Detection of early pancreatic ductal adenocarcinoma using thrombospondin-2 and CA19-9 blood markers

Jungsun Kim[1], **William R. Bamlet**[2], **Ann L. Oberg**[2], **Kari G. Chaffee**[2], **Greg Donahue**[1], **Xing-Jun Cao**[3], **Suresh Chari**[4], **Benjamin A. Garcia**[3], **Gloria M. Petersen**[5,*], and **Kenneth S. Zaret**[1,*]

[1]Institute for Regenerative Medicine, Department of Cell and Developmental Biology, Abramson Cancer Center (Tumor Biology Program), Perelman School of Medicine, University of Pennsylvania, 9-131 SCTR, 3400 Civic Center Blvd., Philadelphia, PA 19104-5157

[2]Department of Health Sciences Research, Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota 55905

[3]Epigenetics Program, Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[4]Department of Medicine, Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, Minnesota 55905

[5]Department of Health Sciences Research, Division of Epidemiology, Mayo Clinic, Rochester, Minnesota 55905

Abstract

Markers are needed to facilitate early detection of pancreatic ductal adenocarcinoma (PDAC), which is often diagnosed too late for effective therapy. Starting with a PDAC cell reprogramming model that recapitulated the progression of human PDAC, we identified secreted proteins and tested and validated a subset of them as potential markers of PDAC. We optimized an ELISA assay using plasma samples from patients with various stages of PDAC, from individuals with benign pancreatic disease, and from healthy controls. Clinical studies including a phase 1 discovery study (N=20 patients), a phase 2a validation study (N=189), and a second phase 2b validation study (N=537) revealed that concentrations of plasma thrombospondin-2 (THBS2) discriminated among all stages of PDAC consistently over the three studies with a Receiver Operating Characteristic (ROC) c-statistic of 0.76 in Phase 1, 0.842 in Phase 2a, and 0.875 in Phase 2b. The concentration of THBS2 in plasma performed as well at discriminating resectable stage I cancer as stage III/IV PDAC. THBS2 concentrations combined with those for CA19-9, a

*Corresponding author: Dr. Ken Zaret, 215-573-5813, zaret@upenn.edu.

previously identified PDAC marker, yielded a c-statistic of 0.956 in the Phase 2a study and 0.970 in the Phase 2b study. THBS2 data improved the ability of CA19-9 to distinguish PDAC from pancreatitis. With a specificity of 98%, the combination of THBS2 and CA19-9 yielded a sensitivity of 87% for PDAC in the Phase 2b study. Given this, a THBS2 and CA19-9 panel assessed in human blood using a conventional ELISA assay may improve the detection of patients at high risk for PDAC.

INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) is projected to become the second leading cause of cancer death in the United States by 2020 (1). The majority of PDAC patients are diagnosed at an advanced stage of disease and their tumors are not surgically resectable, contributing to an overall 5-year survival rate of 7% (2). The lack of early diagnostics has made it challenging to develop therapeutics to slow or reverse PDAC (3). The CA19-9 serum marker is used to assess disease progression in PDAC patients (4, 5), but is not recommended for general screening (5, 6) because it is elevated in non-malignant pancreatic conditions such as chronic pancreatitis (7) and can produce false negatives in individuals who do not express Lewis blood group antigens (8). Other secreted markers have been reported for PDAC (9–12) including blood or urine proteins (13–15), exosomes (11), miRNAs (16), and epigenetic marks in circulating nucleosomes (17). However, challenges include lack of translation to the clinic, small sample sizes precluding statistical robustness, lack of blinded design, or inappropriate construction of datasets for development-to-validation (15–19). Most biomarkers were discovered in advanced PDAC or cell lines that are not representative of earlier stages, when detection would be most relevant, although recent candidates have been tested or discovered in pre-diagnostic samples of PDAC (20–22). When agnostic biomarker panels are assessed in validation samples, the need to aggregate samples from multiple sources hampers achieving statistical power (23).

We reasoned that proteins released from progressing precursor lesions, such as pancreatic intraepithelial neoplasias (PanIN) at stages 2 and 3 (e.g., from PanIN2 to PanIN3) (24) to PDAC, might provide an innovative and effective opportunity for discovering diagnostic biomarkers. We previously reprogrammed recurrent, advanced human PDAC cells into an induced pluripotent stem (iPSC) cell-like line (25). The iPSC-like line (designated as 10–22) can be propagated indefinitely, yet preferentially generates PanIN2/3 ductal lesions after growing for 3 months as teratomas in immunodeficient mice. The lesions progress to invasive PDAC by 6–9 months. Proteomic analysis of conditioned medium from 10–22 cell-derived precursor PanINs cultured as organoids, compared to medium from 10–22 cells grown under pluripotency conditions, revealed 107 secreted human proteins specific to the PanIN2/3 tumors (25). Of these, 43 proteins fell into interconnected TGFβ and integrin networks for PDAC progression (26, 27) and 25 proteins were within a network for the transcription factor HNF4α, which also dynamically showed an increase in expression. (25).

Here, we report an analysis of proteins secreted or released from the 10–22 cell-derived PanIN organoids using a phased cancer marker development design that incorporated criteria for prospective specimen collection and retrospective blinded evaluation (PRoBE) (28, 29).

RESULTS

Discovery of candidate biomarkers

Of the 107 proteins secreted and released selectively by human PanIN organoids (25), we focused on 53 proteins with a low abundance (2 nmol) in the healthy human plasma proteome and RNA-Seq databases (30–32) (Table S1). Enzyme-linked immunosorbent assays (ELISA) from validated sources (33) were not available for most of these rarely expressed proteins. Of the proteins for which reliable ELISA kits were available and that were not implicated as markers in other diseases, we focused on MMP2, MMP10, and thrombospondin-2 (THBS2) because they occur in integrated networks for TGF-β and integrin signaling, which drive PDAC progression (25). We investigated these three candidate markers in a screen of human plasma samples. All procedures were performed using a recommended biomarker phased design following the PRoBE criteria (28, 29). De-identified human plasma samples were provided by the Mayo Clinic pancreas research biospecimen repository. We then performed ELISA analyses blinded to disease status and then returned coded data to the Mayo Clinic team for statistical analysis and interpretation.

Phase 1 validation of candidate biomarkers

We examined whether MMP2, MMP10, or THBS2 could discriminate between cancer cases (n=10) and controls (n=10) with an area under the curve (AUC) analysis of the sensitivity and specificity of the markers. All cancer cases for the Phase 1 study were selected to have CA19-9 concentrations above 55 U/mL. MMP2 was unable to discriminate effectively between cancer cases and controls, and MMP10 signals were undetectable in all plasma samples (Fig. 1A). By contrast, THBS2 exhibited a c-statistic of 0.76 considering all cases versus controls (n=10) and a c-statistic of 0.886 when considering resectable and locally advanced PDAC (n=7). While human THBS2 has 80% amino acid sequence homology with THBS1, we demonstrated the specificity of each of the reagents in the THBS2 ELISA assay (Figures S1 and S2).

After the Phase 1 validation analysis, we performed a mass spectrometry study of the pooled cancer plasma samples (n=10) and the pooled control plasma samples (n=10) after the plasma samples were individually depleted of the 14 most abundant plasma proteins (e.g., serum albumin). At 5% false discovery rate (FDR), four unique peptides for THBS2 were identified, of which two were from sequences specific to THBS2 and the other two were from sequences that are conserved between THBS1 and THBS2. One of two peptides specific to THBS2 was present at a 3-fold greater concentration in the cancer plasma pool compared to the control plasma pool, and another peptide specific to THBS2 was detected only in the cancer pool and not the control pool (Tables S2A,S2B). At 1% FDR, a THBS2-specific peptide was detectable only in the cancer pool (Table S2C). Computational analysis of RNA expression data deposited by The Cancer Genome Atlas (TCGA) project (https://cancergenome.nih.gov/) shows that, of all cancers tested, PDAC (n=134) was second only to mesothelioma for expression of *THBS2* mRNA (Fig. 1B, compare medians denoted by vertical black bars within the red boxes). Taking together the Phase 1 validation study by ELISA, mass spectrometry data, and TCGA RNA-seq data, we concluded that THBS2 merited further study.

Phase 2a validation of THBS2 and CA19-9

Further validation of THBS2 involved human plasma samples in a Phase 2a study (Table 1) that contained CA19-9 negative and positive cases. The median ELISA value for THBS2 at all PDAC stages (N=81) in the Phase 2a group, 29.7 ng/ml, was 12.2 ng/ml higher than observed in controls (N=80) (Figure 2A), consistent with the mass spectrometry data for the Phase 1 study. THBS2 exhibited a c-statistic of 0.842 for all PDAC samples compared to controls (n=161, Table 2 all stages). In the same sample set, CA19-9 had a comparable c-statistic of 0.846 for all PDAC samples compared to controls (Table 2 all stages).

The data for the THBS2 analyses were reproducible across three different lot numbers of ELISA kits tested on the same subset of Phase 2a human plasma samples over a two-year period, with an average 10% coefficient of variation (CV) across the samples (Figure S3). The samples included 4 plasma samples that were re-frozen and thawed twice, and 3 plasma samples that were re-frozen and thawed three times. We concluded that the THBS2 assay was robust from the point of view of differences in plasma sample handling and assessment.

To determine whether CA19-9 and THBS2 together could constitute a more discriminatory panel than either marker alone, we performed logistic regression to estimate the combined probability of their discriminatory ability. A combination of CA19-9 and THBS2 for all cases versus controls with the Phase 2a data yielded a c-statistic of 0.956 (95% CI 0.93, 0.98) (Figure 2B, all stages), indicating the utility of the two-marker panel.

Phase 2b validation studies for CA19-9 and THBS2 as a combined marker for PDAC

We performed an independent Phase 2b validation study (see Table 1 for specimens) with an increased sample size. We accomplished temporal validation (18) as the Phase 2b analysis was conducted more than one year after the Phase 2a study. The distribution of THBS2 values across the Phase 2a and 2b studies is shown in Figure S4 and the range and median values of THBS2 and CA19-9 are shown in Table S3.

