Lead-time trajectory of CA19-9 as an anchor marker for pancreatic cancer early detection

Johannes F. Fahrmann[1], **C. Max Schmidt**[2], **Xiangying Mao**[1], **Ehsan Irajizad**[1], **Maureen Loftus**[3], **Jinming Zhang**[3], **Nikul Patel**[1], **Jody Vykoukal**[1], **Jennifer B. Dennison**[1], **James P. Long**[4], **Kim-Anh Do**[4], **Jianjun Zhang**[2], **John A. Chabot**[5], **Michael D. Kluger**[5], **Fay Kastrinos**[6,7], **Lauren Brais**[3], **Ana Babic**[3], **Kunal Jajoo**[8], **Linda S. Lee**[8], **Thomas E. Clancy**[9], **Kimmie Ng**[3], **Andrea Bullock**[10], **Jeanine Genkinger**[5], **Michele T. Yip-Schneider**[2], **Anirban Maitra**[6], **Brian M. Wolpin**[3], **Samir Hanash**[1]

[1]Department of Clinical Cancer Prevention, The University of Texas M. D. Anderson Cancer Center, Houston, Texas.

[2]Department of Surgery, Indiana University School of Medicine, Indianapolis, Indiana.

[3]Dana-Farber Brigham and Women's Cancer Center, Division of Gastrointestinal Oncology, Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts.

4Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, Houston, Texas.

5Herbert Irving Comprehensive Cancer Center, Columbia University Medical Center, Department of Epidemiology, Columbia Mailman School of Public Health, New York, New York.

6Division of Digestive and Liver Diseases, Columbia University Irving Medical Cancer and the Vagelos College of Physicians and Surgeons, New York, New York, Department of Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas.

7Herbert Irving Comprehensive Cancer Center, Columbia University Medical Center, Department of Surgery, New York, New York.

8Division of Gastroenterology, Hepatology and Endoscopy, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.

9Dana-Farber Brigham and Women's Cancer Center, Division of Surgical Oncology, Department of Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.

10Division of Hematology/Oncology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts

Abstract

Background & Aims: There is substantial interest in liquid biopsy approaches for cancer early detection, among subjects at risk, using multi-marker panels. CA19-9 is an established circulating biomarker for pancreatic cancer. However, its relevance for pancreatic cancer early detection or for monitoring subjects at risk has not been established.

Methods: CA19-9 levels were assessed in blinded sera from 175 subjects collected up to 5 years prior to diagnosis of pancreatic cancer and from 875 matched controls from the PLCO Cancer Screening Trial. For comparison of performance, CA19-9 was assayed in blinded independent sets of samples collected at diagnosis from 129 subjects with resectable pancreatic cancer and 275 controls (100 healthy subjects; 50 with chronic pancreatitis; and 125 with non-cancerous pancreatic cysts). The complementary value of two additional protein markers, TIMP1 and LRG1, was determined.

Results: In the PLCO cohort, levels of CA19-9 increased exponentially starting at two years prior to diagnosis with sensitivities reaching 60% at 99% specificity within 0–6 months prior to diagnosis for all cases and 50% at 99% specificity for cases diagnosed with early-stage disease. Performance was comparable for distinguishing newly diagnosed cases with resectable pancreatic cancer from healthy controls (64% sensitivity at 99% specificity). Comparison of resectable pancreatic cancer cases to subjects with chronic pancreatitis yielded 46% sensitivity at 99% specificity and for subjects with non-cancerous cysts 30% sensitivity at 99% specificity. For pre-diagnostic cases below cut-off value for CA19-9, the combination with LRG1 and TIMP1 yielded an increment of 13.2% in sensitivity at 99% specificity (p=0.031) in identifying cases diagnosed within 1 year of blood collection.

Conclusion: CA19-9 can serve as an anchor marker for pancreatic cancer early detection applications.

Lay Summary

CA19-9 can serve as an anchor marker for pancreatic cancer early detection. Inclusion of additional markers such as LRG1 and TIMP1 may have value for identifying cases missed by CA19-9 alone.

Keywords

Biomarker; Detection; Pancreatic Cancer

Introduction

Pancreatic ductal adenocarcinoma (PDAC) remains an intractable disease with dismal 5-year survival rates[1]. Poor survival rates are attributed to the fact that most patients with PDAC present with locally advanced or distant disease (80–85%), with only a minority of patients initially presenting with surgically resectable disease (15–20%)[2, 3]. Several lines of evidence clearly indicate that survival rates can be improved if PDAC is identified at early-stage, when the disease is localized and still amenable to surgical intervention[1, 4–7].

Serum CA19-9 is the main clinically validated tumor marker used in the management of PDAC. Concerns regarding CA19-9 have included limited sensitivity for early-stage disease[8], false-negatives values for subjects that do not produce CA19-9 on account of a fucosyltransferase deficiency[9, 10] and false-positives values in cases with biliary obstruction[11–13]. Nevertheless, CA19-9 has proven useful for predicting tumor stage and resectability, overall survival, and patient response to therapy[14].

There is currently substantial interest in liquid biopsy approaches for cancer screening whether in a general population setting or for subjects at increased risk. A wide range of blood-based biomarkers are currently being investigated including circulating tumor DNA (ctDNA) representing mutations or altered methylation patterns; RNA, proteins, metabolites, autoantibodies to tumor antigens; exosomes; and circulating tumor cells[15–20]. Performance requirements depend on subjects' risk, with a need for very high specificity for subjects at average risk. In the case of pancreatic cancer, risk is increased based on a personal and family history of cancer, as well as other risk factors[21–26].

CA19-9 has the potential to serve as an anchor marker for pancreatic cancer screening that may be complemented with other types of markers. CA19-9 has been reported to be elevated in pre-diagnostic PDAC cases among women recruited to the UKCTOCS cohort[27]. The cohort consisted of post-menopausal women with limited information regarding tumor stage at the time of diagnosis[27]. In another study, CA19-9 was found to be elevated in pre-diagnostic PDAC cases among participants recruited to the European EPIC cohort[28]. In this cohort, most cases presented with advanced stage disease which did not allow assessment of CA19-9 for identifying resectable tumors. In a third study, performance of CA19-9 in the pre-diagnostic setting when multiplexed with 66 additional potential markers was more modest, which may possibly be attributable to assay format[29].

In this study, based on a single-plex ELISA assay to reduce non-specific reactivity, we assessed the lead-time trajectory of CA19-9 using blinded pre-diagnostic sera from 175 PDAC cases and 875 matched cancer-free controls from the PLCO cohort. Similarities and differences in performance in the pre-diagnostic vs diagnostic settings were assessed using samples from newly diagnosed subjects with resectable disease. Samples from subjects presenting with chronic pancreatitis (CP) or cystic lesion(s) of the pancreas and from healthy subjects were used as controls. We additionally evaluated the potential of two additional previously validated protein markers LRG1 and TIMP1 using samples from newly diagnosed subjects with early-stage disease[18], for their complementary performance with CA19-9 in the pre-diagnostic setting.

Methods

PLCO Cohort

The PLCO Cancer Screening Trial is a randomized multicenter trial in the United States which aimed to evaluate the impact of early detection procedures for prostate, lung, colorectal and ovarian cancer on disease-specific mortality. Detailed information regarding the PLCO cohort is provided elsewhere[30]. All subjects involved in this study were enrolled with written consent as a criterion for eligibility to participate in the PLCO trial. Study recruitment and randomization began November 1993 and was completed in July 2001. PLCO eligibility criteria excluded subjects with a previous personal history of PLCO cancers, ongoing cancer treatment (excluding basal-cell and squamous-cell skin cancer), participation in another cancer screening or cancer primary prevention trial, and a recent screening test for prostate or colorectal cancer. The cohort comprises approximately 155,000 men and women aged 55 to 74 years old at baseline entry. Study participants completed a baseline questionnaire at study entry that includes demographic, personal, and medical information including diabetes status.

