

Article



A High Performing Biomarker Signature for Detecting Early-Stage Pancreatic Ductal Adenocarcinoma in High-Risk Individuals

Norma A. Palma ^{1,*}, Aimee L. Lucas ², Bryson W. Katona ³, Alcibiade Athanasiou ⁴, Natasha M. Kureshi ¹, Lisa Ford ¹, Thomas Keller ⁵, Stephen Weber ⁵, Ralph Schiess ⁴, Thomas King ¹, Diane M. Simeone ⁶ and Randall Brand ⁷

- ¹ Immunovia, Inc., Durham, NC 27713, USA
- ² Division of Gastroenterology and Hepatology, Mount Sinai West and Sinai Morningside Hospitals, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
- ³ Department of Gastroenterology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104, USA
- ⁴ Proteomedix AG, 8952 Schlieren, Switzerland
- ⁵ ACOMED Statistic, 04109 Leipzig, Germany
- ⁶ Department of Surgery, University of California, San Diego Health, Moores Cancer Center, San Diego, CA 92037, USA
- ⁷ Division of Gastroenterology, Hepatology and Nutrition, University of Pittsburgh Medical Center, Pittsburgh, PA 15213, USA
- * Correspondence: norma.palma@immunovia.com

Simple Summary: Using machine learning, this study identified a 5-plex serum-based biomarker signature with high sensitivity and specificity for detecting pancreatic ductal adenocarcinoma (PDAC) in a cohort of Stage I and II cases and high-risk controls undergoing surveillance. To date, no blood-based biomarkers have been approved specifically for PDAC surveillance or to aid in its diagnosis. Underrepresentation of high-risk controls in study populations has been cited as a primary reason for the failure of PDAC biomarkers in clinical trials. By addressing this shortcoming in our study cohort, our findings demonstrate a blood-based test for detection of early-stage PDAC that performs well in the clinical population likely to gain the most benefit from its use. Overall, these findings represent an important step towards improving early diagnostic success and extending survival of this devastating disease.

Abstract: Background/Objectives: Early detection of pancreatic cancer can improve patient survival, and blood-based biomarkers to aid in this are a significant need. The goal of this study was to develop and evaluate the performance of a 4- to 6-plex biomarker signature for detection of early-stage pancreatic ductal adenocarcinoma (PDAC) that performs well in high-risk controls. Methods: Enzyme-linked immunosorbent assays were used to measure 10 previously identified serum protein biomarker candidates in Stage I and II PDAC cases (n = 128), high-risk controls (n = 465), and normal-risk controls (n = 30). Various combinations of biomarker candidates (models) were trained using machine learning and tested for robustness in differentiating cases from controls on the full cohort and in clinically relevant sub-types including those with diabetes, those \geq 65 years of age, and low producers of carbohydrate antigen 19-9 (CA 19-9). Results: At 98% specificity, the top performing model, which was comprised of tissue inhibitor of metalloproteinase 1 (TIMP1), intracellular adhesion molecule 1 (ICAM1), thrombospondin 1 (THBS1), cathepsin D (CTSD), and CA 19-9, achieved 85% sensitivity in the full cohort and sensitivities of 91% in diabetics, 90% in \geq 65 years of age, and 60% in low CA 19-9 producers. This model demonstrated significantly higher sensitivity in detecting PDAC in the full cohort and all sub-populations compared to CA 19-9 alone (p < 0.001). **Conclusions**: Our findings demonstrate the feasibility of a blood-based assay for detecting early-stage PDAC in high-risk individuals and key



Academic Editor: Atsushi Masamune

Received: 9 May 2025 Revised: 30 May 2025 Accepted: 31 May 2025 Published: 2 June 2025

Citation: Palma, N.A.; Lucas, A.L.; Katona, B.W.; Athanasiou, A.; Kureshi, N.M.; Ford, L.; Keller, T.; Weber, S.; Schiess, R.; King, T.; et al. A High Performing Biomarker Signature for Detecting Early-Stage Pancreatic Ductal Adenocarcinoma in High-Risk Individuals. *Cancers* **2025**, *17*, 1866. https://doi.org/10.3390/ cancers17111866

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/ licenses/by/4.0/). sub-populations, representing an important step towards improving diagnostic success for early-stage disease.

Keywords: pancreatic ductal adenocarcinoma; biomarkers; high-risk; surveillance; potentially resectable; familial pancreatic cancer; germline mutations

1. Introduction

Pancreatic ductal adenocarcinoma (PDAC) is a highly aggressive malignancy that claims approximately 50,000 lives each year and currently ranks as the third leading cause of cancer-related death in the United States [1]. Despite advancements in the understanding of PDAC biology and efforts to develop new therapies, its incidence continues to rise, and the average 5-year survival rate across all stages combined is only 13% [2,3].