The c-statistics for CA19-9 and THBS2 alone, 0.881 and 0.875, respectively, were slightly better with the larger sample size of the Phase 2b study (n=337), compared to the Phase 2a study (n=161), and the combination of the two markers yielded a c-statistic of 0.970 (95% CI=0.96, 0.98) (Table 2, all stages; Fig. 2C, D). With regard to the distribution variability, the 75th percentile of the control values fell below the 25th percentile of the case values. Furthermore, the 95th percentile of the controls fell below the median measure observed in the case samples. The fact that 50% of the case values exceeded 95% of the control values was likely driving the AUC we observed for THBS2 with regard to being able to discern between cases and controls. We compared individual and combined marker performance in the Phase 2a and 2b studies at the stages of resectable PDAC (stages I, II) and locally advanced and metastatic PDAC (stages III, IV). Notably, the combination panel of CA19-9 and THBS2 performed well across all stages of PDAC (Table 2).

More detailed analysis of the distribution of ELISA signals provided insight into how the combination of CA19-9 and THBS2 performed. As observed in the scatter plots in Figures 2E and 2F, various cases (red +) had essentially zero CA19-9 signal (i.e., along the bottom of the plot), consistent with the likelihood that they were from PDAC patients who were

Lewis antigen negative; however, many of these cases had elevated THBS2 concentrations. Similarly, several cases exhibited THBS2 concentrations that overlapped with the upper range of the group of controls, and these cases exhibited high CA19-9. Thus, the two markers appeared complementary in their ability to detect PDAC.

While stages I, IIA, and IIB are classified as resectable tumors in the 6th Edition of the American Joint Committee on Cancer (AJCC) Pancreatic Cancer Staging System (34), only stages I and IIA are considered to be early. Therefore, we directly compared the AUCs for combinations of stages: I+IIA+IIB+II (unspecified), I+IIA+II (unspecified), and I+IIA in our phase 2a and 2b study. The AUC and 95% CI values were comparable for the two-marker combination across these three subsets, indicating that the exclusion of the questionable early stage IIB samples had limited impact on marker performance (Table S4).

In Tables S5A and S5B, we evaluated the relationship between THBS2 plasma values and age, sex, and presence of diabetes mellitus in the cohort. We observed no apparent associations between these parameters for any of the diagnosis groups of PDAC Adenocarcinoma stages I/II, Adenocarcinoma stages III/IV, pancreatitis, intraepithelial pancreatic mucinous neoplasm, insulinoma (islet cell), and healthy controls. Given the overall lack of association, we did not include any of these factors as adjustor variables in subsequent modeling.

Establishing a provisional cutoff point for THBS2 for clinical use

To determine a THBS2 plasma concentration to use as a cutoff point for discriminating healthy controls from PDAC cases in the clinic, we first considered the distribution of THBS2 values based upon the 230 healthy controls from our combined Phase 1, 2a, and 2b studies. From this distribution, we chose six cutoffs that represented a range of approximate false positive rates from 0 to 5 percent. These cutoffs were then evaluated for their sensitivity in detecting PDAC in the Phase 2a and 2b samples. As seen in Table 3 for the Phase 2b study, a concentration of THBS2 at or above 42 ng/ml detected about half of the PDAC cases (sensitivity) with 99% specificity. Combining the conventional CA19-9 cutoff of 55U/ml and a cutoff of 42 ng/ml for THBS2 in the Phase 2b samples, we observed 98% specificity and 87% sensitivity (Table 3).

Comparison of the THBS2/CA19-9 panel against other benign pancreatic conditions

When considering all PDAC cases versus chronic pancreatitis (Phase 2a, n=109; Phase 2b, n=252), c-statistics including CA19-9 increased from 0.774 or 0.816 (alone) to 0.842 or 0.867 (with THBS2) with the Phase 2a or Phase 2b data, respectively (Table 4, Figure 3A). The THBS2/CA19-9 panel was able to discriminate all PDAC cases tested (stages I–IV) versus intraductal papillary mucinous neoplasms (N=312) with a c-statistic of 0.952 (Table 4, Figure, B). Thus, the THBS2/CA19-9 panel was able to distinguish PDAC from intraductal papillary mucinous neoplasms and helped to distinguish PDAC from pancreatitis compared to CA19-9 alone.

THBS2 lacked the ability to discriminate between all PDAC cases and pancreatic neuroendocrine tumors and hindered, rather than enhanced, the c-statistic of CA19-9 (Table 4, Figure 3C). Considering a lack of markers available for pancreatic neuroendocrine tumors

and the poor performance of THBS2 in the discrimination of PDAC from pancreatic neuroendocrine tumors, we examined whether THBS2 could discriminate pancreatic neuroendocrine tumors (N=30) from healthy normal controls (N=149). CA19-9 alone did not discriminate pancreatic neuroendocrine tumor samples, as previously reported (35). However, THBS2 alone could discriminate pancreatic neuroendocrine tumors from healthy normal controls with a c-statistic of 0.751 (Table 4, Figure 3D).

PDAC can result in obstructive jaundice that can confound plasma assays (20, 36). Of the 288 adenocarcinoma cases included in these studies, we retrieved clinical total serum bilirubin information for 279 cases (96.9%) (Table S6A). Of the 279 samples with this information, 70 (25.1%) were inferred to have obstructive jaundice, based on total bilirubin concentrations of 3.5 mg/dl. Slightly lower median CA19-9 concentrations (208.5 vs 220 U/ml) as well as elevated median THBS2 concentrations (56.4 vs 33.0 ng/ml) were observed in PDAC subjects with jaundice when compared to those without jaundice, indicating that obstructive jaundice influenced both CA19-9 and THBS2 concentrations. Yet 14 out 55 (25%) of the PDAC patients with normal CA19-9 and without jaundice had elevated THBS2 concentrations (42 ng/ml) (Table S6B). Also, 8 out of 13 (62%) patients with normal CA19-9 and with jaundice showed elevated THBS2 concentrations (42 ng/ml) (Table S6B). Therefore, the THBS2 concentration in plasma identified a subset of non-jaundiced adenocarcinoma cases with normal CA19-9 concentrations. Furthermore, stratifying the marker panel performance by overall PDAC or PDAC without jaundice, versus controls, in the Phase 2a and 2b studies affected the AUCs by less than 0.01, (Table S6C). Due to limited availability of benign biliary disease samples, we did not compare THBS2 and CA19-9 concentrations between benign biliary disease, non-jaundice PDAC, and jaundice PDAC.

Cross-validation of THBS2 measurements in different laboratories

Following these analyses, an independent biomarker development laboratory at the University of Pennsylvania tested a subset of the Phase 2b samples for THBS2 concentrations. Thirty-eight samples were randomly selected to cover the entire range of THBS2 concentrations, focusing on those around the cutoff value. The samples were de-identified and provided without communication other than the manufacturer's instructions for the ELISA assay and the methods section of this paper. The ELISA assays for THBS2 were performed over a year after the original study and with reagents with different lot numbers. The THBS2 concentrations in the original and cross-validated assays were highly concordant and yielded Pearson and Spearman correlation coefficients of 0.95 and 0.968, respectively (Fig. S5A).

We noticed that the THBS2 signals were slightly lower in the cross-validated data, including for the human normal control plasma samples used on each plate. The original studiesyielded an average value of 17 ng/ml for the normal control plasma samples, whereas the cross-validation study yielded a value of 13.25 for the normal control plasma samples. The lower overall values of unknowns caused 4 of the 38 samples that were just over the 42 ng/ml cutoff, to fall below the cutoff (Table S7A).

To accommodate for operational differences, we created a scalar where our original 42 ng/ml cutoff to detect cancer was divided by the original 17 ng/ml average, the normal

plasma control value, to yield a scalar cutoff of 2.47. We therefore divided the THBS2 result for each unknown in the cross-validation study by the value (13.25 ng/ml) of the normal control plasma, where the cutoff value would be 2.47 times the control value. Scaling did not affect the correlation coefficient (Figure S5B). With the data scaled in this fashion, two samples that were below the 42 ng/ml cutoff in the original samples were now above the cutoff in the cross-validation data (Table S7B). Thus, while the scaling method improved the outcome of the cross-validation assay, careful calibration was needed to ensure consistency in the assay results over time and with different batches of reagent, once a cutoff for clinical practice was determined.

Expression of THBS2 in different stages of human PDAC

We sought to determine the cells expressing THBS2 in a total of 42 cases of human PDAC and 4 cases of incidental pancreatic intraepithelial neoplasia (PanIN) and intraepithelial pancreatic mucinous neoplasms by immunohistochemistry. All 42 cases of PDAC and all 4 cases of incidental PanIN/intraepithelial pancreatic neoplasms exhibited detectable THBS2 (Figure 4; Figure S6, Table S8). Two different antibodies detected THBS2 in PanIN2 epithelia found incidentally in the PDAC tumor, but little THBS2 was found in PanIN1 epithelia (Figure 4A, B). Both antibodies also detected THBS2 in Stage II and Stage III PDAC, and a 10-fold excess of peptide specific to the second antibody blocked the antibody (Figure 4C–K). Epithelial cells, but not stromal cells, were predominantly labeled with THBS2 in PanIN/intraepithelial pancreatic mucinous neoplasm tissue (4 out of 4) (Figure 4A–B, Figure S6B–C). In PDAC tumor tissue, 32 cases were labeled with THBS2 in epithelial cells, 21 cases were labeled in both epithelial and stromal cells, and in 8 cases the staining was mostly in stromal cells of poorly differentiated PDAC tissue (Table S8).