Pancreatic cancer cases were identified by self-report in annual mail-in surveys, state cancer registries, death certificates, physician referrals and reports from next of kin for deceased individuals. All medical and pathologic records related to pancreatic cancer diagnosis and supporting documentation were obtained and confirmed by PLCO staff.

Herein, we analyzed sera from 175 pancreatic cancer cases from the PLCO cohort that were diagnosed within 5 years of blood draw and 875 matched controls (Table 1). Pancreatic cancers were classified as localized, regional, distant, or un-staged using the National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) historic staging system. Localized disease indicates confinement of the tumor to the pancreas, regional disease indicates a tumor that has involved the regional lymph nodes with or without direct extension into the surrounding organs or tissues, distant refers to a tumor that has spread to part of the body remote from the primary tumor, and un-staged indicates that there was insufficient information in the medical records to stage the tumor. Controls, alive at the time when the index case was diagnosed, were matched to cases at a ratio of 5:1 (control:case) based on the distribution of age, race, sex, and calendar date of blood draw in 2-month blocks within the case cohort.

Newly diagnosed pancreatic cancer cohorts

To compare classifier performance of CA19-9 in the pre-diagnostic and diagnostic settings and assess its performance for distinguishing pancreatic cancer from patients presenting with chronic pancreatitis (CP) or non-cancerous cystic lesions that were not represented in the PLCO cohort, we evaluated plasma samples with two test sets. Test Set #1 consisted of 99 patients with early-stage, resected pancreatic adenocarcinoma, 100 healthy controls, and 50 patients with CP. Patient characteristics are provided in Table S5. Pancreatic cancer patients provided informed written consent to collection of pre-treatment plasma samples and clinical data abstraction. PDAC patients were recruited from cancer clinics at Dana-Farber Cancer Institute/Brigham and Women's Hospital (DFCI/BWH; N=69), Beth Israel Deaconess Medical Center (BIDMC; N=15), and Columbia University Irving Medical Center (CUMC; N=15). Healthy control patients were recruited from DFCI/BWH (N=94) and CUMC (N=6). Healthy controls were undergoing screening colonoscopy (N=91) or accompanying a non-blood related relative to an appointment at a gastrointestinal cancer clinic (N=9). Healthy controls had no history of cancer in the 5 years prior to sample collection. Pancreatic cancer patients and healthy controls were matched on gender and age at the time of blood collection. Fifty chronic pancreatitis patients were recruited from gastroenterology clinics at DFCI/BWH (N=30), BIDMC (N=15), and CUMC (N=5). Patients were included if clinic notes from a gastroenterologist indicated a diagnosis of CP. Pancreatic cancer and CP patients were not gender or age matched. Clinical data abstraction was performed identically across the sites, with data uploaded to a password-protected REDCap database. All plasma samples were collected and processed according to a uniform, standardized protocol across the sites and patient groups.

Test set #2 consisted of 125 patients with low dysplastic grade pancreatic cyst and 30 patients with resectable pancreatic cancer (8 PDAC and 22 intraductal papillary mucinous neoplasms (IPMNs) with an associated invasive ductal adenocarcinoma) from the Indiana University School of Medicine. Patient characteristics are provided in Table S6. All patients underwent surgical resection of their cystic lesion, and plasma samples were collected prior to surgery. Dysplastic grade was histopathology confirmed after surgical resection and determined according to WHO criteria.

Enzyme-linked Immunosorbent Assays

Plasma protein concentrations for CA19-9, LRG1 and TIMP1 were determined by bead-based ELISA assays using Luminex multiplexed assay technology ((CA19-9) HCCBP1-58MAG, (LRG1) HCVD6MAG-67K, (TIMP1) HTMP1MAG-54K, Millipore) as previously described[18, 31]. We note that the assays for CA19-9, LRG1 and TIMP1 described herein are for intended research use only and are not for diagnostic applications. For all ELISA experiments, each sample was assayed in singlet and the absorbance or chemiluminescence measured with a SpectraMax M5 microplate reader (Molecular Devices, Sunnyvale, CA). Samples were analyzed in a blinded fashion, ratio of case:control was equilibrated across each analytical plate to mitigate potential bias. An internal control sample was run in every plate and each value of the samples was divided by the mean value of the internal control in the same plate to correct for interpolate variability. Biomarker scores for the 3-marker panel of LRG1, TIMP1 and CA19-9 were derived using fixed β-

coefficients from a previously developed logistic regression model[18]. Coefficient of variation (CV) values for CA19-9, LRG1 and TIMP1 in pooled quality control samples were 11.0, 8.0 and 9.7, respectively.

Statistical Analyses

Model discrimination was assessed by receiver operating characteristic curve (ROC). Time-dependent ROC analyses were performed using pROC (version 1.15.3) in the R software environment (version 3.6.1, The R Foundation, https://www.r-project.org). The 95% confidence intervals (CI) for AUCs were estimated using Delong method[32]. A fourth order fitted spline curve was used to fit the trajectory of the AUC and sensitivity at 99% specificity for CA19-9 performance in relation to time to diagnosis from baseline blood draw. For these analyses, we considered cases relative to matched controls. Specificities at 95% and 99% sensitivities as well as sensitivities at 95% and 99% specificities were determined from ROC curves. The 95% CI of sensitivities and specificities at specific cut-offs were calculated based on the Exact binomial confidence limits, using epi.tests function from epiR (version 1.0.15). P values are reported based on 2-sided Wilcoxon rank sum test unless otherwise specified.

In order to determine the 1-year risk of PDAC and discrimination estimates that reflect the background population of the entire PLCO study, we used the case-control approach of Schlesselman 1982[33]. Briefly, risk scores based on continuous CA19-9 values were calculated based on a logistic regression model[34]. Samples assayed for CA19-9 herein reflect a nested case-control cohort that enriches for cases and, therefore, do not reflect the true risk of PDAC in the general population. Thus, to calculate the estimated absolute risk value based on continuous values of CA19-9 that reflect the true risk of PDAC in the entire PLCO population, we used the following formula to calculate absolute risk:

$$Risk = \frac{\exp\left(\beta'_0 + \beta_1 \times \left(CA19.9 \frac{U}{ml}\right)\right)}{1 + \exp\left(\beta'_0 + \beta_1 \times \left(CA19.9 \frac{U}{ml}\right)\right)}$$

where

$$\beta'_0 = \beta_0 - \log\left(\frac{P_{data}}{1 - P_{data}}\right) + \log\left(\frac{P_{Population}}{1 - P_{Population}}\right)$$

In this equation, β_0 is the intercept derived from logistic regression in the nested case control cohort, P_{data} is the prevalence of the disease in our case-enriched dataset $Prob(X|D_{data} = 1)$ and $P_{Population}$ is the prevalence of the disease in the general population $Prob(X|D_{Population} = 1)$. Quartiles were determined based on CA19-9 measurements in cases and controls to provide a comparison of the estimated 1-year risk for an individual with a CA19-9 value in the first quartile and an individual with a CA19-9 value in the fourth quartile.

