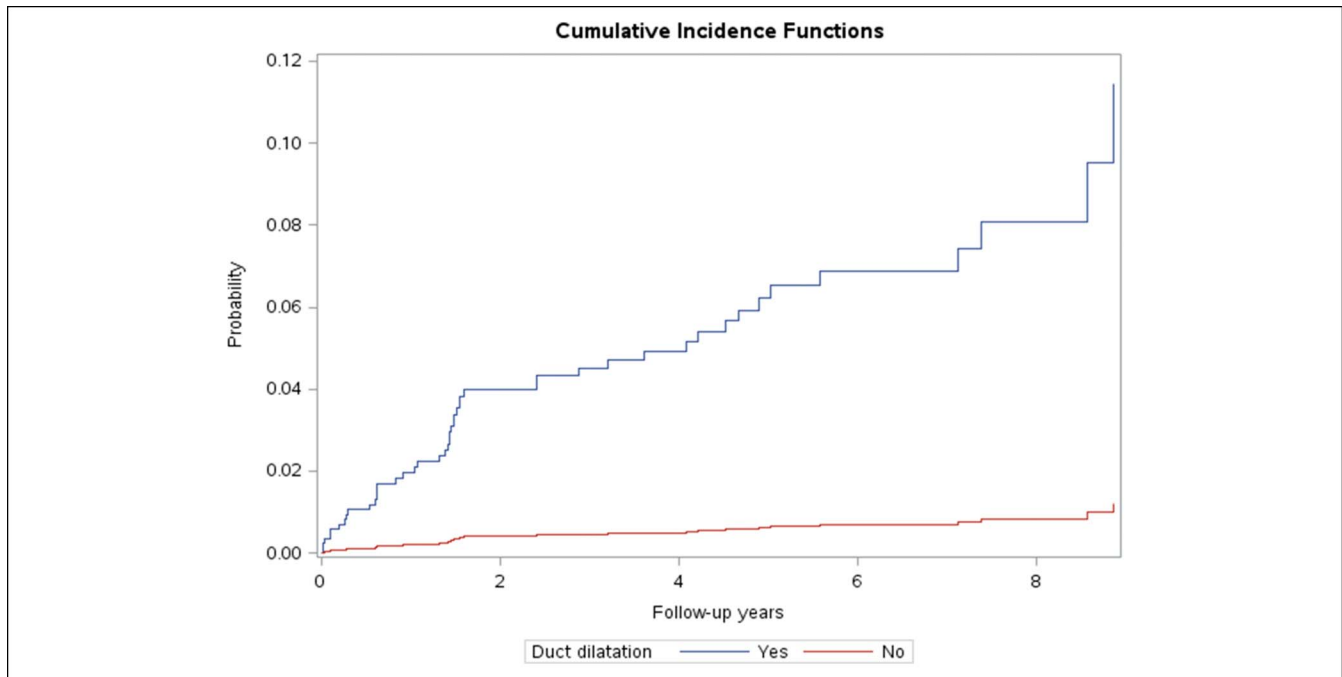
PANCREAS

Figure 3. Cumulative incidence curve of pancreatic cancer ≥1 year after abnormal pancreas imaging in patients with suspected chronic pancreatitis by presence of ductal dilatation.

less than 10% underwent MRI, accurate determination of duct diameter may not have been possible. Based on a manual review of the first abnormal pancreas radiograph of the patients with CP, only 42% of noncalcific CP showed a classic manifestation of CP, whereas 50% of patients had or developed acute pancreatitis without developing definite features of CP. Our manual review consisted of only 10% of the noncalcific patients with CP, and our algorithm has not been externally validated at other health systems. The accuracy of ICD coding we found is in line with the previous observation that ICD-coded diagnoses are limited in accurately identifying definite cases of CP (11). Among patients without a history of acute pancreatitis, timing of the disease onset regarding CP is very difficult to ascertain (28). This represents an underlying challenge in studying the natural history of CP because the ability to establish a diagnosis in the early stages of disease remains elusive (8). Given these limitations, our study population would be best described as persons with suspected CP rather than comprising solely of persons with definitive CP. In addition, there were only 46 patients with incident pancreatic cancer, thus limiting the power of the analysis to identify predictors of pancreatic cancer. Finally, a substantial portion of the study population were missing data on smoking (25%) and alcohol consumption (11%), which could have limited power and the accuracy of the relative risk estimates for smokers and alcohol consumers.