DISCUSSION

With a 5-year survival rate for patients with stage I PDAC being at least four times that of overall survival rates for PDAC (34, 37), the THBS2/CA19-9 marker panel may help to detect early stage tumors that are resectable and should improve the prognosis of PDAC. The performance of THBS2 in early stage cancer may be a consequence of our discovery that it is secreted or released from live human precursor PanIN organoids (25), reflecting the value of the iPS-like PDAC cell line, 10–22 cells, to recapitulate human pancreatic cancer progression. While most PDAC patients are diagnosed at advanced stages, the time from the occurrence of the initiating mutation to the birth of PDAC founder cells (38) can be a decade, suggesting that there may be a time period to identify progressing disease before PDAC can be clinically imaged. PanINs have been identified in pancreas up to 10 years before the development of infiltrating PDAC (39), underscoring the importance of early diagnosis. However, we note that PanINs can also be observed in the absence of PDAC.

We propose that high specificity outweighs considerations of increased sensitivity because of heightened anxiety in patients over suspected pancreatic cancer plus the costs of subsequent diagnostic evaluation. We found that with a THBS2 concentration cutoff of 42 ng/ml, THBS2 could discriminate PDAC patients from healthy primary care controls with a specificity of 99% (1% FPR) and a sensitivity of 52%. Impressively, combining CA19-9

(>55 U/ml) with THBS2 (>42 ng/ml) showed a specificity of 98% and a sensitivity of 87% in our larger phase 2b study. An important strength of this study was the ability to use large defined plasma samples obtained from a single institution that followed standardized processing protocols (23).

Decreased thrombospondin-1 (THBS1) concentrations by mass spectroscopy analysis have been reported in plasma samples from PDAC patients and in samples obtained prior to the cancer diagnosis (20, 40). Jenkinson et al. found that reduced concentrations of THBS1 occurred in PDAC patients with diabetes, but not in PDAC patients without diabetes (20). In contrast, we observed that elevated concentrations of THBS2 were associated with PDAC but we did not find an association of elevated THBS2 concentrations with diabetes mellitus, age, or sex (Table S5A, S5B). THBS1 and THBS2 share 80% of their protein sequence, but have diverged in function and in their genetic regulation (41, 42). We showed that elevated THBS2 did not correspond to THBS1 in PDAC. First, the antibodies in the ELISA kit used for our study had negligible cross-reactivity with THBS1 (Figs. S1, S2). Second, we confirmed by mass spectrometry that the peptides specific to THBS2 were more abundant in cancer patient plasma than in plasma from normal healthy controls (Table S2). Thus, elevated THBS2 concentrations in PDAC were independent of THBS1 concentrations as reported in the literature.

THBS2 is a glycoprotein that may be an angiogenesis inhibitor, and mutation of the mouse *TSP-2* gene increases susceptibility to cancer (43). We found that THBS2 antigen is expressed in normal pancreas cells but the baseline concentration of THBS2 is very low in normal human plasma by both mass spectrometry and ELISA. We found that THBS2 antigen is robustly expressed in PDAC tumor tissue, perhaps concordant with the poor vascularization associated with PDAC, and is also increased in plasma of PDAC patients. Interestingly, THBS2 is downregulated in gastric cancer cells (44). Further work is needed to understand how the release of THBS2 into plasma is increased in the patients cohorts studied here.

A group of scientists has initiated the STARD (Standards for Reporting of Diagnostic Accuracy) with guidelines to improve the reporting of diagnostic accuracy (45). It will be useful to follow these standard guidelines in the clinic by reporting imprecision as the coefficient of variation (CV%) and precision as 95% confidence intervals near clinical decision points obtained by repeating the test over several independent days. Also, to reduce even small differences in the assay occurring between different laboratories, presenting the likelihood ratio with 95% confidence intervals along with specificity and sensitivity at several cut-off points is recommended. Our cross-validation study was an initial attempt to address these issues and more work is needed for a determination of clinical decision points with confidence.

The combination of THBS2 and CA19-9 improved the discrimination of patients with PDAC from those with chronic pancreatitis. Longitudinal studies are needed to determine if a subset of the pancreatitis patients who scored positive for THBS2, but were clinically assessed to be PDAC negative, indeed harbored early stage PDAC. Likewise, larger studies are needed to determine the effectiveness of THBS2 for diagnosing pancreatic

neuroendocrine tumors (PNET) where CA19-9 is not applicable, and other cancers showing high *THBS2* mRNA expression (Figure 1B). Further research with larger numbers of pancreatic cancer cases without jaundice as well as patients without cancer but with jaundice will be necessary to quantify the value of THBS2 and CA19-9 for detecting non-jaundice pancreatic cancer.

There are limitations to our study. The prevalence of PDAC in different populations affects the positive and negative predictive values for determining the utility of a biomarker in a population. The positive predictive value (PPV) is the probability that subjects with a positive screening test have the disease, and the negative predictive value (NPV) is the probability that subjects with a negative screening test do not have the disease. Given the low prevalence of 4 to 12.4 cases of pancreatic cancer per 100,000 in the general population (https://seer.cancer.gov/statfacts/html/pancreas.html), our marker panel with a combined 98% specificity and 87% sensitivity would have a PPV of 0.002, with an NPV of 1.0 (2). Yet, when viewed in terms of the 1.5% lifetime risk of PDAC in the general population (2), the PPV becomes about 0.4 with an NPV of 0.99. For patients older than 55 years who are newly diagnosed with diabetes (46), with a prevalence of 1% in the general population for PDAC, the PPV is 0.31 and the NPV is 1.0. For first degree relatives of PDAC patients and smokers in the general population, each group with a lifetime risk of 3.75% (47), the PPV is 0.63 and NPV is 0.99. For carriers with relevant germline mutations (in aggregate, including *BRCA1, BRCA2, CDKN2A, PALB2*), lifetime risk is 40% (47) and the PPV rises to 0.97 and NPV is 0.92. Based on these considerations, we suggest that the THBS2/CA19-9 marker panel could serve as a low cost, non-intervention screening tool in asymptomatic individuals who have a high risk of developing PDAC (3, 47, 48), or in patients who are newly diagnosed with diabetes mellitus that develops as a result of pancreatic injury (49), but not in the general population.

Another limitation of our work is that our histological analysis of THBS2 expression at different stages of pancreatic cancer was limited by the portion of tissue available from each of the resections. Furthermore, it is unclear how expression of THBS2 in cells under normal or pathological conditions may relate to the extent to which the protein is secreted or released into the plasma, and stable there. Also, further work is needed to refine clinical decision points for high-risk individuals, to determine the panel's utility for detecting earlier stage progression to PDAC, and to determine the specificity for pancreatic cancer versus cancers of other tissue types and other disease states.

Materials and Methods

Study Design and Populations

All procedures were performed using a recommended biomarker phased design following the PRoBE criteria (28, 29). De-identified human plasma samples from the Mayo Clinic pancreas research biospecimen repository were shipped to our laboratory, which performed ELISA analyses blinded to disease status, and then returned coded data to the Mayo Clinic team for statistical analysis and interpretation.

Collection of plasma samples was approved by the Mayo Clinic Institutional Review Board. Following rapid case finding (50) and informed consent, participants with PDAC provided venous blood samples prior to initiation of cancer therapy. Samples were frozen at −80°C until used. Similarly, blood samples were obtained from the Mayo Clinic through primary care (healthy controls) and gastroenterology clinics (participants diagnosed with chronic pancreatitis, intraductal papillary mucinous neoplasms, and pancreatic neuroendocrine tumors). An aliquot of serum was assayed for CA19-9 at the Mayo Clinic Immunochemical Core Laboratory as recommended by the ELISA kit manufacturer (Cobas/Roche). Demographic and clinical characteristics in each group are shown in Table 1.

Exploratory Set (Phase 1)—Plasma samples from 20 non-Hispanic Caucasian subjects recruited at Mayo Clinic included 10 healthy primary care controls and 10 [6 early stage (I/II), 4 late stage (III/IV)] patients with clinically or histologically proven PDAC. All cancer cases for Phase 1 were selected to have CA19-9 concentrations above 55 U/mL.

Validation Set (Phase 2a)—Plasma samples from 189 non-Hispanic Caucasian subjects recruited at Mayo Clinic included 81 (58 early stage, 23 late stage) patients with clinically or histologically proven PDAC, 80 healthy primary care controls, and 28 patients with a personal history of chronic pancreatitis; patients with hereditary pancreatitis were excluded given their increased risk for PDAC. The controls were matched to the cases by age and sex. Approximately 15% of the healthy controls self-reported a personal history of diabetes.

Validation Set (Phase 2b)—Plasma samples collected from 537 non-Hispanic Caucasian subjects recruited at Mayo Clinic included 197 (88 early stage, 109 late stage) patients with clinically or histologically proven PDAC, 140 healthy primary care controls, 115 patients with intraepithelial pancreatic mucinous neoplasm without PDAC, 30 patients with pancreatic neuroendocrine tumor (PNET), and 55 patients with a self-reported personal history of chronic pancreatitis; patients with hereditary pancreatitis were excluded. Approximately 11% of the controls self-reported a personal history of diabetes.