The CA19-9 cut-off for the 99% specificity population (CA19-9 > 97.385 U/mL) was defined as the threshold value using the measurements of CA19-9 in the one-year PLCO

population. Emphasis was placed on 99% specificity given the low prevalence of PDAC and the need to mitigate false-positives. Confusion matrices were utilized to describe the classification model of the 3-marker panel or CA19-9 alone at a >99% specificity cutoff. Rows of the matrix display the predicted classes (case or control) whereas columns represent the actual classes (case or control). To test whether 3-marker panel yielded statistically significant classifier improvement over CA19-9, the McNemar exact test was applied to the 2×2 contingency table wherein the first cell represents the number of patients that both markers predict correctly (a), second one represents the number of patients correctly identified by CA19-9 and misclassified by 3-marker panel (b), third one represents the number of patients misclassified by CA19-9 but correctly identified by 3-marker panel (c) and the last cell represents the number of patients misclassified by both markers (d). Therefore, the null and alternative hypothesize are as follows:

$$H_0 : P_b = P_c$$

$$H_a : P_b < P_c$$

Herein, P_i denotes the probability of occurrence in cell i. Exact binomial test was used to achieve p-value[35].

Figures were generated using GraphPad Prism Version 8.0.0 (GraphPad Software, Inc).

Results

CA19-9 performance trajectory in the pre-diagnostic setting

Testing of CA19-9 was performed using blinded pre-diagnostic sera from 175 pancreatic cancer cases and 875 matched cancer-free controls from the PLCO cohort. Among the 175 pancreatic cancer cases, 62 were diagnosed within the first year and 33 were diagnosed one to two years after blood collection. The remainder were diagnosed 2–5 years after blood collection. Controls were matched based on age, sex and race (Table 1).

When considering all cases compared to all controls, CA19-9 yielded an Area Under the Receiver Operating Characteristic Curve (AUC) of 0.68 (95% CI: 0.64–0.73) with 16.6% sensitivity at 99% specificity. Whereas CA19-9 exhibited marginal performance two to five years prior to diagnosis, an exponential rise in the classifier performance of CA19-9 was revealed beginning at 2 years prior to diagnosis with an AUC of 0.77 (95% CI: 0.69–0.86) 1–2 years prior to diagnosis reaching an AUC of 0.87 (95% CI: 0.78–0.96) when considering those cases diagnosed within 0–6M of baseline blood draw (Figure 1 and Table S1). Corresponding sensitivities of CA19-9 at 99% specificity for 6M-12M, and 0–6M prior to diagnosis were 24.3%, and 60%, respectively (Figure 1B and Tables S1, S2).

The one-year PDAC cancer risk estimates for each study participant according to their CA19-9 values adjusted for prevalence of disease based on the entire intervention arm of the PLCO population[30] are shown in Figure 2. We estimated 1-year risks at 0.013% and 0.049% assuming CA19.9 value equals to the first and fourth quartile, respectively. The

25% percentile, median and 75% percentile one-year risk estimates for participants who later diagnosed with PDAC were 0.02%, 0.10% and 0.62%, respectively; these values were considerably lower for controls where we have 0.01%, 0.02% and 0.04%, respectively.

Of the 175 pancreatic cancer cases in the PLCO cohort, 35 had localized disease, 33 had regional disease and 92 had distant metastases at the time of clinical diagnosis (Table 1). CA19-9 performance for distinguishing cases presenting with localized disease 1–2 years, 0–1 year and 0–6 months prior to diagnosis from time-interval matched controls yielded sensitivities of 12.5%, 15.4% and 50% at 99% specificity (Figure 3; Table S3, S4). When considering cases presenting with regional disease, CA19-9 yielded sensitivity at 99% specificity for cases 1–2 years, 0–1 year and 0–6M prior to diagnosis of 0%, 27.3% and 50%, respectively (Figure 3; Table S3, S4). Resultant sensitivity at 99% specificity of CA19-9 for distinguishing cases presenting with distant disease 1–2 years, 0–1 year and 0–6M prior to diagnosis compared to time-interval matched controls were 10.5%, 52.9% and 68.8% at 99% specificity, respectively (Figure 3, Figure S1, Table S3, S4).

Notably, cases presenting with localized disease yielded a mean (+/− StDev) CA19-9 value of 116.24 +/− 94.62 U/mL 0–6 months prior to diagnosis whereas cases presenting with distant disease exhibited a mean (+/− StDev) CA19-9 value of 106.3 +/− 110.8 U/mL 6–12M prior to diagnosis (Figure S1A–D). These data indicate a six-month time-interval latency to achieve comparable levels of CA19-9 between cases presenting with localized disease and those with distant disease, suggesting that rapid disease progression may occur during this period.

CA19-9 classifier performance for newly diagnosed resectable PDAC

We compared performance of CA19-9 for detection of cases subsequently diagnosed with early-stage pancreatic cancer with performance at the time of diagnosis of early-stage pancreatic cancer. Samples from an independent cohort consisting of 99 subjects with resectable pancreatic cancer, 100 matched healthy controls and 50 subjects with CP were tested (Test Set #1). Patient characteristics are provided in Table S5. Classifier performance of CA19-9 for distinguishing cases from matched healthy controls was 0.90 (95% CI: 0.86–0.94) with 64.7% sensitivity at 99% specificity (Figure 4A) and 0.86 (95% 0.80–0.92) with 46.5% sensitivity at 99% specificity when considering cases compared to controls with CP (Figure 4B). Notably, the classifier performance of CA19-9 for identifying of cases versus healthy subjects as determined in Test Set #1 was markedly like the performance of CA19-9 for cases presenting with localized disease and that were diagnosed within 0–6M of baseline blood draw in the PLCO cohort (AUC= 0.89) (Figure 1A). These findings reinforce lead-time utility of CA19-9 for identification of early-stage pancreatic cancer when the disease is still operable.

We further evaluated classifier performance of CA19-9 for distinguishing pancreatic cancer cases from subjects harboring non-cancerous IPMN using in an independent cohort of 30 resectable pancreatic cancer cases (8 PDAC and 22 IPMNs with an associated invasive ductal adenocarcinoma) and 125 subjects with non-cancerous IPMN (Test Set #2) (Table S6). Classifier performance of CA19-9 for distinguishing cases from subjects harboring non-cancerous cysts was 0.76 (95% CI: 0.66–0.85) with 30% sensitivity at 99% specificity

(Figure 4C). When stratifying cases according to those with adenocarcinoma or invasive IPMN, CA19-9 yielded respective AUCs of 0.85 (95% CI: 0.73–0.97) and 0.72 (95% CI: 0.61–0.84) with corresponding sensitivity of 37.5% and 27.3% at 99% specificity compared to controls (Figure 4D–E).

Additive performance of LRG1 and TIMP1 for identifying pancreatic cancer cases not detected with CA19-9 alone

Selection of the cutoff point to provide exceptionally high specificity to minimize false-positive results in attenuated sensitivity and cases that are missed. Further, approximately ~10% of PDAC subjects have fucosyltransferase deficiency and therefore do not produce CA19-9[9, 10]. Thus, additional biomarker(s) are needed to complement CA19-9 and improve sensitivity while maintaining high specificity. We previously validated two protein markers, LRG1 and TIMP1 for detection of symptomatic PDAC at an early-stage. A combination rule of LRG1, TIMP1 and CA19-9 yielded improved classifier performance for detection of PDAC compared to CA19-9 alone[18]. In the PLCO cohort, for cases diagnosed within 1 year of baseline blood draw that were 'negative' for CA19-9 (38 out of 62 subjects) based on a 99% specificity cutoff (CA19-9 < 97.385 U/mL), the three marker protein panel, using previously fixed beta-coefficients of the logistic regression model[18], yielded an additional 13.2% sensitivity at >99% specificity, compared to CA19-9 alone in identifying cases diagnosed within one year of baseline blood draw (Figure 5A; Table S7). A confusion matrix reporting the performance of the classification model corresponding to the three-marker panel and CA19-9 alone indicates that the three-marker panel correctly identified an additional five cases that were missed by CA19-9 alone (1-sided McNemar exact test p: 0.031) when allowing for equivalent number of false positives (1) (Figure 5B).