PDAC usually remains asymptomatic until it has advanced to later stages, when treatments are considerably less effective [4,5]. Early detection of disease impacts patient outcomes such that diagnosis at Stage I or II results in a 44% 5-year survival rate versus a 3.1% 5-year survival rate, where surgical intervention is no longer an option [6]. The low incidence of PDAC in the general population limits recommended surveillance to those who carry a germline mutation that puts them at high risk for PDAC (i.e., CDKN2A, STK11, BRCA2, etc.), have a family history meeting the definition of familial pancreatic cancer, or have a mucinous pancreatic cyst [7–9]. Unfortunately, the most widely used surveillance tests, which include computed tomography (CT), magnetic resonance imaging (MRI), and endoscopic ultrasound (EUS), have limited sensitivity against small, early-stage tumors [10]. Carbohydrate antigen 19-9 (CA 19-9), the one biomarker approved for use in PDAC patients [11], can inform response to treatment but lacks the necessary sensitivity or specificity to be recommended for diagnosis or surveillance [12,13]. Additionally, 5–10% of the general population and up to 25% of some ethnic populations harbor homozygous alterations in the FUT2/3 genes, which renders them incapable of producing and/or secreting CA 19-9 [14,15]. These individuals, as well as those with FUT2/3 variants that result in low, but not nil, production of CA 19-9 are disadvantaged by relying solely on the use of this biomarker. Other populations, including diabetics and the elderly are also disadvantaged by relying on CA 19-9 because its levels tend to increase in response to both aging and diabetes [16,17].

To improve upon existing diagnostic tools and extend survival rates, significant efforts have been made to identify noninvasive biomarkers for detecting early-stage PDAC [18–21]. Although many biomarker candidates have been published, none have been approved specifically to aid in the diagnosis of PDAC, emphasizing the importance of continuing these efforts.

In a previous study, we evaluated candidate biomarkers for early-stage PDAC using Olink multiplex technology and targeted immunoassays in a cohort of Stage I and II PDAC cases and age- and sex-matched controls [22]. This study identified 41 serum-based proteins that differentiated PDAC from controls (p < 0.05). The primary endpoint of the current study was to develop a four- to six-plex biomarker signature from this initial pool of 41 candidates then assess its ability to discern early-stage PDAC in a genetic/familial high-risk patient population. Secondary study endpoints included developing a model that outperformed CA 19-9 alone in the full cohort and sub-populations studied.

2. Materials and Methods

2.1. Study Design and Participants

This retrospective study was conducted according to protocols that adhered to ethical standards outlined in the Declaration of Helsinki and were approved by the Institutional Review Boards/Ethics Committees of all 12 participating study sites before the study began (Supplemental Table S1). All participants provided written informed consent. The study complied with all relevant ethical regulations for research involving human participants. Banked serum samples were blinded for their clinical information using a unique clinical identifier such that no patient data or disease status was known by laboratory personnel until after protein biomarkers had been measured and the data set had been locked. The study population included: (1) patients recently diagnosed with Stage I or II PDAC (n = 128) (2) high-risk controls identified through PDAC surveillance programs (n = 465), and (3) normal-risk controls (n = 30). In this manuscript high-risk and normal-risk controls are collectively referred to as controls. Samples were retrospectively selected according to PDAC status from different centers and therefore were not randomized in relation to the screening population. The presence of diabetes (regardless of duration of diagnosis), age, and CA 19-9 status were evaluated in the groups examined. Where indicated, diabetics, those \geq 65 years of age, and low CA 19-9 producing groups (<37 U/mL) were evaluated as sub-populations. 65 years was chosen as the cutoff to define older study participants because it would allow data from the Medicare population to be evaluated. Inclusion and exclusion criteria for study participation and methods used to select samples for the study are shown in Supplemental Figure S1.

2.2. Candidate Protein Selection and Measurement

From the 41 biomarker candidates identified in our previous study [22], 10 were selected for follow-up analysis based on their reported roles in PDAC, ability to distinguish PDAC cases from controls in the previous study, and/or the availability of proteinspecific immunoassay materials that meet technical suitability for clinical use. CA 19-9 was also included for analysis due to its status as the only FDA-approved marker for PDAC patients. We targeted a 4- to 6-plex biomarker panel in ELISA format to ensure a reproducible and efficient application in clinical testing. CA 19-9 was measured by Dr. Risch Services AG (Buchs, Switzerland) using a COBAS 8000 modular analyzer (Roche Diagnostics, Rotkreuz, Switzerland) according to manufacturer's instructions. ELISAs were optimized and the following proteins were quantitated by Proteomedix (Schlieren-Zurich, Switzerland): carboxypeptidase B1 (CPB1) (Aviva Systems Biology, San Diego, CA, USA #OKEH07759, RRID:AB_3678911), cathepsin D (CTSD) (Proteomedix AG, Schlieren, Switzerland #Proclarix® CTSD ASSAY), intracellular adhesion molecule-1 (ICAM1) (Proteomedix, AG, # ICAM1), tissue inhibitor of metalloproteinases 1 (TIMP1) (Biotechne R&D Systems, Minneapolis, MN, USA #MAB97, RRID:AB_3678914 (capture antibody), #BAF970, RRID:AB_356824 (detection antibody), #970-TM (protein)), thrombospondin 1 (THBS1) (Proteomedix AG, Schlieren, Switzerland #Proclarix® THBS1 ASSAY), glycoprotein nonmetastatic melanoma protein B (GPNMB) (Biotechne R&D Systems, Minneapolis, MN, USA #DY2550, RRID:AB3678917), kallikrein related peptidase 10 (KLK10) (Aviva, San Diego, CA, USA #OKDD04348, RRID:AB_3678918), mer tyrosine kinase (MERTK) (LS-Bio, #LS-F462, RRID:3678919), phospholipase A2 group 1B (PLA2G1B) (LS-Bio, Seattle, WA, USA #LS-F33272, RRID:AB_3678920), and latent transforming growth factor-β binding protein 2 (LTBP2) (Abcam, Waltham, MA, USA #ab13988, RRID:AB_300815).