Measurement of markers in human plasma

After the ELISA assays for the Phase 1 study was completed, the remaining PDAC (n=10) and control samples (n=10) were each separately depleted of abundant serum proteins by filtration and then High Performance Liquid Chromatography (HPLC) using a Seppro IgY14 LC 10 column (Sigma Aldrich). The resulting 10 samples of cancer plasmas, depleted of abundant proteins, were pooled separately from a pool of the controls and the two pools were subjected to 2D Strong Cation Exchange Chromatography (SCX)/tandem mass spectrometry analysis as previously described (51). In brief, aSCX tip column was made with a 200 ul tip packed with 20 ul PolySULFOETHYL resin (Nest Group). The SCX tip was pre-washed with buffer B (500 mM KCl, 10mM NaH2PO4, 30% acetonitrile, pH 2.6), followed by equilibration with buffer A (10 mM NaH2PO4, 30% acetonitrile, pH 2.6). The lyophilized digested peptides (100 ug) were reconstituted in 50 ul buffer A. The reconstituted digested peptide solution was loaded into the SCX tip column twice, followed by washing with 50 ul buffer A. All flow-through fractions were combined ("flow-through"). The following 100 ul KCl concentration buffers, made by mixing the different

proportions of buffer A and B, were used to successively wash the column: 30 mM, 40 mM, 50 mM, 60 mM, 70 mM, 85 mM, 100 mM, 150 mM and 500 mM. A total of 10 fractions was dried and de-salted using the homemade C18 Stage Tips. About 3 ug digested peptides were injected into a 75 um I.D. X 25cm C18 column with a pulled tip. Easy nLC 1000 was run at 300 nl/min flow rate for 180 min gradient. Online nanospray was used to spray the separated peptides into an Orbitrap Fusion Tribrid mass spectrometer (Thermo Electron). The raw data were acquired with Xcalibur and pFind2.8 software was used to search human Uniprot database. A 5% false discovery rate for the protein spectrum measurement (PSM) was used initially to filter the peptide search results. Supplementary Table 2 shows the results for a total of four THBS2 peptides that were detected in the pooled cancer plasma samples. The pFind2.8 search engine revealed two unique peptides in each of the pooled plasmas; the cancer pool had a peptide specific THBS2 (VCNSPEPQYGGK) and a peptide shared between THBS1 and THBS2 (NALWHTGNTPGQVR) and the normal pool had a peptide specific to THBS2 (TRNMSACWQDGR) and a peptide shared between THBS1 and THBS2 (FYVVMWK) (Supplementary Table 2A). We quantified THBS2 levels in each of the pooled plasmas by measuring the area under the peptide signals based on mass and retention time in original ms 1 and ms2 windows. We then normalized to the total spectral counts in each sample (Supplementary Table 2A).

To more stringently assess THBS2 peptide levels, we searched with 1% and 5% FDR settings against the Uniprot Human database (89,796 entries in total) with the updated pFind3.0 search engine (52). Search parameters were set for a precursor mass tolerance of ±7 ppm, fragment mass tolerance of ±0.4 Da, trypsin cleaving after lysine and arginine with up to 2 miscleavages, carbamidomethyl [C]/+57.021 as the fixed modification, and acetyl [proteinN-term]/+42.011, deamidated [NQ]/+0.984, and xxidation [M]/+15.995 as the variable modifications. The target-decoy approach was used to filter the search results, in which the false discovery rate (FDR) was less than 1% or 5% at both the peptide and protein level. At both 1% and 5% FDR, the peptide specific to THBS2 (VCNSPEPQYGGK) and the peptide shared between THBS1 and THBS2 (NALWHTGNTPGQVR) was seen in the cancer pooled sample, consistent with the original pFind2.8 search (Supplementary Table 2B and 2C). However, at either 1% or 5% FDR with the pFind3.0 search, no peptides specific to THBS2 sequence were identified and only peptides shared between THBS1 and THBS2 were identified in the normal pooled sample (Supplementary Table 2B and 2C). Therefore, more stringent analysis verified THBS2 sequence in cancer pooled sample.

ELISA kits for human MMP2 (Millipore), human MMP10 (Ray Biotech), and human Thrombospondin-2 Quantikine (DTSP20, R&D systems) were used as described by the manufacturers' instructions. Duplicate 5 ul plasma samples were diluted 10 fold with calibrator diluent RD5P buffer and all 50 ul used for THBS2. Marker concentrations were determined from standard curves of positive control proteins from the kits with a 4 parameter logistic nonlinear regression model using SoftMax Pro Software (Molecular Device). Normal pooled human plasma (IPLA-N, Innovative Research) was tested in duplicate on each ELISA plate. Across 15 independent ELISA plate assays, THBS2 in duplicate control samples of commercial normal pooled human plasma ranged between 15 and 21 ng/ml, with a coefficient of variation of 13%. Also, the inclusion (or exclusion) of

occasional plasma samples that were orange or reddish in color, indicating hemolysis, had a negligible impact on the data.

RNA-Seq analysis from The Cancer Genome Atlas (TCGA)

THBS2 mRNA amounts were assessed in TGCA RNA-Seq datasets (http://cancergenome.nih.gov/) using the cBioPortal for Cancer Genome (53, 54). Data were downloaded from the UCSC Xena data hub and sample IDs curated using the Broad Institute's Genome Data Analysis Center Firehose. *THBS2* mRNA values were estimated by the RSEM algorithm (55), and log 2 (RSEM+1) transformed for Figure 1B, as parsed and plotted using scripts in Python, R.

Western Blot and ELISA for validation of THBS2 ELISA kits for cross-reactivity with THBS1

The recombinant THBS proteins were obtained from R&D systems and performed western blot with polyclonal goat anti-THBS2 (detection antibody, working conc 0.15nM) and monoclonal mouse anti-THBS2 (capture antibody, working conc. 3nM) to check the cross-reactivity. A detection antibody and a 100-fold molar excess of recombinant THBS1 or THBS2 proteins were incubated in 5% non-fat milk for 30 min at RT for competition assay. The incubated solution was centrifuged at 10K RPM for 15 min to remove any immunocomplexes prior to applying onto a PVDF membrane a total of 2, 10 ng proteins were transferred for detection antibodies. For competition assay of capture antibody, a 10-fold molar excess of recombinant THBS1 or THBS2 proteins with incubated with capture antibody in 5% BSA for 30 min at RT. The incubated solution was centrifuged at 15K RPM for 15 min to remove any immunocomplexes prior to applying onto a PVDF membrane a total of 10, 50 ng proteins were transferred for detection antibodies. The presence of THBS2 in a gel were confirmed by silver staining or re-probing membranes with detection antibody in THBS2-competed membranes. To determine whether presence of THBS1 interferes with the THBS2 ELISA, a 200ng/ml of recombinant THBS1 protein was spiked into various concentration of recombinant THBS2 proteins (0ng/ml to 20ng/ml) or human plasma of wide range of THBS2 in THBS2 ELISA assay.

Immunostaining of THBS2 in human pancreatic cancer tissue

The pancreatic tumor tissue sections were obtained from US Biomax (cat #PA1002) and each tissue spot was individually examined by their own pathologists certified according to WHO published standardizations of diagnosis, classification and pathological grade. Incidental PanIN I–II tissue section was derived from the head and neck of pancreas of pancreatic periampullary cancer patient at the Fox Chase Cancer Center (FCCC) under IRB 09-801 to K.S.Z. and its histology was confirmed by pathologist Dr. Joseph Anderson at FCCC. This tissue blocks do not correspond to plasma samples where we measured plasma THBS2 concentrations. The paraffin embedded tissues were antigen-retrieved by boiling in pH 6.0 Citric Acid buffer after de-paraffination. Next, the endogenous peroxidase activity in tissue slides was quenched in hydrogen peroxide solution for 15 min at RT. Tissues were blocked with avidin/biotin blocking (Vector lab, Burlingame, CA) for 15 min each, followed by non-protein blocker (Thermo Scientific) for 30 min at RT. Primary antibodies were applied and incubated for 12–16 hours at 4° C. Two primary antibodies for THBS2 were used for our study: Goat polyclonal THBS2 antibody (dilution 1:25, sc-7655, Santa Cruz)

and rabbit polyclonal THBS2 antibody (dilution 1:100, TA590658, Origene). It is not clear where TA590658 antibody recognize and whether it detects secreted THBS2. Yet, SC-7655 antibody can detect both secreted and cytoplasmic THBS2 since it targets the epitopes of 15–20 amino acids in length that are located within the first 50 amino acids of the peptide sequence for Thrombospondin-2, whose signal peptides are located between 1-18 amino acids. Only 2 amino acids of the epitope are overlapped to the signal peptides and the remaining are overlapped over the main body of the peptides. A peptide was available for sc-7655 from Santa Cruz, thus the SC-7655 antibodies were incubated with the corresponding peptides in 10-fold excess for 30 min prior to being applied onto tissue section to confirm the specificity of signals. Also, no primary antibody controls for sc-7655 and TA590658 antibodies were used for negative controls. After washing twice, tissues were incubated with biotinylated anti-goat IgG or rabbit IgG (Vector lab) at 37° C for 30 min. Tissue sections were conjugated with avidin-Horseradish peroxidase (HRP) by using VectaStain Elite ABC kit (vector lab) at 37° C for 30 min, followed by developing with DAB peroxidase substrate kit (Vector Lab) for peroxidase for 4–5 min. Developed tissue sections were stained with hematoxylin for nucleus, dehydrated, and mounted. We confirmed the THBS2 sequence of the peptide by mass spectrometry.

Statistical Analysis

The primary comparison for this study was defined as PDAC cases (all stages) vs. healthy controls. In order to explore any relationship between patient demographic information and THBS2 plasma concentrations, a Spearman correlation coefficient was calculated for continuous variables (age) and median expression concentrations were calculated for categorical variables (sex: male vs female, presence or absence of diabetes mellitus). Based upon the data obtained from our Phase 1 and 2 studies, we observed no apparent associations between age, sex, diabetes mellitus, jaundice, and THBS2. Given this lack of association, any concerns regarding the potential for confounding were mitigated and these clinical factors were not included in subsequent multivariable modeling.

Univariate and multivariable logistic regression models were developed to consider each candidate marker (THBS2) alone and combined with CA19-9. The response variable was coded as 1 to indicate the presence of cancer, 0 for controls. Candidate markers (THBS2) were entered as continuous variables. CA19-9 was dichotomized as 0=normal (<55 U/mL) or 1=elevated (55 U/mL). The area under the ROC curve (AUC) was calculated for each model considered. In order to asses if the difference observed between AUCs from the CA19-9 and THBS2 and the CA19-9 alone models was statistically significantly different from 0, we considered a test statistic T ($T = AUC_{Ca199} - AUC_{Ca199+THBS2})^2 / (s^2_{Ca199} + s^2_{Ca199+THBS2})$)(56), which looks at the difference in AUC between the two models divided by the sum of the variances from the two models. The fact that this test statistic followed a chi-squared distribution with 1 degree of freedom under the null hypothesis was used to calculate a resulting p-value. A bootstrap percentile confidence interval approach was used to estimate a 95% confidence interval (CI) for the AUC. This approach re-sampled the dataset (1000 times) then ran the logistic regression models to calculate area under the ROC curve (AUC) on each bootstrapped dataset to approximate the sampling distribution of

the AUC. The 2.5th and 97.5th percentiles from this distribution of AUC values was then used as estimates of lower and upper bounds for the 95% CI for the AUC.