Discussion

Our findings establish the utility of CA19-9 for detecting pancreatic cancer amongst asymptomatic individuals. Elevated CA19-9 levels began to manifest as early as 2 years prior to clinical diagnosis. Importantly, CA19-9 provided important lead-time for detection of resectable disease, when multimodality treatment can lead to improved long-term survival[1, 4–6].

The low incidence of PDAC in the average risk population (~8–12 per 100,000) with a 1–3% lifetime risk of developing disease[36, 37] makes it challenging to implement screening for pancreatic cancer in the general population[36, 37]. Screening individuals that are at high-risk for pancreatic cancer increases the positive predictive value of the test and reduces the absolute number of false-positive tests. To-date, several groups have been identified at high-risk of PDAC including subjects with inherited risk[23, 38], individuals with mucinous cystic lesions of the pancreas[39], subjects with CP[22] as well as individuals older than 50 years of age and presenting with new-onset diabetes[24, 40]. To this end, our data implies that progression of disease can occur rapidly and that the window of opportunity to identify pancreatic cancer when disease is still operable is narrow. Consequently, testing of CA19-9 amongst high-risk subjects should be performed regularly with testing intervals matching their degree of risk. An initial rise in CA19-9 could then prompt more intensive follow-

up whereas a positive test based on a defined threshold would trigger an imaging-based modality such as contrast-enhanced pancreas protocol CT or MRI/MRCP.

We note some limitations to our study. Time-dependent AUCs and sensitivity at 99% specificity trajectories of CA19-9 were derived based on the availability of plasma samples at varying time points before cancer diagnosis from individual patients. Availability of serial samples would allow development of longitudinal algorithms and assessment of significance of incremental increases in CA19-9 levels. Sub-analyses assessing performance of CA19-9 for identifying diabetic subjects that went on to develop pancreatic cancer from diabetic subjects that did not go on to develop pancreatic cancer was not possible due to limited sample size. Data on benign conditions such as CP, pancreatic cystic lesions or obstructive jaundice were lacking in the PLCO cohort hence our reliance on additional sources of samples which yielded good performance for CA19-9 in distinguishing cases from individuals presenting with CP or subjects harboring mucinous cystic lesions of the pancreas with resultant AUCs consistent with estimates determined for cases compared to healthy controls. The prevalence of cysts in the general population is reported to be between 2.4 to13.5% and increases with age[41–44]. IPMNs are the most frequent type of pancreatic cyst[45, 46]. A recent systematic review and meta-analysis found that the probability of IPMN progressing to pancreatic cancer within 10 years is approximately 8% for low-risk IPMN and 25% for high-risk IPMN[47]. Current consensus guidelines recommend either resection of IPMN with high-risk of malignancy or surveillance of IPMN without surgical indications[48]. However, current radiological and clinical guidelines yield satisfactory sensitivity but lack specificity for predicting malignant IPMN with high-grade dysplasia or invasive carcinoma, compared with surgical pathology[49, 50]. In this regard, a recent study demonstrated that only 23% of resected IPMNs contain invasive or high-grade histology[51]. Thus, clinical management of IPMN patients remains a substantial challenge. Consequently, the utility of a screening test that can identify individuals harboring mucinous cystic lesions of the pancreas that are at high risk of malignant transformation or that actively harbor malignancy is clinically desirable. Our findings demonstrate that CA19-9 can serve as a lead marker for distinguishing pancreatic cancer from non-cancerous pancreatic cysts.

To-date, several blood-based tests have been proposed for interception of pancreatic cancer and other cancer types[15, 52, 53]. CancerSEEK, a multi-analyte panel, has shown promise for detection of pancreatic cancer at an early-stage with reported ~72% sensitivity at >99% specificity[16]. Studies exploring utility of cell-free DNA methylation patterns for detection of pancreatic cancer have reported sensitivity of 63% and 83% at >99% specificity for stage I and II disease, respectively[53]. We note that these studies were performed in the diagnostic setting and didn't consider individuals presenting with benign conditions of the pancreas. Moreover, the performance of CancerSEEK and cfDNA methylation, based on sensitivity at high specificity, was comparable with that of CA19-9 observed in our study, providing a compelling rationale for CA19-9 utility as an anchor marker for screening of pancreatic cancer. However, given that ~10% of individuals lack the ability to produce CA19-9[54], other markers would have complementary benefits as demonstrated for LRG1 and TIMP1 in our study resulting in capture of additional cases diagnosed within one year that were missed with CA19-9 alone.

In conclusion, our findings demonstrate utility of CA19-9 for screening of pancreatic cancer. Inclusion of additional markers as demonstrated for LRG1 and TIMP1 and potentially in combination with other marker types[15, 17, 19, 52, 53], would have value for identifying cases that do not meet CA19-9 cutoff thresholds.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

Conflicts-of-Interest:

B.M.W. receives research funding from Celgene and Eli Lilly and consulting for BioLineRx, Celgene, and GRAIL.

Abbreviations:

AUC	area under the curve
CP	chronic pancreatitis
IPMN	intraductal papillary mucinous neoplasms
PLCO	Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial
ROC	receiver operating characteristic curve
SEER	Surveillance, Epidemiology, and End Results

References

1. Pereira SP, Oldfield L, Ney A, et al. Early detection of pancreatic cancer. Lancet Gastroenterol Hepatol 2020.

2. Ryan DP, Hong TS, Bardeesy N. Pancreatic adenocarcinoma. N Engl J Med 2014;371:1039–49. [PubMed: 25207767]

3. Kleeff J, Korc M, Apte M, et al. Pancreatic cancer. Nat Rev Dis Primers 2016;2:16022. [PubMed: 27158978]

4. Neoptolemos JP, Kleeff J, Michl P, et al. Therapeutic developments in pancreatic cancer: current and future perspectives. Nat Rev Gastroenterol Hepatol 2018;15:333–348. [PubMed: 29717230]

5. Cloyd JM, Katz MH, Prakash L, et al. Preoperative Therapy and Pancreatoduodenectomy for Pancreatic Ductal Adenocarcinoma: a 25-Year Single-Institution Experience. J Gastrointest Surg 2017;21:164–174. [PubMed: 27778257]

6. Murphy JE, Wo JY, Ryan DP, et al. Total Neoadjuvant Therapy With FOLFIRINOX Followed by Individualized Chemoradiotherapy for Borderline Resectable Pancreatic Adenocarcinoma: A Phase 2 Clinical Trial. JAMA Oncol 2018;4:963–969. [PubMed: 29800971]

7. Blackford AL, Canto MI, Klein AP, et al. Recent trends in the incidence and survival of Stage 1A Pancreatic Cancer: A Surveillance, Epidemiology, and End Results analysis. J Natl Cancer Inst 2020.