2.3. Statistical Analyses

A sample size estimation was performed at 80% power and an assumed disease prevalence of 20% as described [23]. The relationship between individual protein biomarkers and PDAC diagnostic status was evaluated using Mann–Whitney U tests. The training of the algorithm was performed as previously described [22]. Briefly, markers were used to train generalized linear models with embedded feature selection that included all possible 2- to 6-plex combinations and relevant imputations of candidate proteins using PDAC diagnosis as the binary outcome. All models were trained and cross validated on the full cohort and tested for overfitting using a 7 to 10 ratio of training/test splits from the full cohort. Their performance was then evaluated on the diabetic, \geq 65 years, and low CA 19-9 producer sub-populations based on the sensitivity they achieved at 98% specificity. Area under receiver operating characteristic (ROC) curves (AUC) for the final model and its two variations were compared using the Delong method [24]. Sensitivities and specificities were compared using the McNemar test [25] or two-tailed student's *t*-tests. For all tests, *p*-values < 0.05 were considered statistically significant. All statistical analyses were performed in R, version 4.4.1., RRID:PRJ_000000000 [26].

3. Results

3.1. Patient Characteristics

Serum samples from 623 study participants (n = 128 cases, n = 495 controls) were analyzed. Demographics of the study population are shown in Table 1 (full cohort) and Supplemental Table S2 (sub-populations). The cohort's median age was 64 years (SD: \pm 8 years), with cases (68 years, SD: \pm 10 years) being older than controls (63 years, SD: \pm 7 years). No significant differences in the expression of final candidate proteins based on age were found (Supplemental Figure S2).

Clinical Variable	All (n = 623)	Cases (n = 128)	Controls (n = 495)
Age, median (min-max)	64 (40–95)	68 (46–95)	63 (40-83)
CA 19-9 U/mL, median (min-max)	10 (0.6–1000) *	78 (0.6–1000)	9 (0.6–70)
Sex male, n (% of column heading)	244 (39%)	69 (54%)	175 (35%)
PDAC Stage I, n (% of column heading)	88 (14%)	88 (69%)	N/A
PDAC Stage II, n (% of column heading)	40 (6%)	40 (31%)	N/A
White (% of column heading)	481 (77%)	52 (41%)	429 (86%)
Not White (% of column heading)	28 (4%)	4 (3%)	28 (22%)
Not reported (% of column heading)	114 (18%)	72 (56%)	42 (33%)

Table 1. Study participant demographics, full cohort.

* Minimum CA 19-9 levels (0.6 U/mL) represent the limit of detection of the COBAS instrument. N/A = not applicable.

A disparity in sex distribution was observed: 379 participants (61%) were female and 244 (39%) were male. This was largely due to the female predominance of control subjects (65% female vs. 35% male) with a smaller male predominance in cases (46% female vs. 54% male). Sex disparities result from the typical composition of high-risk surveillance programs (predominantly female) [27] and the somewhat higher incidence of PDAC in males [28,29]. One of the candidate proteins, CBP1, demonstrated a significant difference in serum concentration between male and female controls and was eliminated from the final model selection. The final biomarkers used in the models showed no difference in expression between male and female controls or between male and female cases (Supplemental Figure S3).

3.2. Protein Analysis and Selection of Final Candidates

Using commercially available materials and reagents from various manufacturers, we developed ELISAs to quantitate 10 pre-selected biomarker candidates identified in our discovery study [22]: CTSD, CPB1, ICAM1, TIMP1, GPNMB, MERTK, PLA2G1B, KLK10, LTBP2, and THBS1. Due to poor analyte stability, high lot-to-lot variability of the assay, and/or extensive covariance with other potential final model candidates KLK10, MERTK, GPNMB, and PLA2G1B were excluded from further analysis.

We characterized expression of the remaining 6 proteins and CA 19-9 by comparing their serum concentrations between PDAC cases and controls. The abundance of every analyte except THBS1 was significantly different between the two groups (p < 0.001) (Table 2). The expression pattern of each analyte matched those observed in our previous biomarker discovery study (Supplemental Table S3). We next evaluated the diagnostic performance of the individual markers. At 98% specificity, CTSD, TIMP1, ICAM1, and CA 19-9 achieved 31–65% sensitivity while LTBP2, CPB1, and THBS1 achieved 2–12% sensitivity (Table 2). Compared to their performance in the full cohort, each analyte performed similarly or better in the diabetic, \geq 65 years, and low CA 19-9 sub-populations (Supplemental Table S4).

Table 2. Expression levels and sensitivities of individual candidate markers when applied to the full cohort.