A similar approach was considered for each of the sub-analyses that stratified by stage (early stage, late stage) and other comparison groups (intraepithelial pancreatic mucinous neoplasms, chronic pancreatitis, or pancreatic neuroendocrine tumor). A Kappa statistic was calculated to assess agreement (below cutoff vs above cutoff) in the THBS2 assay results from each of the 2 independent labs in the cross-validation study. Analyses were performed using SAS 9.4 on Linux.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. Cancer Res. 2014; 74:2913–2921. [PubMed: 24840647]

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016; 66:7–30. [PubMed: 26742998]

3. Chari ST, Kelly K, Hollingsworth MA, Thayer SP, Ahlquist DA, Andersen DK, Batra SK, Brentnall TA, Canto M, Cleeter DF, Firpo MA, Gambhir SS, Go VL, Hines OJ, Kenner BJ, Klimstra DS, Lerch MM, Levy MJ, Maitra A, Mulvihill SJ, Petersen GM, Rhim AD, Simeone DM, Srivastava S, Tanaka M, Vinik AI, Wong D. Early detection of sporadic pancreatic cancer: summative review. Pancreas. 2015; 44:693–712. [PubMed: 25931254]

4. Satake K, Kanazawa G, Kho I, Chung YS, Umeyama K. A clinical evaluation of carbohydrate antigen 19-9 and carcinoembryonic antigen in patients with pancreatic carcinoma. J Surg Oncol. 1985; 29:15–21. [PubMed: 3857396]

5. Duffy MJ, Sturgeon C, Lamerz R, Haglund C, Holubec VL, Klapdor R, Nicolini A, Topolcan O, Heinemann V. Tumor markers in pancreatic cancer: a European Group on Tumor Markers (EGTM) status report. Ann Oncol. 2010; 21:441–447. [PubMed: 19690057]

6. Locker GY, Hamilton S, Harris J, Jessup JM, Kemeny N, Macdonald JS, Somerfield MR, Hayes DF, Bast RC Jr. AscovASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. J Clin Oncol. 2006; 24:5313–5327. [PubMed: 17060676]

7. Goonetilleke KS, Siriwardena AK. Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer. Eur J Surg Oncol. 2007; 33:266–270. [PubMed: 17097848]

8. Tempero MA, Uchida E, Takasaki H, Burnett DA, Steplewski Z, Pour PM. Relationship of carbohydrate antigen 19-9 and Lewis antigens in pancreatic cancer. Cancer Res. 1987; 47:5501–5503. [PubMed: 3308077]

9. Harsha HC, Kandasamy K, Ranganathan P, Rani S, Ramabadran S, Gollapudi S, Balakrishnan L, Dwivedi SB, Telikicherla D, Selvan LD, Goel R, Mathivanan S, Marimuthu A, Kashyap M, Vizza RF, Mayer RJ, Decaprio JA, Srivastava S, Hanash SM, Hruban RH, Pandey A. A compendium of potential biomarkers of pancreatic cancer. PLoS Med. 2009; 6:e1000046. [PubMed: 19360088]

10. Mayers JR, Wu C, Clish CB, Kraft P, Torrence ME, Fiske BP, Yuan C, Bao Y, Townsend MK, Tworoger SS, Davidson SM, Papagiannakopoulos T, Yang A, Dayton TL, Ogino S, Stampfer MJ, Giovannucci EL, Qian ZR, Rubinson DA, Ma J, Sesso HD, Gaziano JM, Cochrane BB, Liu S, Wactawski-Wende J, Manson JE, Pollak MN, Kimmelman AC, Souza A, Pierce K, Wang TJ, Gerszten RE, Fuchs CS, Vander Heiden MG, Wolpin BM. Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development. Nat Med. 2014; 20:1193–1198. [PubMed: 25261994]

11. Melo SA, Luecke LB, Kahlert C, Fernandez AF, Gammon ST, Kaye J, LeBleu VS, Mittendorf EA, Weitz J, Rahbari N, Reissfelder C, Pilarsky C, Fraga MF, Piwnica-Worms D, Kalluri R. Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. Nature. 2015; 523:177–182. [PubMed: 26106858]

12. Yu KH, Barry CG, Austin D, Busch CM, Sangar V, Rustgi AK, Blair IA. Stable isotope dilution multidimensional liquid chromatography-tandem mass spectrometry for pancreatic cancer serum biomarker discovery. J Proteome Res. 2009; 8:1565–1576. [PubMed: 19199705]

13. Chan A, Prassas I, Dimitromanolakis A, Brand RE, Serra S, Diamandis EP, Blasutig IM. Validation of biomarkers that complement CA19.9 in detecting early pancreatic cancer. Clin Cancer Res. 2014; 20:5787–5795. [PubMed: 25239611]

14. Makawita S, Dimitromanolakis A, Soosaipillai A, Soleas I, Chan A, Gallinger S, Haun RS, Blasutig IM, Diamandis EP. Validation of four candidate pancreatic cancer serological biomarkers that improve the performance of CA19.9. BMC Cancer. 2013; 13:404. [PubMed: 24007603]

15. Radon TP, Massat NJ, Jones R, Alrawashdeh W, Dumartin L, Ennis D, Duffy SW, Kocher HM, Pereira SP, Guarner posthumous L, Murta-Nascimento C, Real FX, Malats N, Neoptolemos J, Costello E, Greenhalf W, Lemoine NR, Crnogorac-Jurcevic T. Identification of a Three-Biomarker Panel in Urine for Early Detection of Pancreatic Adenocarcinoma. Clin Cancer Res. 2015; 21:3512–3521. [PubMed: 26240291]

16. Schultz NA, Dehlendorff C, Jensen BV, Bjerregaard JK, Nielsen KR, Bojesen SE, Calatayud D, Nielsen SE, Yilmaz M, Hollander NH, Andersen KK, Johansen JS. MicroRNA biomarkers in whole blood for detection of pancreatic cancer. JAMA. 2014; 311:392–404. [PubMed: 24449318]

17. Bauden M, Pamart D, Ansari D, Herzog M, Eccleston M, Micallef J, Andersson B, Andersson R. Circulating nucleosomes as epigenetic biomarkers in pancreatic cancer. Clin Epigenetics. 2015; 7:106. [PubMed: 26451166]

18. Miller ME, Hui SL, Tierney WM. Validation techniques for logistic regression models. Stat Med. 1991; 10:1213–1226. [PubMed: 1925153]

19. Rao, RB., Fung, G., Rosales, R. An Experimental Evaluation. Proceedings of the 2008 SIAM International Conference on Data Mining; 2008. p. 588-598.

20. Jenkinson C, Elliott VL, Evans A, Oldfield L, Jenkins RE, O'Brien DP, Apostolidou S, Gentry-Maharaj A, Fourkala EO, Jacobs IJ, Menon U, Cox T, Campbell F, Pereira SP, Tuveson DA, Park BK, Greenhalf W, Sutton R, Timms JF, Neoptolemos JP, Costello E. Decreased Serum Thrombospondin-1 Levels in Pancreatic Cancer Patients Up to 24 Months Prior to Clinical Diagnosis: Association with Diabetes Mellitus. Clin Cancer Res. 2016; 22:1734–1743. [PubMed: 26573598]

21. Mirus JE, Zhang Y, Li CI, Lokshin AE, Prentice RL, Hingorani SR, Lampe PD. Cross-species antibody microarray interrogation identifies a 3-protein panel of plasma biomarkers for early diagnosis of pancreas cancer. Clin Cancer Res. 2015; 21:1764–1771. [PubMed: 25589628]

22. Nolen BM, Brand RE, Prosser D, Velikokhatnaya L, Allen PJ, Zeh HJ, Grizzle WE, Huang Y, Lomakin A, Lokshin AE. Prediagnostic serum biomarkers as early detection tools for pancreatic cancer in a large prospective cohort study. PLoS One. 2014; 9:e94928. [PubMed: 24747429]

23. Capello M, Bantis LE, Scelo G, Zhao Y, Li P, Dhillon DS, Patel NJ, Kundnani DL, Wang H, Abbruzzese JL, Maitra A, Tempero MA, Brand R, Brennan L, Feng E, Taguchi I, Janout V, Firpo MA, Mulvihill SJ, Katz MH, Hanash SM. Sequential Validation of Blood-Based Protein Biomarker Candidates for Early-Stage Pancreatic Cancer. J Natl Cancer Inst. 2017; 109

24. Maitra A, Hruban RH. Pancreatic cancer. Annu Rev Pathol. 2008; 3:157–188. [PubMed: 18039136]

25. Kim J, Hoffman JP, Alpaugh RK, Rhim AD, Reichert M, Stanger BZ, Furth EE, Sepulveda AR, Yuan CX, Won KJ, Donahue G, Sands J, Gumbs AA, Zaret KS. An iPSC line from human pancreatic ductal adenocarcinoma undergoes early to invasive stages of pancreatic cancer progression. Cell Rep. 2013; 3:2088–2099. [PubMed: 23791528]

26. Bardeesy N, Cheng KH, Berger JH, Chu GC, Pahler J, Olson P, Hezel AF, Horner J, Lauwers GY, Hanahan D, DePinho RA. Smad4 is dispensable for normal pancreas development yet critical in progression and tumor biology of pancreas cancer. Genes Dev. 2006; 20:3130–3146. [PubMed: 17114584]

27. Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, Mankoo P, Carter H, Kamiyama H, Jimeno A, Hong SM, Fu B, Lin MT, Calhoun ES, Kamiyama M, Walter K, Nikolskaya T, Nikolsky Y, Hartigan J, Smith DR, Hidalgo M, Leach SD, Klein AP, Jaffee EM, Goggins M, Maitra A, Iacobuzio-Donahue C, Eshleman JR, Kern SE, Hruban RH, Karchin R, Papadopoulos N, Parmigiani G, Vogelstein B, Velculescu VE, Kinzler KW. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science. 2008; 321:1801–1806. [PubMed: 18772397]

28. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. J Natl Cancer Inst. 2008; 100:1432–1438. [PubMed: 18840817]

29. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, Winget M, Yasui Y. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst. 2001; 93:1054–1061. [PubMed: 11459866]

30. Kolker E, Higdon R, Haynes W, Welch D, Broomall W, Lancet D, Stanberry L, Kolker N. MOPED: Model Organism Protein Expression Database. Nucleic Acids Res. 2012; 40:D1093–1099. [PubMed: 22139914]

31. Nanjappa V, Thomas JK, Marimuthu A, Muthusamy B, Radhakrishnan A, Sharma R, Ahmad Khan A, Balakrishnan L, Sahasrabuddhe NA, Kumar S, Jhaveri BN, Sheth KV, Kumar Khatana R, Shaw PG, Srikanth SM, Mathur PP, Shankar S, Nagaraja D, Christopher R, Mathivanan S, Raju R, Sirdeshmukh R, Chatterjee A, Simpson RJ, Harsha HC, Pandey A, Prasad TS. Plasma Proteome Database as a resource for proteomics research: 2014 update. Nucleic Acids Res. 2014; 42:D959–965. [PubMed: 24304897]

32. Wu C, Orozco C, Boyer J, Leglise M, Goodale J, Batalov S, Hodge CL, Haase J, Janes J, Huss JW 3rd, Su AI. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. Genome Biol. 2009; 10:R130. [PubMed: 19919682]

33. Prassas I, Brinc D, Farkona S, Leung F, Dimitromanolakis A, Chrystoja CC, Brand R, Kulasingam V, Blasutig IM, Diamandis EP. False biomarker discovery due to reactivity of a commercial ELISA for CUZD1 with cancer antigen CA125. Clin Chem. 2014; 60:381–388. [PubMed: 24097894]

34. Bilimoria KY, Bentrem DJ, Ko CY, Ritchey J, Stewart AK, Winchester DP, Talamonti MS. Validation of the 6th edition AJCC Pancreatic Cancer Staging System: report from the National Cancer Database. Cancer. 2007; 110:738–744. [PubMed: 17580363]

35. Molina V, Visa L, Conill C, Navarro S, Escudero JM, Auge JM, Filella X, Lopez-Boado MA, Ferrer J, Fernandez-Cruz L, Molina R. CA 19-9 in pancreatic cancer: retrospective evaluation of patients with suspicion of pancreatic cancer. Tumour Biol. 2012; 33:799–807. [PubMed: 22203495]

36. Yan L, Tonack S, Smith R, Dodd S, Jenkins RE, Kitteringham N, Greenhalf W, Ghaneh P, Neoptolemos JP, Costello E. Confounding effect of obstructive jaundice in the interpretation of proteomic plasma profiling data for pancreatic cancer. J Proteome Res. 2009; 8:142–148. [PubMed: 19055369]

37. Wagner M, Redaelli C, Lietz M, Seiler CA, Friess H, Buchler MW. Curative resection is the single most important factor determining outcome in patients with pancreatic adenocarcinoma. Br J Surg. 2004; 91:586–594. [PubMed: 15122610]

38. Yachida S, Jones S, Bozic I, Antal T, Leary R, Fu B, Kamiyama M, Hruban RH, Eshleman JR, Nowak MA, Velculescu VE, Kinzler KW, Vogelstein B, Iacobuzio-Donahue CA. Distant

metastasis occurs late during the genetic evolution of pancreatic cancer. Nature. 2010; 467:1114–1117. [PubMed: 20981102]

39. Brat DJ, Lillemoe KD, Yeo CJ, Warfield PB, Hruban RH. Progression of pancreatic intraductal neoplasias to infiltrating adenocarcinoma of the pancreas. Am J Surg Pathol. 1998; 22:163–169. [PubMed: 9500216]

40. Nie S, Lo A, Wu J, Zhu J, Tan Z, Simeone DM, Anderson MA, Shedden KA, Ruffin MT, Lubman DM. Glycoprotein biomarker panel for pancreatic cancer discovered by quantitative proteomics analysis. J Proteome Res. 2014; 13:1873–1884. [PubMed: 24571389]

41. Bornstein P, Devarayalu S, Li P, Disteche CM, Framson P. A second thrombospondin gene in the mouse is similar in organization to thrombospondin 1 but does not respond to serum. Proc Natl Acad Sci U S A. 1991; 88:8636–8640. [PubMed: 1924325]

42. Good DJ, Polverini PJ, Rastinejad F, Le Beau MM, Lemons RS, Frazier WA, Bouck NP. A tumor suppressor-dependent inhibitor of angiogenesis is immunologically and functionally indistinguishable from a fragment of thrombospondin. Proc Natl Acad Sci U S A. 1990; 87:6624–6628. [PubMed: 1697685]

43. Hawighorst T, Velasco P, Streit M, Hong YK, Kyriakides TR, Brown LF, Bornstein P, Detmar M. Thrombospondin-2 plays a protective role in multistep carcinogenesis: a novel host anti-tumor defense mechanism. EMBO J. 2001; 20:2631–2640. [PubMed: 11387198]

44. Sun R, Wu J, Chen Y, Lu M, Zhang S, Lu D, Li Y. Down regulation of Thrombospondin2 predicts poor prognosis in patients with gastric cancer. Mol Cancer. 2014; 13:225. [PubMed: 25262009]

45. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HC, Lijmer JG. A. Standards for Reporting of Diagnostic. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Clin Chem. 2003; 49:7–18. [PubMed: 12507954]

46. Chari ST, Leibson CL, Rabe KG, Ransom J, de Andrade M, Petersen GM. Probability of pancreatic cancer following diabetes: a population-based study. Gastroenterology. 2005; 129:504–511. [PubMed: 16083707]

47. Petersen GM. Familial Pancreatic Adenocarcinoma. Hematol Oncol Clin North Am. 2015; 29:641–653. [PubMed: 26226902]

48. Rustgi AK. Familial pancreatic cancer: genetic advances. Genes Dev. 2014; 28:1–7. [PubMed: 24395243]

49. Aggarwal G, Rabe KG, Petersen GM, Chari ST. New-onset diabetes in pancreatic cancer: a study in the primary care setting. Pancreatology. 2012; 12:156–161. [PubMed: 22487526]

50. McWilliams RR, Bamlet WR, de Andrade M, Rider DN, Cunningham JM, Petersen GM. Nucleotide excision repair pathway polymorphisms and pancreatic cancer risk: evidence for role of MMS19L. Cancer Epidemiol Biomarkers Prev. 2009; 18:1295–1302. [PubMed: 19318433]

51. Qian WJ, Kaleta DT, Petritis BO, Jiang H, Liu T, Zhang X, Mottaz HM, Varnum SM, Camp DG 2nd, Huang L, Fang X, Zhang WW, Smith RD. Enhanced detection of low abundance human plasma proteins using a tandem IgY12-SuperMix immunoaffinity separation strategy. Mol Cell Proteomics. 2008; 7:1963–1973. [PubMed: 18632595]

52. Chi H, He K, Yang B, Chen Z, Sun RX, Fan SB, Zhang K, Liu C, Yuan ZF, Wang QH, Liu SQ, Dong MQ, He SM. pFind-Alioth: A novel unrestricted database search algorithm to improve the interpretation of high-resolution MS/MS data. J Proteomics. 2015; 125:89–97. [PubMed: 25979774]

53. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012; 2:401–404. [PubMed: 22588877]

54. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013; 6:pl1. [PubMed: 23550210]

55. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011; 12:323. [PubMed: 21816040]

56. Gonen, Mithat. Analyzing Receiver Operating Characteristic Curves with SAS. SAS Institute; 2007.

Accessible Summary: Detecting pancreatic cancer

Pancreatic ductal adenocarcinoma (PDAC) has a dismal prognosis due to a lack of diagnostics for detecting early stage disease. Kim et al. genetically reprogrammed late stage human PDAC cells to a stem-cell like state, enabling the reprogrammed cells to recapitulate human PDAC progression and revealing secreted candidate markers of early stage disease. The protein thrombospondin-2 (THBS2) was screened against 746 cancer and control human plasma samples in a multi-phase study, and it was found that in combination with the marker CA19-9 to boost detection of the early stages of PDAC in high-risk populations.