8. Steinberg WM, Gelfand R, Anderson KK, et al. Comparison of the sensitivity and specificity of the CA19-9 and carcinoembryonic antigen assays in detecting cancer of the pancreas. Gastroenterology 1986;90:343–9. [PubMed: 2416628]

9. Guo M, Luo G, Lu R, et al. Distribution of Lewis and Secretor polymorphisms and corresponding CA19-9 antigen expression in a Chinese population. FEBS Open Bio 2017;7:1660–1671.

10. Liu C, Deng S, Jin K, et al. Lewis antigen-negative pancreatic cancer: An aggressive subgroup. Int J Oncol 2020;56:900–908. [PubMed: 32319567]

11. Kim S, Park BK, Seo JH, et al. Carbohydrate antigen 19–9 elevation without evidence of malignant or pancreatobiliary diseases. Sci Rep 2020;10:8820. [PubMed: 32483216]

12. Ballehaninna UK, Chamberlain RS. The clinical utility of serum CA 19–9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal. J Gastrointest Oncol 2012;3:105–19. [PubMed: 22811878]

13. Ballehaninna UK, Chamberlain RS. Serum CA 19–9 as a Biomarker for Pancreatic Cancer-A Comprehensive Review. Indian J Surg Oncol 2011;2:88–100. [PubMed: 22693400]

14. Poruk KE, Gay DZ, Brown K, et al. The clinical utility of CA 19–9 in pancreatic adenocarcinoma: diagnostic and prognostic updates. Curr Mol Med 2013;13:340–51. [PubMed: 23331006]

15. Cohen JD, Javed AA, Thoburn C, et al. Combined circulating tumor DNA and protein biomarker-based liquid biopsy for the earlier detection of pancreatic cancers. Proc Natl Acad Sci U S A 2017;114:10202–10207. [PubMed: 28874546]

16. Cohen JD, Li L, Wang Y, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science 2018;359:926–930. [PubMed: 29348365]

17. Fahrmann JF, Bantis LE, Capello M, et al. A Plasma-Derived Protein-Metabolite Multiplexed Panel for Early-Stage Pancreatic Cancer. J Natl Cancer Inst 2018.

18. Capello M, Bantis LE, Scelo G, et al. Sequential Validation of Blood-Based Protein Biomarker Candidates for Early-Stage Pancreatic Cancer. J Natl Cancer Inst 2017;109.

19. Fahrmann JF, Mao X, Irajizad E, et al. Plasma-Derived Extracellular Vesicles Convey Protein Signatures that Reflect Pathophysiology in Lung and Pancreatic Adenocarcinomas. Cancers (Basel) 2020;12.

20. Kobayashi M, Katayama H, Irajizad E, et al. Proteome Profiling Uncovers an Autoimmune Response Signature That Reflects Ovarian Cancer Pathogenesis. Cancers (Basel) 2020;12.

21. Jacobs EJ, Chanock SJ, Fuchs CS, et al. Family history of cancer and risk of pancreatic cancer: a pooled analysis from the Pancreatic Cancer Cohort Consortium (PanScan). Int J Cancer 2010;127:1421–8. [PubMed: 20049842]

22. Kirkegård J, Mortensen FV, Cronin-Fenton D. Chronic Pancreatitis and Pancreatic Cancer Risk: A Systematic Review and Meta-analysis. Am J Gastroenterol 2017;112:1366–1372. [PubMed: 28762376]

23. Canto MI, Almario JA, Schulick RD, et al. Risk of Neoplastic Progression in Individuals at High Risk for Pancreatic Cancer Undergoing Long-term Surveillance. Gastroenterology 2018;155:740–751.e2. [PubMed: 29803839]

24. Sharma A, Kandlakunta H, Nagpal SJS, et al. Model to Determine Risk of Pancreatic Cancer in Patients With New-Onset Diabetes. Gastroenterology 2018;155:730–739.e3. [PubMed: 29775599]

25. Aslanian HR, Lee JH, Canto MI. AGA Clinical Practice Update on Pancreas Cancer Screening in High-Risk Individuals: Expert Review. Gastroenterology 2020.

26. Kim J, Yuan C, Babic A, et al. Genetic and Circulating Biomarker Data Improve Risk Prediction for Pancreatic Cancer in the General Population. Cancer Epidemiol Biomarkers Prev 2020;29:999–1008. [PubMed: 32321713]

27. O'Brien DP, Sandanayake NS, Jenkinson C, et al. Serum CA19-9 is significantly upregulated up to 2 years before diagnosis with pancreatic cancer: implications for early disease detection. Clin Cancer Res 2015;21:622–31. [PubMed: 24938522]

28. Honda K, Katzke VA, Hüsing A, et al. CA19-9 and apolipoprotein-A2 isoforms as detection markers for pancreatic cancer: a prospective evaluation. Int J Cancer 2019;144:1877–1887. [PubMed: 30259989]

29. Nolen BM, Brand RE, Prosser D, et al. Prediagnostic serum biomarkers as early detection tools for pancreatic cancer in a large prospective cohort study. PLoS One 2014;9:e94928. [PubMed: 24747429]

30. Prorok PC, Andriole GL, Bresalier RS, et al. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. Control Clin Trials 2000;21:273s–309s. [PubMed: 11189684]

31. Fahrmann JF, Bantis LE, Capello M, et al. A Plasma-Derived Protein-Metabolite Multiplexed Panel for Early-Stage Pancreatic Cancer. J Natl Cancer Inst 2019;111:372–379. [PubMed: 30137376]

32. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45. [PubMed: 3203132]

33. Schlesselman JJ. Case–Control Studies. Oxford: Oxford University Press 1982.

34. Kovalchik SA, Tammemagi M, Berg CD, et al. Targeting of low-dose CT screening according to the risk of lung-cancer death. N Engl J Med 2013;369:245–254. [PubMed: 23863051]

35. Breslow NE, Nicholas E. Day, and Elisabeth Heseltine. Statistical methods in cancer research. 1980;1.

36. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394–424. [PubMed: 30207593]

37. Rawla P, Sunkara T, Gaduputi V. Epidemiology of Pancreatic Cancer: Global Trends, Etiology and Risk Factors. World J Oncol 2019;10:10–27. [PubMed: 30834048]

38. Petersen GM. Familial pancreatic cancer. Semin Oncol 2016;43:548–553. [PubMed: 27899186]

39. Ohno E, Hirooka Y, Kawashima H, et al. Natural history of pancreatic cystic lesions: A multicenter prospective observational study for evaluating the risk of pancreatic cancer. J Gastroenterol Hepatol 2018;33:320–328. [PubMed: 28872701]

40. Pannala R, Basu A, Petersen GM, et al. New-onset diabetes: a potential clue to the early diagnosis of pancreatic cancer. Lancet Oncol 2009;10:88–95. [PubMed: 19111249]

41. de Jong K, Nio CY, Hermans JJ, et al. High prevalence of pancreatic cysts detected by screening magnetic resonance imaging examinations. Clin Gastroenterol Hepatol 2010;8:806–11. [PubMed: 20621679]

42. Kromrey ML, Bülow R, Hübner J, et al. Prospective study on the incidence, prevalence and 5-year pancreatic-related mortality of pancreatic cysts in a population-based study. Gut 2018;67:138–145. [PubMed: 28877981]

43. Soroida Y, Sato M, Hikita H, et al. Pancreatic cysts in general population on ultrasonography: Prevalence and development of risk score. Journal of Gastroenterology 2016;51:1133–1140. [PubMed: 26988361]

44. Lee KS, Sekhar A, Rofsky NM, et al. Prevalence of incidental pancreatic cysts in the adult population on MR imaging. Am J Gastroenterol 2010;105:2079–84. [PubMed: 20354507]