Marker	<i>p</i> -Value (Controls vs. Cases)	Concentration Means, Cases (CA19-9, U/mL; All Others, ng/mL)	Concentration Means, Controls (CA19-9, U/mL; All Others, ng/mL)	Sensitivity at 98% Specificity, % (95% CI)
CA 19-9	< 0.001	248	12	65 (57–73)
ICAM1	<0.001	363	161	52 (44–61)
TIMP1	< 0.001	657	375	48 (39–56)
CTSD	<0.001	646	333	31 (23–39)
LTBP2	<0.001	215	133	12 (6–17)
CPB1	< 0.001	8139	3199	6 (2–10)
THBS1	0.220	30613	28844	2 (0–5)

Next, we compared the distribution of protein levels between all cases and all controls. CTSD, TIMP1, ICAM1, CPB1, LTBP2, and CA 19-9 demonstrated clear separation, while THBS1 demonstrated nearly complete overlap between the two groups, showing that its amounts were not significantly different between cases and controls when analyzed in the full cohort (Figure 1). However, in the diabetic sub-population THBS1 demonstrated a significant separation in its distribution (p = 0.031) (Supplemental Table S4 and Supplemental Figure S4), suggesting that it could be helpful in detecting PDAC in this sub-group despite its low overall sensitivity. For this reason, THBS1 was included in the model development studies.



Figure 1. Density plots showing distribution of analyte concentrations in all cases and controls (full cohort).

3.3. Model Development and Cross Validation

To identify which biomarker signatures (i.e., models) achieve the best diagnostic performance, we used machine learning to train all possible 2- to 6-plex combinations of ICAM1, THBS1, TIMP1, CTSD, LTBP2, and CPB1 with and without CA 19-9 on the full cohort.

The top performing model (Model v1), which was comprised of TIMP1, ICAM1, CTSD, THBS1, and CA 19-9, achieved 85% sensitivity at 98% specificity (95% CI) (Figure 2A,B) compared to 65% (95% CI) achieved by CA 19-9 alone (p < 0.001) (Figure 2B). We assessed the robustness of the model using a standard repeated data split and cross-validation strategy. Random training and validation splits (7 to 10 ratio) were repeated 1,000 times for each model. Boxplot sensitivities for the model were evaluated on the full cohort. Sensitivity of the training cohort at 98% specificity consistently fell within the range of performance of the validation cohort from the random splits across the model (Figure 2C), suggesting that the model was not over-fitted. Overall, this model (Model v1) diagnosed 10 of the 495 controls as PDAC positive (2% false positive rate) and 19 of the 128 PDAC cases as negative (15% false negative rate) (Figure 2D). Two variations of this model, which lacked either THBS1 (Model v2) or CTSD (Model v3) demonstrated similar performance (Supplemental Figure S5).



Β.

r					
Comparison of s	ensitivitie	s at 9	8% specificity.		
Biomarker(s)		Sensitivity, % (95% CI)		3)	p-value
Model: TIMP1 + ICAM1 + CTSD + THBS1 + CA 19-9		85 (79–91)			< 0.001
CA 19-9		65 (57–73)		F	Reference
Comparison of AUCs.					
	AUC		95% CI	F	o-value
Model	0.945		0.917–0.973		
CA 19-9	0.857		0.808-0.906		
Difference between Model and CA 19-9	0.089		0.037–0.139	(0.0007

D.

₹

Misclassification rates of Model v1.					
	Controls		Cases		
	Model v1		Model v1		
	Positive	Negative	Positive	Negative	
CA 19-9 > 37 U/mL	1	14	82	1	
CA 19-9 ≤ 37 U/mL	9	471	27	18	

Figure 2. Model v1 performance and robustness in the full cohort. (A) ROC curves of the top performing model and CA 19-9 only. Gray line (comparator) depicts no model discrimination between groups (AUC = 0.5). (B) Top, Comparison of model sensitivity at 98% specificity presented as a percentage with a 95% CI. p-value according to McNemar indicates the difference in model sensitivity compared to CA 19-9 alone. Bottom, Comparison of AUCs for the Model and CA 19-9 showing 95% CIs, the difference between the Model and CA 19-9 AUCs, 95% CI, and associated *p*-value. (C) Density plot and boxplots of the 1000x random splits showing robustness of Model v1. The yellow star indicates model performance when trained using the full cohort. (D) Table showing the number of cases and controls that were misclassified by Model v1 and their CA 19-9 status.

3.4. Model Performance in Sub-Populations

We next assessed the performance of Model v1 in the three pre-defined subpopulations. In the diabetic and \geq 65 years groups, the model achieved sensitivities of 91% and 90%, respectively, at 98% specificity (95% CI), which outperformed CA 19-9 alone by 12 percentage points in the diabetic group and 19 percentage points in the \geq 65 group (p < 0.001) (Figure 3). In the low CA 19-9 group, the model achieved 60% sensitivity at 98% specificity (95% CI), significantly outperforming CA 19-9 alone (9% sensitivity) at 98% specificity (p < 0.001) (Figure 3).



Figure 3. Model performance on sub-populations. (**A**) ROC curves showing model performance in the diabetic, ≥ 65 years, and low CA 19-9 sub-populations. Gray lines (comparators) depict no model discrimination between groups (AUC = 0.5). Dashed gray lines represent 90% specificity. (**B**) Sensitivities at 98% specificity are presented as a percentage with a 95% CI. Sensitivity achieved by Model v1 in all sub-populations compared to CA 19-9 alone, p < 0.001.

4. Discussion

A blood-based biomarker assay capable of detecting early-stage PDAC could address a significant clinical need. CA 19-9 alone lacks adequate sensitivity and specificity to serve as a surveillance marker and routine imaging to screen for PDAC is not recommended for the general population [30]. Furthermore, currently available diagnostic imaging techniques have limited sensitivity for detecting small, early-stage tumors. Our goal is to develop a noninvasive test that can aid in the detection of early-stage PDAC by implementing biomarkers that have robust sensitivity and specificity against Stage I and II disease in those at high risk for PDAC and in sub-populations that may be challenging to diagnose.