A	All Stages (N=10)			Stage I/II/III (N=7)			Stage I/II (N=6)		
	AUC	95% CI		AUC	95% CI		AUC	95% CI	
MMP2	0.57	0.43	0.81	0.557	0.45	0.80	0.533	0.43	0.80
MMP10	no signal			no signal			no signal		
THBS2	0.76	0.56	0.95	0.886	0.71	1.00	0.867	0.64	1.00
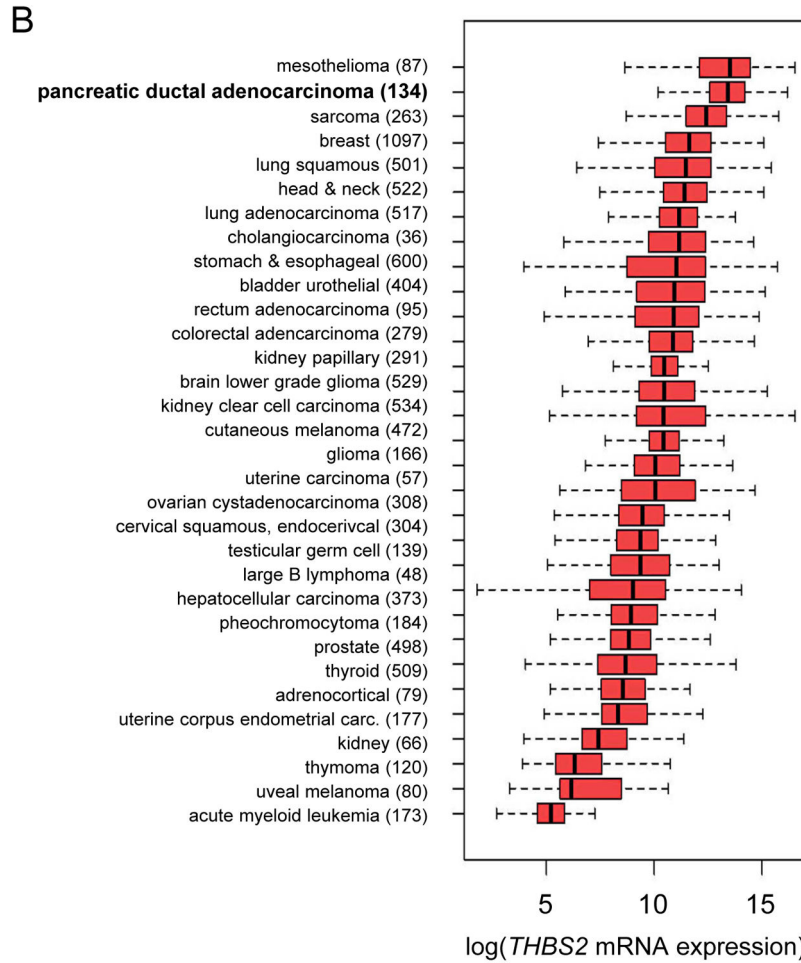
Figure 1. Phase 1 validation studies and THBS2 expression in PDAC and other human tumors

(A) AUC analysis of blinded ELISA data for the proteinsMMP2, MMP10, and THBS2 in plasma samples from 10 patients with PDAC at various stages of disease compared to 10 healthy controls. (B) Boxplots of *THBS2* mRNA expression measured in various human tumors (sample sizes in parentheses) assessed by RNA-seq. Tumors are sorted in order of decreasing median expression of *THBS2* mRNA. Of the pancreatic cancer samples from the TGCA database (n=179), we analyzed only PDAC (n=134). All expression values are log2(RSEM values =1) transformed.
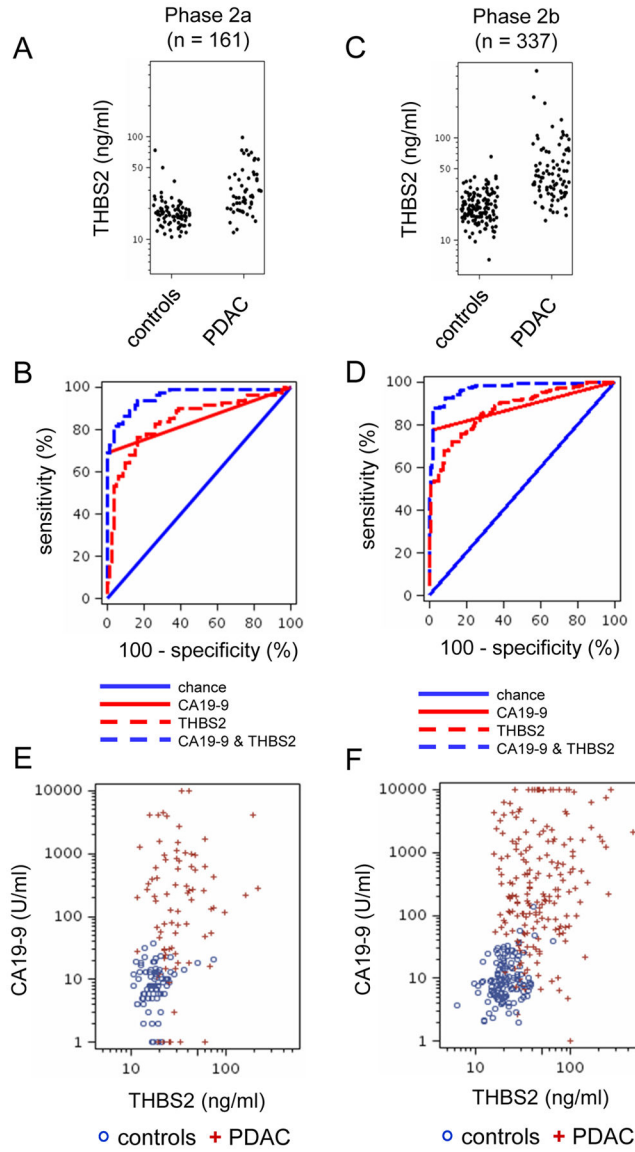
Figure 2. THBS2 and CA19-9 concentrations in plasma samples from patients with PDAC versus healthy controls

(A, CD) Scatter plots of THBS2 concentrations in plasma samples from patients at all stages of PDAC versus controls for the phase 2a (A) and phase 2b (D) validation studies. (B, D, E) ROC curves of THBS2, CA19-9, and THBS2+CA19-9 concentrations in plasma samples from patients with all stages of PDAC versus healthy controls for phase 2a (PDAC n=81, controls n=80) (B) and phase2b (PDAC n=197, controls n=140) (D) studies. P values are shown. (E, F) Scatter plots showing THBS2 and CA19-9 concentrations in plasma samples in patients with all stages of PDAC cases versus healthy controls for Phase 2a (F) and Phase 2b (G) studies.
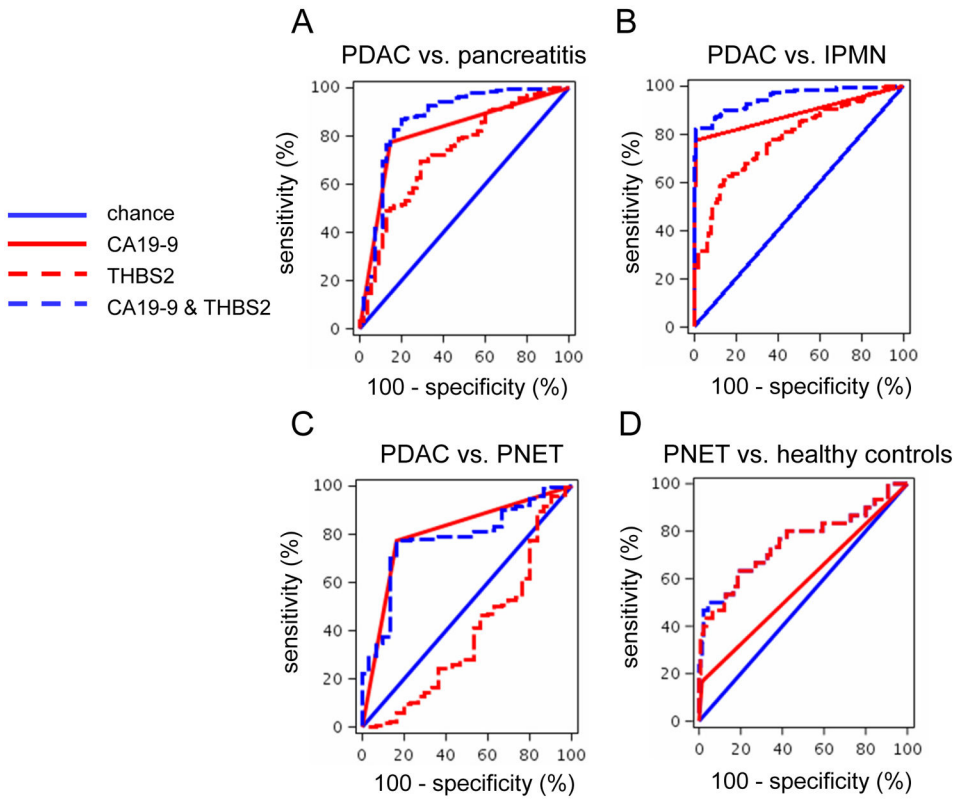
Figure 3. THBS2 and CA19-9 concentrations in plasma samples from all PDAC cases versus versus benign pancreatic disease cases