45. Laffan TA, Horton KM, Klein AP, et al. Prevalence of unsuspected pancreatic cysts on MDCT. AJR Am J Roentgenol 2008;191:802–7. [PubMed: 18716113]

46. Maitra A, Fukushima N, Takaori K, et al. Precursors to invasive pancreatic cancer. Adv Anat Pathol 2005;12:81–91. [PubMed: 15731576]

47. Choi SH, Park SH, Kim KW, et al. Progression of Unresected Intraductal Papillary Mucinous Neoplasms of the Pancreas to Cancer: A Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol 2017;15:1509–1520.e4. [PubMed: 28342950]

48. Tanaka M, Fernandez-del Castillo C, Adsay V, et al. International consensus guidelines 2012 for the management of IPMN and MCN of the pancreas. Pancreatology 2012;12:183–97. [PubMed: 22687371]

49. Tang RS, Weinberg B, Dawson DW, et al. Evaluation of the guidelines for management of pancreatic branch-duct intraductal papillary mucinous neoplasm. Clin Gastroenterol Hepatol 2008;6:815–9; quiz 719. [PubMed: 18602036]

50. Jang JY, Park T, Lee S, et al. Validation of international consensus guidelines for the resection of branch duct-type intraductal papillary mucinous neoplasms. Br J Surg 2014;101:686–92. [PubMed: 24668442]

51. Khoury RE, Kabir C, Maker VK, et al. What is the Incidence of Malignancy in Resected Intraductal Papillary Mucinous Neoplasms? An Analysis of Over 100 US Institutions in a Single Year. Ann Surg Oncol 2018;25:1746–1751. [PubMed: 29560572]

52. Lennon AM, Buchanan AH, Kinde I, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. Science 2020;369:eabb9601. [PubMed: 32345712]

53. Liu MC, Oxnard GR, Klein EA, et al. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Annals of Oncology 2020;31:745–759. [PubMed: 33506766]

54. Bhat K, Wang F, Ma Q, et al. Advances in biomarker research for pancreatic cancer. Curr Pharm Des 2012;18:2439–51. [PubMed: 22372502]

What You Need to Know

BACKGROUND AND CONTEXT:

Several lines of evidence clearly indicate that survival rates for pancreatic cancer can be improved through identification of disease at an earlier stage. Herein, we assessed the lead-time trajectory of CA19-9 for early detection of pancreatic cancer using pre-diagnostic sera from the PLCO Cancer Screening Cohort. We also evaluated the potential of two additional previously validated protein markers, LRG1 and TIMP1, for their complementary performance with CA19-9 in the pre-diagnostic setting.
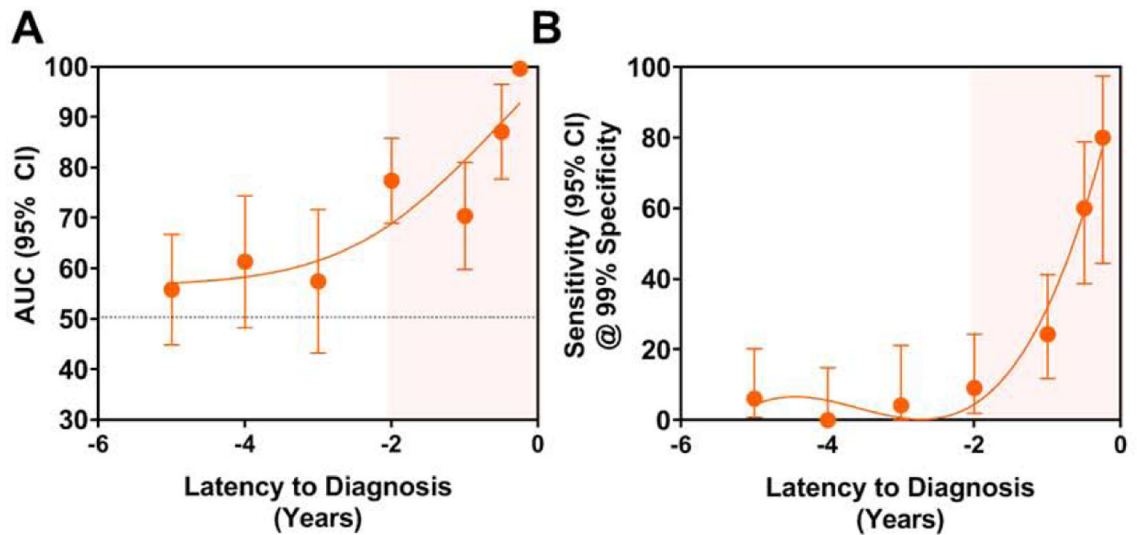
NEW FINDINGS:

Levels of CA19-9 increased markedly starting at two years prior to diagnosis with sensitivities reaching 50% at 99% specificity within 0–6 months prior to diagnosis for cases diagnosed with early-stage disease. For pre-diagnostic cases below the cut-off value for CA19-9, we further demonstrate that LRG1 and TIMP1 together are capable of complementing CA19-9 to enable identification of additional cases missed by CA19-9 alone.

LIMITATIONS:

Lack of serial samples did not allow for development of longitudinal algorithms and assessment of significance of incremental increases in CA19-9 levels. Limited data on benign conditions such as CP, pancreatic cystic lesions, new onset diabetes or obstructive jaundice were lacking in the PLCO pre-diagnostic cohort.

IMPACT:

CA19-9 can serve as an anchor marker for pancreatic cancer early detection applications. Inclusion of additional markers such as LRG1 and TIMP1 may have value for identifying cases that do not meet the CA19-9 cutoff.