Here, we identified a model comprised of TIMP1, ICAM1, CTSD, THBS1, and CA 19-9 that distinguished PDAC cases from controls with a sensitivity of 85% at 98% specificity, thereby achieving the primary endpoint of this study. At 98% specificity this model also achieved sensitivities of 91%, 90%, and 60% in diabetics, those \geq 65 years of age, and low CA 19-9 producers, respectively, to outperform CA 19-9 alone in both the full cohort and all clinical sub-populations, thereby achieving the secondary study endpoints. Two variations of the model that omitted either THBS1 or CTSD showed comparable performance, indicating that an abbreviated signature comprised of 4 biomarkers may be sufficient for clinical use.

TIMP1, ICAM1, CTSD, THBS1, and CA 19-9 are differentially expressed in PDAC tumors compared to noncancerous tissue and each play various roles in PDAC pathophysiology (Figure 4). High expression of TIMP1 and ICAM1 is associated with poor prognosis in PDAC patients [31–34], and both proteins are thought to promote metastasis by facilitating neutrophil extracellular trap (NET) formation as well as neutrophil and macrophage infiltration into the pancreas and liver where they create a premetastatic niche

for tumors [35–37]. Like TIMP1 and ICAM1, CTSD is typically overexpressed in PDAC and has been implicated in cancer cell proliferation, migration, and tumor invasion [38,39]. THBS1 has been shown to both suppress and promote angiogenesis in PDAC through mechanisms that are currently under investigation [40–42]. THBS1 also promotes tumor metastasis in post-dormancy stages of PDAC by converting latent TGF- β 1 complexes to their active form through loss of the tumor suppressor Smad4 and upregulation of MMP-9 production [43,44].



Figure 4. Signature analytes and their role in pancreatic cancer. CTSD = cathepsin D, THBS1 = thrombospondin 1, TIMP1 = tissue inhibitor of metalloproteinase 1, ICAM-1 = intracellular adhesion molecule 1, CA 19-9 = carbohydrate antigen 19-9.

Due to their roles in PDAC progression TIMP1, ICAM1, CTSD, and THBS1 have all been previously evaluated as diagnostic biomarkers. Several reports show that incorporating any one of these proteins into diverse biomarker signatures, some of which include CA 19-9, improves diagnostic success over CA 19-9 alone [31,36,45–47]. The findings in the current study agree with these previous reports. Unlike previous studies our study focused on early-stage, potentially resectable PDAC cases and investigated diagnostic performance in both high-risk individuals undergoing surveillance and clinically relevant sub-populations. A key finding of our study is a set of biomarkers that performs well in not only the general PDAC population, but also in genetic/familial high-risk surveillance populations, and sub-populations that face diagnostic challenges, especially low CA 19-9 producers.

Investigators have attempted to improve diagnostic success in low CA 19-9 producers by identifying alternative diagnostic biomarkers [48,49] and establishing different clinical cutoff levels for these individuals relative to the general population [50]. Thus far, both approaches have had limited success. The performance of our model in the low CA 19-9 group is particularly noteworthy because for the first time it demonstrates a clinically useful, non-invasive test for early-stage PDAC patients who do not overexpress CA 19-9.

Although our findings are promising, we acknowledge this study had limitations. Due to the retrospective nature of the sample collection, the status of certain clinical variables, including biliary obstruction and pancreatitis at the time of sample collection are unknown, which prevented evaluation of the model's specificity in these clinical contexts. This aspect will be evaluated in follow-up studies. Another limitation was the low racial/ethnic diversity in the test population, which remains a challenge in both clinical trials and high-

risk pancreatic cancer surveillance [51]. We are addressing this issue in our current work by targeting study populations that more closely resemble the ethnic composition of the United States. As a third limitation, we note that, to appropriately power the study, 10 case samples that had been in storage significantly longer than the rest of the cohort (up to 10 years) had to be included in analysis. Of these, four were classified as false negatives (sensitivity 60%) suggesting that long-term sample storage may adversely affect test performance. This possibility is not surprising considering serum proteins are known to degrade over time even when stored at -80 °C. We are testing the effect of long-term storage in ongoing validation studies by comparing test sensitivity and specificity between the numerically oldest (50%) and newest (50%) cases and controls. We note that prolonged sample storage would not impact clinical utilization of this test because in practice, only recently collected samples would be analyzed. Finally, we emphasize that although statistical evaluation supports the robustness of this model, it is essential to validate its clinical performance with an independent set of cases and controls before its clinical implementation, and these validation studies are currently underway.

5. Conclusions

Our findings demonstrate a robust model with high sensitivity and specificity in differentiating Stage I and II PDAC cases from high-risk and normal-risk controls. The model also performs well in diabetic, \geq 65 years, and low CA 19-9 producing sub-populations. Overall, these results justify advancement to the next phase of analytical and independent clinical validation studies. This biomarker panel represents an important step towards improving detection of early-stage PDAC and survival outcomes for this devastating disease.