(A–D)Shown are ROC curves for THBS2, CA19-9, and THBS2+CA19-9 concentrations in plasma samples from patients with PDAC in the Phase 2b study (n=197) versus pancreatitis (n=55, A), PDAC vs. intraepithelial pancreatic mucinous neoplasm (n=115, B), PDAC vs. PNET (n=30, C), and PNET (n=30) vs. healthy controls (n=140, D).
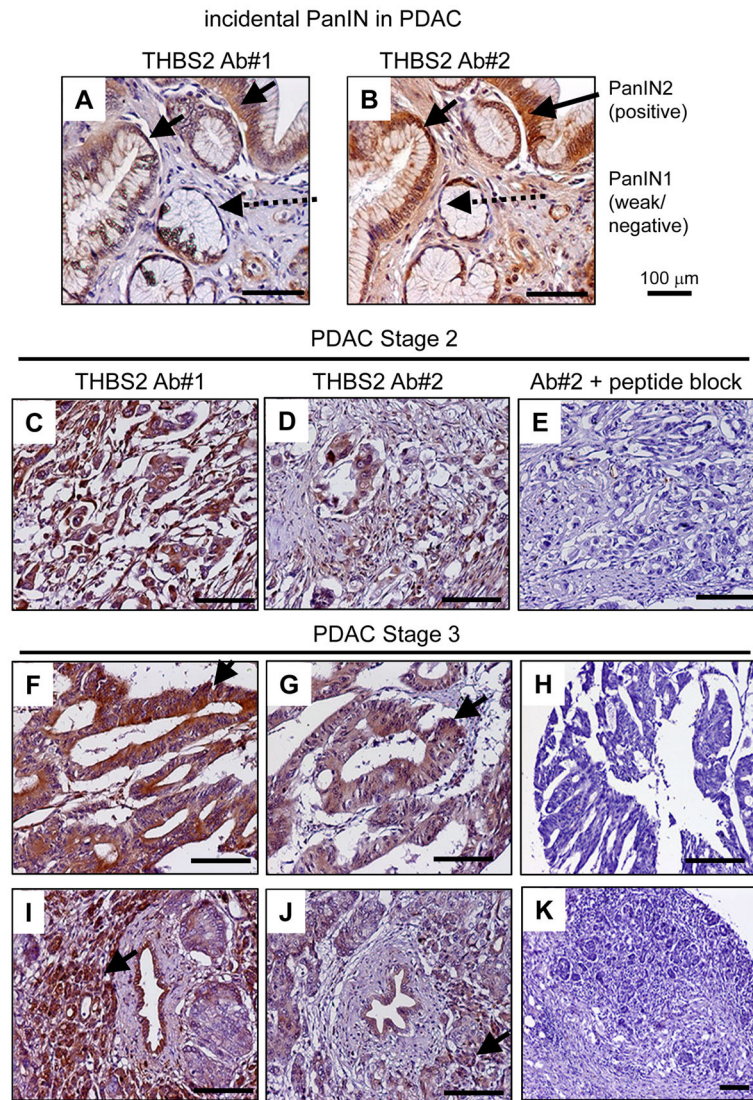
Figure 4. Expression of THBS2 in human PanIN tissue and PDAC tumor tissue
(A) Representative THBS2 immunohistochemistry analysis of incidental PanIN stage I–II tissue derived from the head and neck of a pancreas from a patient with pancreatic periampullary cancer using two different antibodies. The arrows indicate PanIN2 tissue staining positively for THBS2; dotted arrows indicate weak or negative staining of PanIN1 tissue. THBS2 expression, designated by arrows, was also confirmed in stage II PDAC (C–E) and stage III (F–K) pancreatic cancer tissue arrays. Competitive assays were performed for antibody #2 by pre-incubating the antibody with a 10-fold excess of antigen peptide (E, H, K), to confirm target specificity. Brown color indicates THBS2 staining and blue color indicates hematoxylin nuclear staining. THBS2 was detected in the epithelial cells of non-invasive lesions (PanINs and intraepithelial pancreatic neoplasms) and poorly differentiated PDAC tissue as well as in fibroblasts in invasive PDAC tissue (see Table S8 and Figure S6).

Table 1

Demographic and clinical characteristics of patients

Discovery Phase 1	Adenocarcinoma Stage I/II	Adenocarcinoma Stage III/IV	Controls	IPMN, no Adenocarcinoma	PNET	Pancreatitis
	N=6	N=4	N=10			
Age	56.8 (7.5)	65.0 (13.0)	62.2 (15.4)			
Male Gender	4 (66.7%)	2 (50.0%)	5 (50.0%)			
Body Mass Index (kg/m^2)	31.2 (8.6)	26.3 (5.4)	26.0 (3.9)			
Personal History of Diabetes	1 (16.7%)	0	2 (20.0%)			
CA19-9	20770.8 (47060.9)	111224.0 (217927.2)	12.0 (6.4)			
Stage of Disease						
II	1 (16.7%)					
IIA	1 (16.7%)					
IIB	4 (66.7%)					
III		1 (25.0%)				
IV		3 (75.0%)				
Validation Phase 2a						
	N=58	N=23	N=80			N=28
Age	67.6 (9.4)	68.6 (10.9)	67.4 (9.8)			61.5 (9.4)
Male Gender	43 (74.1%)	12 (52.2%)	54 (67.5%)			19 (67.9%)
Body Mass Index (kg/m^2)	28.9 (5.2)	28.5 (5.7)	27.2 (4.7)			26.3 (4.8)
Personal History of DiabetesDM	15 (25.9%)	3 (13.0%)	12 (15.0%)			3 (10.7%)
CA19-9	305.3 (411.1)	2137.5 (2983.7)	10.6 (6.9)			68.9 (165.5)
Stage of Disease						
I	1 (1.7%)					
IA	1 (1.7%)					
IB	6 (10.3%)					
II	16 (27.6%)					
IIA	13 (22.4%)					
IIB	21 (36.2%)					
III		10 (43.5%)				
IV		13 (56.5%)				

Validation Phase 2b

	Adenocarcinoma Stage I/II	Adenocarcinoma Stage III/IV	Controls	IPMN, no Adenocarcinoma	PNET	Pancreatitis
	N=88	N=109	N=140	N=115	N=30	N=55
Age	66.5 (11.3)	64.4 (11.0)	65.8 (10.8)	68.8 (8.7)	63.2 (7.1)	55.9 (17.7)
Male Gender	45 (51.1%)	62 (56.9%)	70 (50.0%)	58 (50.4%)	22 (73.3%)	27 (50.0%)
Body Mass Index (kg/m²)	28.4 (5.4)	29.1 (6.0)	27.1 (4.4)	26.5 (4.0)	29.0 (4.9)	27.8 (5.0)
Personal History of DiabetesDM	34 (38.6%)	25 (22.9%)	15 (10.7%)	20 (17.4%)	10 (33.3%)	7 (13.0%)
CA19-9	633.8 (1665.9)	2399.3 (3481.1)	12.0 (14.5)	15.4 (12.3)	45.3 (98.0)	35.9 (66.0)
Stage of Disease						
I	4 (4.5%)					
IA	2 (2.3%)					
IB	4 (4.5%)					
II	37 (42.0%)					
IIA	15 (17.0%)					
IIB	26 (29.5%)					
III		41 (37.6%)				
IV		68 (62.4%)				

IPMN: Intraductal papillary mucinous neoplasm; PNET: Pancreatic neuroendocrine tumor; Continuous variables (Age, Body Mass Index, and CA19-9 concentration) are presented as mean (standard deviation); Categorical variables (Male Gender, Personal History of Diabetes, and Stage of Disease) are presented as frequency (percentage).

Table 2

Area under the ROC curve (AUC) calculations of ELISA results data of PDAC at different stages, bootstrapped (1000 repetitions) at 95% confidence intervals.

PDAC vs. healthy controls	N	CA19-9(55) AUC	95% CI		THBS2 AUC	95% CI		CA19-9(55) + THBS2 AUC	95% CI		p-value
Validation Phase 2a											
All Stages	81/80	0.846	0.81	0.89	0.842	0.80	0.90	0.956	0.93	0.98	0.0003
Stage I/II	58/80	0.845	0.80	0.89	0.832	0.78	0.89	0.946	0.92	0.98	0.0067
Stage I	8/80	0.938	0.83	1.00	0.839	0.69	0.98	0.977	0.93	1.00	0.5038
Stage II	50/80	0.830	0.78	0.88	0.830	0.78	0.89	0.940	0.91	0.98	0.0080
Stage III/IV	23/80	0.846	0.79	0.90	0.842	0.78	0.90	0.956	0.92	0.99	0.0003
Validation Phase 2b											
All Stages	197/140	0.881	0.86	0.91	0.875	0.85	0.90	0.970	0.96	0.98	< 0.0001
Stage I/II	88/140	0.834	0.79	0.87	0.887	0.85	0.92	0.960	0.94	0.98	< 0.0001
Stage I	10/140	0.793	0.66	0.93	0.896	0.83	0.97	0.952	0.92	0.99	0.0574
Stage II	78/140	0.839	0.80	0.88	0.885	0.85	0.92	0.961	0.94	0.98	< 0.0001
Stage III/IV	109/140	0.919	0.89	0.95	0.866	0.83	0.90	0.980	0.97	0.99	0.0028

Table 3

THBS2 concentration cut off points based upon percentiles of distribution in control plasma samples

Marker	Cutoff	Phase 2a		Phase 2b	
		Sensitivity	Specificity	Sensitivity	Specificity
CA19-9 (55)		69.14	100	77.66	98.57
THBS2 (ng/ml)					
95%	36	33.33	96.25	58.38	93.57
96%	37	33.33	97.50	57.36	95.00
97%	38	30.86	97.50	53.81	96.43
98%	40	28.40	97.50	53.30	97.86
99%	42	24.69	97.50	51.78	99.29
100%	73.4	7.41	100	23.35	100
CA19-9 (55) & THBS2 (ng/ml)					
95%	36	74.07	96.25	88.32	92.86
96%	37	74.07	97.50	88.32	94.29
97%	38	74.07	97.50	87.82	95.71
98%	40	74.07	97.50	87.82	97.14
99%	42	72.84	97.50	87.31	97.86
100%	73.4	69.14	100	81.73	98.57

Table 4

Area under ROC curve (AUC) calculations of ELISA results data for all stages PDAC versus patients with benign pancreatic diseases, bootstrapped (1000 repetitions) at 95% confidence intervals

PDAC vs. benign pancreatic diseases		CA19-9(55)		THBS2		CA19-9(55) + THBS2		
	N	AUC	95% CI	AUC	95% CI	AUC	95% CI	p-value
Validation Phase 2a								
PDAC vs. Pancreatitis	81/28	0.774	0.71 0.84	0.727	0.64 0.81	0.842	0.76 0.92	0.2740
Validation Phase 2b								
PDAC vs. Pancreatitis	197/55	0.816	0.77 0.86	0.731	0.67 0.79	0.867	0.82 0.92	0.2450
PDAC vs. IPMN	197/115	0.884	0.86 0.91	0.784	0.74 0.82	0.952	0.93 0.97	0.0003
PDAC vs. PNET	197/30	0.805	0.75 0.86	0.444	0.36 0.71	0.785	0.72 0.89	0.6502
PNET vs. Controls	30/140	0.576	0.52 0.63	0.751	0.67 0.84	0.755	0.67 0.85	0.0089