Figure 1. Time-dependent classifier performance of CA19-9 in the PLCO Cohort.
A) AUC point estimates (95% CI) of CA19-9 for distinguishing cases stratified by time to diagnosis from baseline blood draw from matched controls. **B)** Sensitivity (95% CI) of CA19-9 at 99% specificity at various lead times. A 4th order fitted spline curve was used to illustrate the trajectory of the classifier performance or sensitivity at 99% specificity of CA19-9 in relation to time to diagnosis from baseline blood draw. Tabulated values are shown in the table beneath the figures.
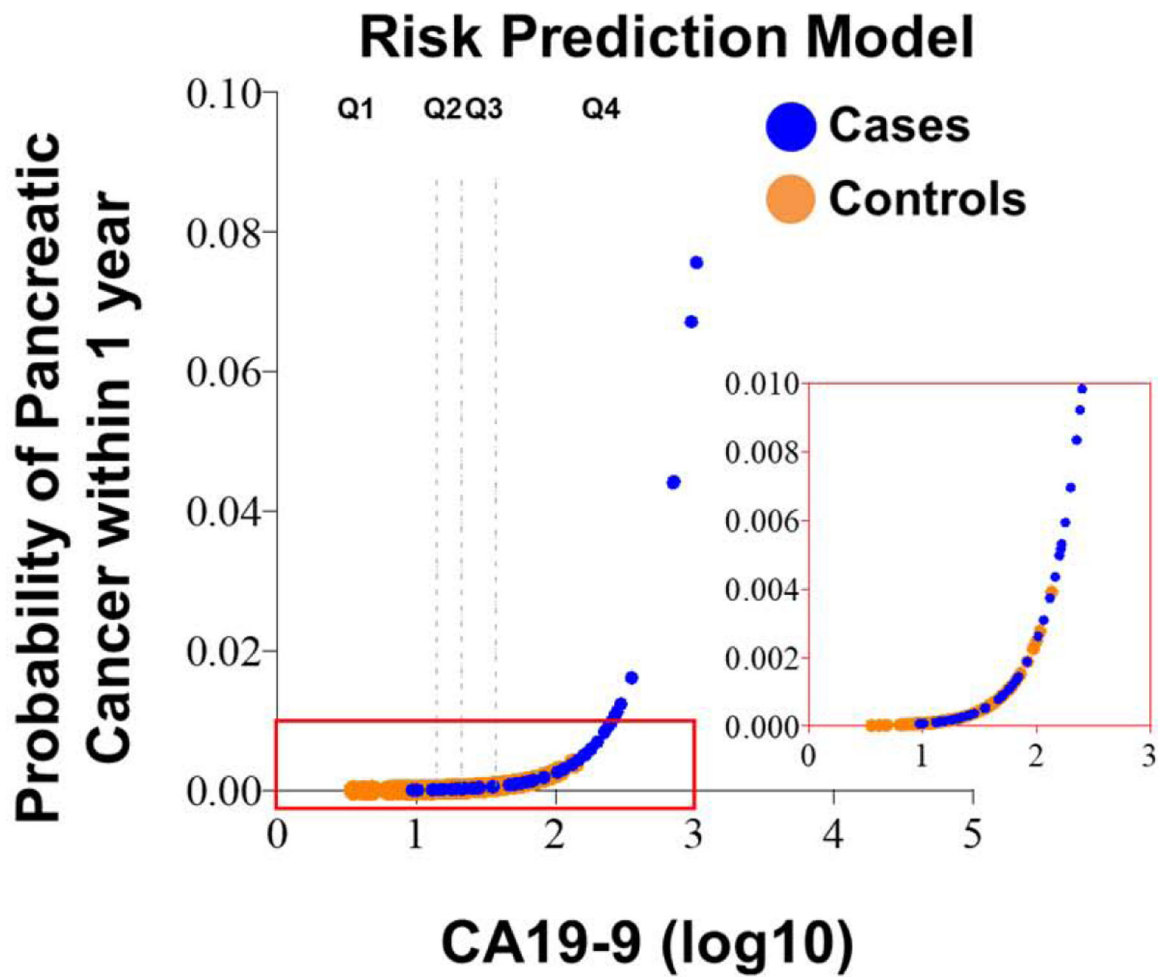
Time to Diagnosis (years)	4+	3-4	2-3	1-2	0.5-1	0-0.5	0-0.25
# of Cases	33	23	24	33	37	25	10
# of Controls	165	115	120	165	185	125	50
AUC (95% CI)	0.56 (0.45-0.67)	0.61 (0.48-0.74)	0.57 (0.43-0.72)	0.77 (0.69-0.86)	0.70 (0.60-0.81)	0.87 (0.78-0.97)	0.996 (0.99-1.00)
Sens @99% Spec (95% CI)	6.06 (0.74-20.23)	0 (0.00-14.82)	4.17 (0.11-21.12)	9.09 (1.92-24.33)	24.32 (11.77-41.20)	60.0 (38.67-78.87)	80.0 (44.39-97.48)
Sens @95% Spec (95% CI)	15.15 (5.11-31.9)	8.7 (1.07-28.04)	12.5 (2.66-32.36)	27.27 (13.3-45.52)	32.43 (18.01-49.79)	72.0 (50.61-87.93)	100.0 (69.15-100.00)