Supplementary Materials: The following supporting information can be downloaded at: https: //www.mdpi.com/article/10.3390/cancers17111866/s1, Table S1: Sample collection sites; Table S2: Demographics of sub-populations; Figure S1: Study participant and sample selection; Figure S2: Analyte expression as a function of age; Figure S3: Analyte expression as a function of sex; Table S3: Comparison of serum candidate protein concentrations between cases and controls in two cohorts; Table S4: Expression levels and individual sensitivities of candidate markers in cohort sub-populations; Figure S4: Density plots showing analyte distribution between cases and controls within sub-populations; Figure S5: Performance of model v2 and model v3.

Author Contributions: Conceptualization, N.A.P., A.L.L., B.W.K., A.A., N.M.K., L.F., T.K. (Thomas Keller), S.W., R.S., T.K. (Thomas King), D.M.S. and R.B.; methodology, A.A., N.M.K., L.F., T.K. (Thomas Keller), S.W., R.S. and T.K. (Thomas King); software, A.A., N.M.K. and R.S.; validation, A.A., N.M.K., L.F., S.W., R.S. and T.K. (Thomas King); formal analysis, A.L.L., A.A., N.M.K., L.F., T.K. (Thomas Keller), S.W., R.S. and T.K. (Thomas King); investigation, A.A., N.M.K., S.W. and R.S.; resources, A.A and R.S.; data curation, A.A. and N.M.K.; writing—review and editing, N.A.P., A.L.L., B.W.K., A.A., N.M.K., L.F., T.K. (Thomas Keller), S.W., R.S. and T.K. (Thomas Keller), S.W., R.S., and R.S.; resources, A.A. and R.S.; data curation, A.A. and N.M.K.; writing—review and editing, N.A.P., A.L.L., B.W.K., A.A., N.M.K., L.F., T.K. (Thomas Keller), S.W., R.S., T.K. (Thomas King), D.M.S. and R.B.; supervision, N.A.P., R.S. and T.K. (Thomas King). All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Boards/Ethics Committees of all 12 study sites. The following is a list of approvals: (1) The University of Chicago Biological Sciences Division/University of Chicago Medical Center BSD IRB Committee A (Protocol #IRB18-1648, approved 3/12/19), (2) Massachusetts General Hospital IRB (Protocol #2018P001028, approved 5/13/21), (3) IRB of the Mount Sinai School of Medicine (Protocol #STUDY-22-00998, approved 9/1/20), (4) New York University Langone Health IRB (Protocol #i22-00937, approved 4/20/08), (5) The Ohio State University IRB (Protocol #2020C0010, approved 8/24/20), (6) Oregon Health & Science University Hospital IRB (Protocol #3609, approved 4/16/18), (7) University of Pittsburgh Human

Research Protection Office (Protocol #MOD19070256-039, approved 12/5/24), (8) Stanford University IRB (Protocol #40897, approved 6/18/19), (9) University of Massachusetts Medical School IRB (Protocol #H00015358, approved 8/27/20), (10) University of Pennsylvania (Protocol #831292, approved 4/15/20), (11) Columbia University Medical Center IRB (Protocol #IRB00010414, approved 2/18/18), and (12) Beth Israel Deaconess Medical Center IRB (Protocol #PLIRB, approved 3/19/21).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors thank the Principal Investigators of each study site for providing the samples analyzed in this study, Anja Wittig and Mei Wei Dutler for providing excellent technical assistance, and Laura Sommerville for providing exceptional professional writing services to draft and edit this manuscript. We wish to acknowledge the study coordinators and research personnel at the study sites, all study participants, and the clinical laboratory teams at Immunovia, Inc. and Proteomedix AG.

Conflicts of Interest: All authors are either employees or consultants of Immunovia, Inc.

Abbreviations

The following abbreviations are used in this manuscript:

CA 19-9	Carbohydrate Antigen 19-9
CI	Confidence Interval
CPB1	Carboxypeptidase B1
CTSD	Cathepsin D
GPNMB	Glycoprotein Non-Metastatic Melanoma Protein B
ICAM1	Intracellular Adhesion Molecule 1
KLK10	Kallikrein Related Peptidase 10
LTBP2	Latent Transforming Growth Factor-Beta Binding Protein 2
MERTK	Mer Tyrosine Kinase
MMP-9	Matrix Metalloproteinase 9
NET	Neutrophil Extracellular Trap
PDAC	Pancreatic Ductal Adenocarcinoma
PLA2G1B	Phospholipase A2 group 1B
TGF-β1	Transforming Growth Factor-Beta 1
THBS1	Thrombospondin 1
TIMP1	Tissue Inhibitor of Metalloproteinase 1