The strengths of the present study include the demonstration of an important role for state-of-the-art approaches to data acquisition such as NLP in helping to provide new insight into cancer prediction. Our methods were developed using open source tools and can be replicated across health systems with appropriate expertise in NLP. Previous applications of NLP on pancreatic disorders are few and have been limited to identifying existing pancreatic cysts (29,30) or IPMNs (31). These include NLP of surgical pathology texts to identify IPMNs beyond known IPMN cases in a manually curated registry (31) and NLP of any text-based reports in the medical records to identify pancreatic cysts or ductal dilatation (29,30), which had close concordance with manual review. Our study demonstrates that NLP can not only identify existing pancreatic abnormalities of immediate concern but also points to features that indicate high future risk of pancreatic cancer, thereby demonstrating utility for surveillance. In addition, our study had applied NLP to only reports of abdomen-focused imaging studies and thus was more selective than application of NLP to any reports in the medical records. Furthermore, our study cohort was drawn from a diverse community-based setting, thus our findings are generalizable to large health systems.

In summary, our study shows that in patients diagnosed with CP, obesity and duct dilatation on radiographic imaging point to persons with elevated longer-term risk of pancreatic cancer.

Study Highlights

WHAT IS KNOWN

✓ Pancreatic cancer occurs more often in patients with CP than in the general population. However, we lack guidelines for identifying which patients with CP should undergo surveillance for pancreatic cancer.

WHAT IS NEW HERE

✓ Among patients suspected for CP, ductal dilatation is a salient feature of future risk of pancreatic cancer.

TRANSLATIONAL IMPACT

✓ Active surveillance may be warranted in patients suspected with CP with ductal dilatation, given the elevated risk for pancreatic cancer.