Figure 2. Predicted probability of pancreatic cancer within 1 year according to CA19-9 values.
The rug plot shows the observed distribution of biomarker scores. The vertical broken lines correspond to the quartiles threshold for CA19-9 values amongst controls (Q1, Q2, Q3, and Q4).
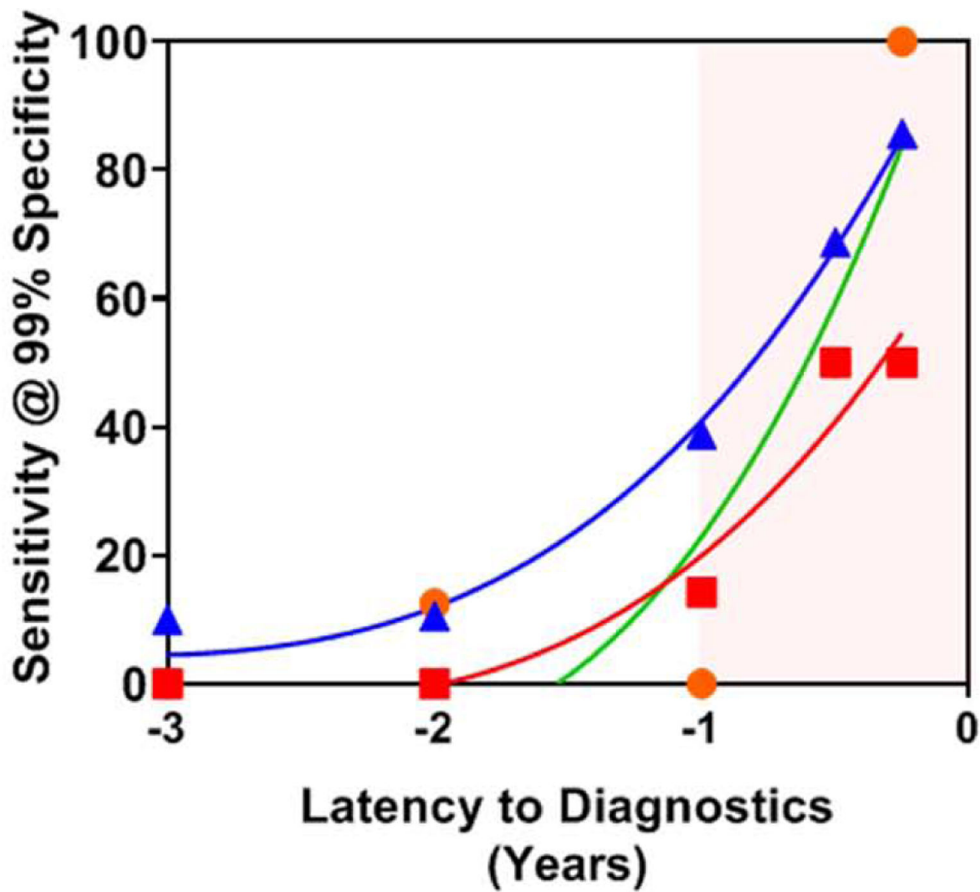
Figure 3. Time-dependent classifier performance of CA19-9 for cases stratified based on presentation of localized, regional or distant disease at time of diagnosis and that were diagnosed within 3 years of baseline blood draw in the PLCO Cohort.
Sensitivity of CA19-9 at 99% specificity for distinguishing cases stratified based on presentation of localized, regional or distant disease at time of diagnosis from time-interval matched controls. A 4th order fitted spline curve was used to illustrate the trajectory of the sensitivity at 99% specificity of CA19-9 in relation to time to diagnosis from baseline blood draw.
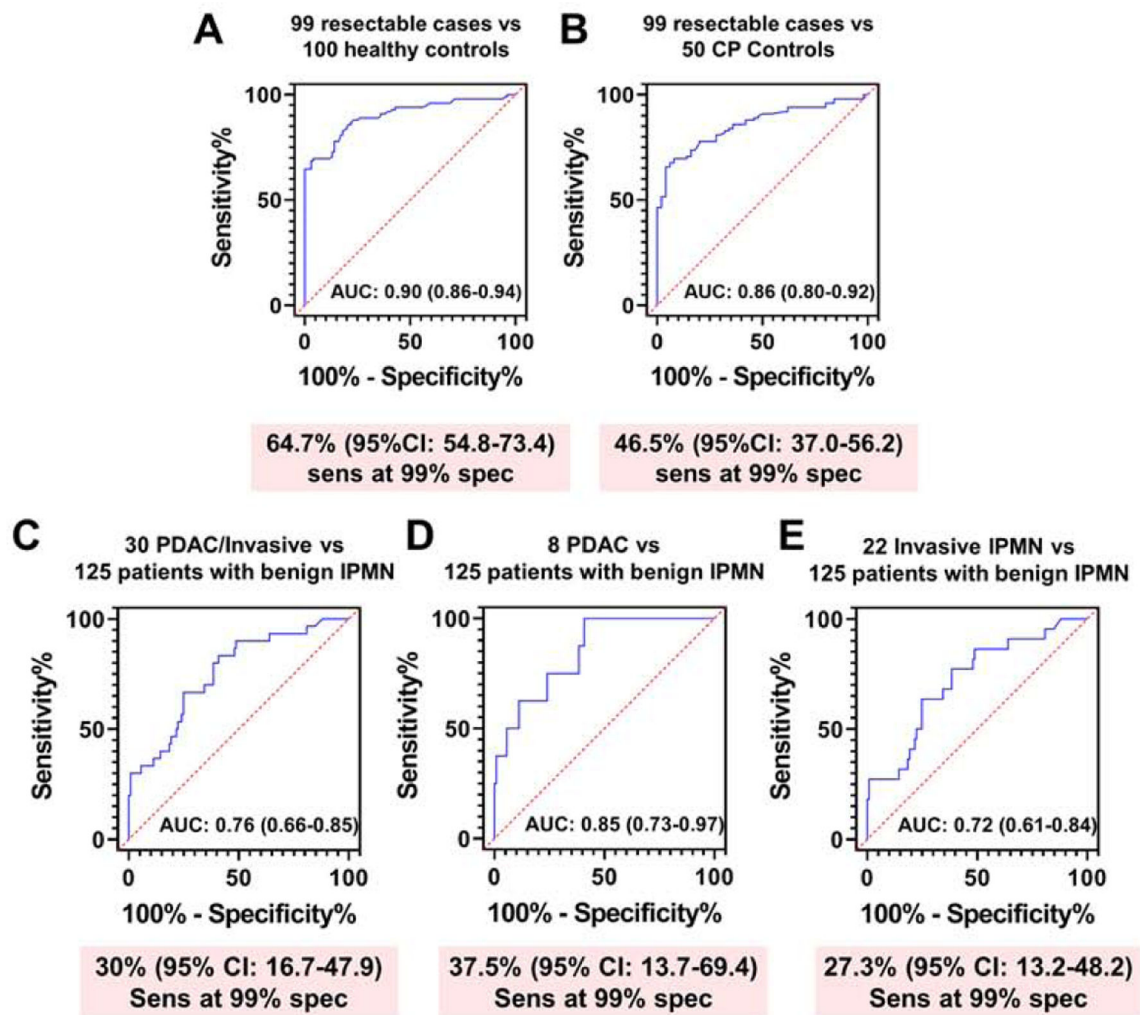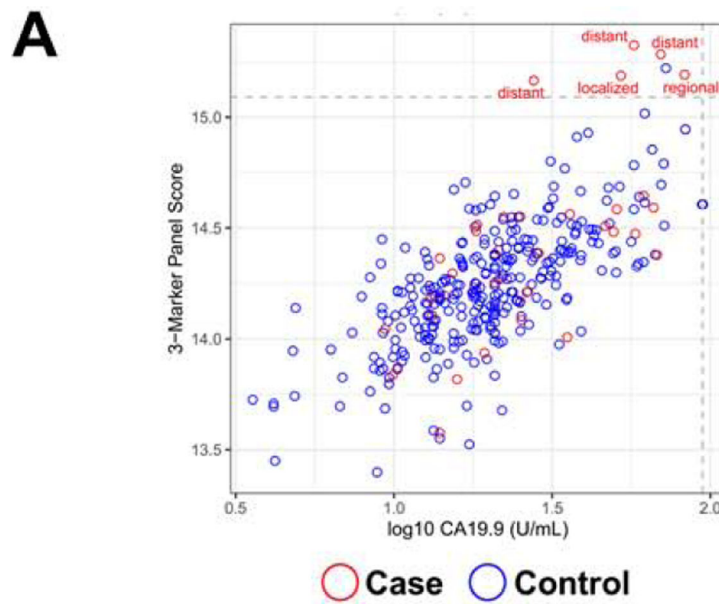
Figure 4. Classifier performance of CA19-9 in Test Set #1 and Test Set #2.
A-B) AUC (95% CI) and sensitivity (95% CI) at 99% specificity of CA19-9 for distinguishing resectable PDAC cases (n=99) from matched healthy controls (n=100) (**A**) or subjects with chronic pancreatitis (n=50) (**B**). **C)** AUC (95% CI) and sensitivity at 99% specificity of CA19-9 for distinguishing resectable PDAC cases (8 PDAC and 22 IPMN with an associated invasive ductal adenocarcinoma) from subjects harboring benign IPMN (n=125). **D-E)** AUC (95% CI) and sensitivity (95% CI) at 99% specificity of CA19-9 for distinguishing PDAC (n=8) (**D**) or Invasive IPMN (n=22) (**E**) from subjects harboring benign IPMN (n=125).

A) Scatter plot illustrating the distribution of the 3-marker panel scores (Y-axis) and log10 CA19-9 values (X-axis). Broken lines represent >99% specificity cutoffs. The 3-marker panel was derived using fixed beta-coefficients from the logistic regression model previously developed elsewhere[18]. **B)** Confusion matrix describing the performance of the classification model corresponding to the 3-marker panel and CA19-9 alone at >99% specificity. Statistical significance was determined by 1-sided McNemar exact test.

Figure 5. Classifier performance of a 3-marker panel consisting of LRG1, TIMP1 and CA19-9 for identifying cases diagnosed within 1 year and that were 'negative' for CA19-9 alone based on a 99% specificity cutoff.

Table 1.

PLCO Patient Characteristics

	Case/Control Status			
	Non-Case		Case	
	N	%	N	%
Total	875	100	175	100
Gender				
Female	365	41.71	73	41.71
Male	510	58.29	102	58.29
Age At Randomization				
<= 59	185	21.14	37	21.14
60–64	210	24	42	24
65–69	325	37.14	65	37.14
>= 70	155	17.71	31	17.71
Race				
White	795	90.86	159	90.86
Black	30	3.43	6	3.43
Other	50	5.71	10	5.71
Cigarette Smoking Status				
Never Smoked Cigarettes	426	48.69	65	37.14
Current Cigarette Smoker	74	8.46	36	20.57
Former Cigarette Smoker	375	42.86	74	42.29
BMI at Baseline (In kg/m2)				
Not Answered	7	0.8	0	0
0–18.5	8	0.91	3	1.71
18.5–25	306	34.97	57	32.57
25–30	368	42.06	72	41.14
30+	186	21.26	43	24.57
Diabetic Status				
Unknown	1	0.11	-	-
Yes	55	6.292	23	13.14
No	819	96.6	152	86.86
SEER Staging (cases only)				
Unknown			9	5.14
Localized			35	20
Regional			33	18.86
Distant			92	52.57