References

- Institute, National Cancer Cancer Stat Facts: Common Cancer Sites. 2024. Available online: https://seer.cancer.gov/statfacts/ html/common.html (accessed on 1 March 2025).
- Society, A.C. Survival Rates for Pancreatic Cancer. 2024. Available online: https://www.cancer.org/cancer/types/pancreaticcancer/detection-diagnosis-staging/survival-rates.html (accessed on 1 March 2025).
- 3. da Costa, W.L., Jr.; Oluyomi, A.O.; Thrift, A.P. Trends in the Incidence of Pancreatic Adenocarcinoma in All 50 United States Examined Through an Age-Period-Cohort Analysis. *JNCI Cancer Spectr.* **2020**, *4*, pkaa033. [CrossRef] [PubMed]
- Klein, A.P.; Lindstrom, S.; Mendelsohn, J.B.; Steplowski, E.; Arslan, A.A.; Bueno-de-Mesquita, H.B.; Fuchs, C.S.; Gallinger, S.; Gross, M.; Helzlsouer, K.; et al. An absolute risk model to identify individuals at elevated risk for pancreatic cancer in the general population. *PLoS ONE* 2013, *8*, e72311. [CrossRef] [PubMed]
- 5. Ryan, D.P.; Hong, T.S.; Bardeesy, N. Pancreatic adenocarcinoma. N. Engl. J. Med. 2014, 371, 1039–1049. [CrossRef] [PubMed]
- Society, A.C. Cancer Facts & Figures 2025. 2025. Available online: https://www.cancer.org/research/cancer-facts-statistics/allcancer-facts-figures/2025-cancer-facts-figures.html (accessed on 10 March 2025).

- Goggins, M.; Overbeek, K.A.; Brand, R.; Syngal, S.; Del Chiaro, M.; Bartsch, D.K.; Bassi, C.; Carrato, A.; Farrell, J.; Fishman, E.K.; et al. Management of patients with increased risk for familial pancreatic cancer: Updated recommendations from the International Cancer of the Pancreas Screening (CAPS) Consortium. *Gut* 2020, *69*, 7–17. [CrossRef]
- Daly, M.B.; Pal, T.; Maxwell, K.N.; Churpek, J.; Kohlmann, W.; AlHilli, Z.; Arun, B.; Buys, S.S.; Cheng, H.; Domchek, S.M.; et al. NCCN Guidelines(R) Insights: Genetic/Familial High-Risk Assessment: Breast, Ovarian, and Pancreatic, Version 2.2024. *J. Natl. Compr. Cancer Netw.* 2023, 21, 1000–1010. [CrossRef]
- Sawhney, M.S.; Calderwood, A.H.; Thosani, N.C.; Rebbeck, T.R.; Wani, S.; Canto, M.I.; Fishman, D.S.; Golan, T.; Hidalgo, M.; Kwon, R.S.; et al. ASGE guideline on screening for pancreatic cancer in individuals with genetic susceptibility: Summary and recommendations. *Gastrointest. Endosc.* 2022, 95, 817–826. [CrossRef]
- Zhang, L.; Sanagapalli, S.; Stoita, A. Challenges in diagnosis of pancreatic cancer. World J. Gastroenterol. 2018, 24, 2047–2060. [CrossRef]
- Locker, G.Y.; Hamilton, S.; Harris, J.; Jessup, J.M.; Kemeny, N.; Macdonald, J.S.; Somerfield, M.R.; Hayes, D.F.; Bast, R.C., Jr. ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J. Clin. Oncol.* 2006, 24, 5313–5327. [CrossRef]
- Mann, D.V.; Edwards, R.; Ho, S.; Lau, W.Y.; Glazer, G. Elevated tumour marker CA19-9: Clinical interpretation and influence of obstructive jaundice. *Eur. J. Surg. Oncol.* 2000, 26, 474–479. [CrossRef]
- 13. Katsanos, K.H.; Kitsanou, M.; Christodoulou, D.K.; Tsianos, E.V. High CA 19-9 levels in benign biliary tract diseases. Report of four cases and review of the literature. *Eur. J. Intern. Med.* **2002**, *13*, 132–135. [CrossRef]
- 14. Koprowski, H.; Herlyn, M.; Steplewski, Z.; Sears, H.F. Specific antigen in serum of patients with colon carcinoma. *Science* **1981**, 212, 53–55. [CrossRef] [PubMed]
- Orntoft, T.F.; Vestergaard, E.M.; Holmes, E.; Jakobsen, J.S.; Grunnet, N.; Mortensen, M.; Johnson, P.; Bross, P.; Gregersen, N.; Skorstengaard, K.; et al. Influence of Lewis alpha1-3/4-L-fucosyltransferase (FUT3) gene mutations on enzyme activity, erythrocyte phenotyping, and circulating tumor marker sialyl-Lewis a levels. *J. Biol. Chem.* 1996, 271, 32260–32268. [CrossRef] [PubMed]
- 16. Zhang, G.M.; Bai, S.M.; Zhang, G.M.; Ma, X.B. Reference intervals of carbohydrate antigen 19-9 in the apparently healthy adult population. *J. Clin. Lab. Anal.* **2018**, *32*, e22380. [CrossRef] [PubMed]
- Lee, S.P.; Sung, I.K.; Kim, J.H.; Lee, S.Y.; Park, H.S.; Shim, C.S. Usefulness of Carbohydrate Antigen 19-9 Test in Healthy People and Necessity of Medical Follow-up in Individuals with Elevated Carbohydrate Antigen 19-9 Level. *Korean J. Fam. Med.* 2019, 40, 314–322. [CrossRef]
- Kane, L.E.; Mellotte, G.S.; Mylod, E.; O'Brien, R.M.; O'Connell, F.; Buckley, C.E.; Arlow, J.; Nguyen, K.; Mockler, D.; Meade, A.D.; et al. Diagnostic Accuracy of Blood-based Biomarkers for Pancreatic Cancer: A Systematic Review and Meta-analysis. *Cancer Res. Commun.* 2022, 2, 1229–1243. [CrossRef]
- Nene, N.R.; Ney, A.; Nazarenko, T.; Blyuss, O.; Johnston, H.E.; Whitwell, H.J.; Sedlak, E.; Gentry-Maharaj, A.; Apostolidou, S.; Costello, E.; et al. Serum biomarker-based early detection of pancreatic ductal adenocarcinomas with ensemble learning. *Commun. Med.* 2023, *3*, 10. [CrossRef]
- 20. Al-Shaheri, F.N.; Alhamdani, M.S.S.; Bauer, A.S.; Giese, N.; Buchler, M.W.; Hackert, T.; Hoheisel, J.D. Blood biomarkers for differential diagnosis and early detection of pancreatic cancer. *Cancer Treat. Rev.* 2021, *96*, 102193. [CrossRef]
- Zhang, W.H.; Wang, W.Q.; Han, X.; Gao, H.L.; Li, T.J.; Xu, S.S.; Li, S.; Xu, H.X.; Li, H.; Ye, L.Y.; et al. Advances on diagnostic biomarkers of pancreatic ductal adenocarcinoma: A systems biology perspective. *Comput. Struct. Biotechnol. J.* 2020, 18, 3606–3614. [CrossRef]
- Athanasiou, A.; Kureshi, N.; Wittig, A.; Sterner, M.; Huber, R.; Palma, N.A.; King, T.; Schiess, R. Biomarker Discovery for Early Detection of Pancreatic Ductal Adenocarcinoma (PDAC) Using Multiplex Proteomics Technology. *J. Proteome Res.* 2025, 24, 315–322. [CrossRef]
- Bujang, M.A.; Adnan, T.H. Requirements for Minimum Sample Size for Sensitivity and Specificity Analysis. J. Clin. Diagn. Res. 2016, 10, YE01–YE06. [CrossRef]
- 24. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* **1988**, *44*, 837–845. [CrossRef] [PubMed]
- 25. Sundjaja, J.H.; Shrestha, R.; Krishan, K. McNemar And Mann-Whitney U Tests. 2024. Available online: https://www.ncbi.nlm. nih.gov/books/NBK560699/ (accessed on 1 February 2025).
- 26. The R Project for Statistical Computing. 2024. Available online: https://www.r-project.org/ (accessed on 1 February 2025).
- Silva-Santisteban, A.; Hernandez Woodbine, M.J.; Noriega, M.A.; Rabinowitz, L.G.; Grimshaw, A.; Farrell, J.J.; Chhoda, A.; Sawhney, M.S. Disparities in race, ethnicity, sex, and age inclusion in pancreatic cancer screening studies: A systematic review and meta-analysis. *Gastrointest. Endosc.* 2024, 100, 1–16.e20. [CrossRef] [PubMed]
- Ferlay, J.; Colombet, M.; Soerjomataram, I.; Parkin, D.M.; Pineros, M.; Znaor, A.; Bray, F. Cancer statistics for the year 2020: An overview. *Int. J. Cancer* 2021, 149, 778–789. [CrossRef] [PubMed]