REFERENCES

1. Kirkegard J, Mortensen FV, Cronin-Fenton D. Chronic pancreatitis and pancreatic cancer risk: A systematic review and meta-analysis. Am J Gastroenterol 2017;112:1366–72.
2. Malka D, Hammel P, Maire F, et al. Risk of pancreatic adenocarcinoma in chronic pancreatitis. Gut 2002;51:849–52.
3. Yadav D, Lowenfels AB. The epidemiology of pancreatitis and pancreatic cancer. Gastroenterology 2013;144:1252–61.
4. Bang UC, Benfield T, Hyldstrup L, et al. Mortality, cancer, and comorbidities associated with chronic pancreatitis: A Danish nationwide matched-cohort study. Gastroenterology 2014;146:989–94.
5. Service up. Final update summary: Pancreatic cancer: Screening. U.S. Preventive Services Task Force (https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/pancreatic-cancer-screening). 2015.
6. Munigala S, Kanwal F, Xian H, et al. New diagnosis of chronic pancreatitis: Risk of missing an underlying pancreatic cancer. Am J Gastroenterol 2014;109:1824–30.
7. Conwell DL, Lee LS, Yadav D, et al. American pancreatic association practice guidelines in chronic pancreatitis: Evidence-based report on diagnostic guidelines. Pancreas 2014;43:1143–62.
8. Bird S, Klein E, Loper E. Natural Language Processing With Python. O'Reilly Media Inc; 2009 (http://nltk.org/book).
9. Chapman BE, Lee S, Kang HP, et al. Document-level classification of CT pulmonary angiography reports based on an extension of the ConText algorithm. J Biomed Inform 2011;44:728–37.
10. Manning C, Surdeanu M, Bauer J, et al. The Stanford CoreNLP Natural Language Processing Toolkit, In Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics. System Demonstrations. Association for Computational Linguistics: Baltimore, MD, 2014.
11. Xiao AY, Tan ML, Plana MN, et al. The use of International Classification of Diseases codes to identify patients with pancreatitis: A systematic review and meta-analysis of diagnostic accuracy studies. Clin Transl Gastroenterol 2018;9:191.
12. Goldacre MJ, Wotton CJ, Yeates D, et al. Liver cirrhosis, other liver diseases, pancreatitis and subsequent cancer: Record linkage study. Eur J Gastroenterol Hepatol 2008;20:384–92.
13. Karlson BM, Ekbom A, Josefsson S, et al. The risk of pancreatic cancer following pancreatitis: An association due to confounding? Gastroenterology 1997;113:587–92.
14. Lowenfels AB, Maisonneuve P, Cavallini G, et al. Pancreatitis and the risk of pancreatic cancer: International Pancreatitis Study Group. N Engl J Med 1993;328:1433–7.
15. Pedrazzoli S, Pasquali C, Guzzinati S, et al. Survival rates and cause of death in 174 patients with chronic pancreatitis. J Gastrointest Surg 2008;12:1930–7.
16. Talamini G, Falconi M, Bassi C, et al. Incidence of cancer in the course of chronic pancreatitis. Am J Gastroenterol 1999;94:1253–60.
17. Ueda J, Tanaka M, Ohtsuka T, et al. Surgery for chronic pancreatitis decreases the risk for pancreatic cancer: A multicenter retrospective analysis. Surgery 2013;153:357–64.
18. Inoue K, Ohuchida J, Ohtsuka T, et al. Severe localized stenosis and marked dilatation of the main pancreatic duct are indicators of pancreatic cancer instead of chronic pancreatitis on endoscopic retrograde balloon pancreatography. Gastrointest Endosc 2003;58:510–5.
19. Lee SH, Ozden N, Pawa R, et al. Periductal hypoechoic sign: An endosonographic finding associated with pancreatic malignancy. Gastrointest Endosc 2010;71:249–55.
20. Bracci PM. Obesity and pancreatic cancer: Overview of epidemiologic evidence and biologic mechanisms. Mol Carcinog 2012;51:53–63.
21. Genkinger JM, Kitahara CM, Bernstein L, et al. Central adiposity, obesity during early adulthood, and pancreatic cancer mortality in a pooled analysis of cohort studies. Ann Oncol 2015;26:2257–66.
22. Genkinger JM, Spiegelman D, Anderson KE, et al. A pooled analysis of 14 cohort studies of anthropometric factors and pancreatic cancer risk. Int J Cancer 2011;129:1708–17.
23. Aune D, Greenwood DC, Chan DS, et al. Body mass index, abdominal fatness and pancreatic cancer risk: A systematic review and non-linear dose-response meta-analysis of prospective studies. Ann Oncol 2012;23:843–52.
24. Butler AE, Galasso R, Matveyenko A, et al. Pancreatic duct replication is increased with obesity and type 2 diabetes in humans. Diabetologia 2010;53:21–6.
25. Incio J, Liu H, Suboj P, et al. Obesity-induced inflammation and desmoplasia promote pancreatic cancer progression and resistance to chemotherapy. Cancer Discov 2016;6:852–69.
26. Saligram S, Lo D, Saul M, et al. Analyses of hospital administrative data that use diagnosis codes overestimate the cases of acute pancreatitis. Clin Gastroenterol Hepatol 2012;10:805–11.e1.
27. Walsh TN, Rode J, Theis BA, et al. Minimal change chronic pancreatitis. Gut 1992;33:1566–71.
28. Witt H, Apte MV, Keim V, et al. Chronic pancreatitis: Challenges and advances in pathogenesis, genetics, diagnosis, and therapy. Gastroenterology 2007;132:1557–73.
29. Mehrabi S, Schmidt CM, Waters JA, et al. An efficient pancreatic cyst identification methodology using natural language processing. Stud Health Technol Inform 2013;192:822–6.
30. Roch AM, Mehrabi S, Krishnan A, et al. Automated pancreatic cyst screening using natural language processing: A new tool in the early detection of pancreatic cancer. HPB (Oxford) 2015;17:447–53.
31. Al-Haddad MA, Friedlin J, Kesterson J, et al. Natural language processing for the development of a clinical registry: A validation study in intraductal papillary mucinous neoplasms. HPB (Oxford) 2010;12:688–95.