- 29. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [CrossRef]
- 30. Network, National Comprehensive Cancer NCCN Guidelines Pancreatic Adenocarcenoma. 2024. Available online: https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1455 (accessed on 10 March 2025).
- Poruk, K.E.; Firpo, M.A.; Scaife, C.L.; Adler, D.G.; Emerson, L.L.; Boucher, K.M.; Mulvihill, S.J. Serum osteopontin and tissue inhibitor of metalloproteinase 1 as diagnostic and prognostic biomarkers for pancreatic adenocarcinoma. *Pancreas* 2013, 42, 193–197. [CrossRef]
- 32. Tian, Z.; Ou, G.; Su, M.; Li, R.; Pan, L.; Lin, X.; Zou, J.; Chen, S.; Li, Y.; Huang, K.; et al. TIMP1 derived from pancreatic cancer cells stimulates Schwann cells and promotes the occurrence of perineural invasion. *Cancer Lett.* **2022**, *546*, 215863. [CrossRef]
- Huang, C.; Li, N.; Li, Z.; Chang, A.; Chen, Y.; Zhao, T.; Li, Y.; Wang, X.; Zhang, W.; Wang, Z.; et al. Tumour-derived Interleukin 35 promotes pancreatic ductal adenocarcinoma cell extravasation and metastasis by inducing ICAM1 expression. *Nat. Commun.* 2017, *8*, 14035. [CrossRef]
- Roland, C.L.; Dineen, S.P.; Toombs, J.E.; Carbon, J.G.; Smith, C.W.; Brekken, R.A.; Barnett, C.C., Jr. Tumor-derived intercellular adhesion molecule-1 mediates tumor-associated leukocyte infiltration in orthotopic pancreatic xenografts. *Exp. Biol. Med.* 2010, 235, 263–270. [CrossRef]
- 35. Grunwald, B.; Harant, V.; Schaten, S.; Fruhschutz, M.; Spallek, R.; Hochst, B.; Stutzer, K.; Berchtold, S.; Erkan, M.; Prokopchuk, O.; et al. Pancreatic Premalignant Lesions Secrete Tissue Inhibitor of Metalloproteinases-1, Which Activates Hepatic Stellate Cells Via CD63 Signaling to Create a Premetastatic Niche in the Liver. *Gastroenterology* **2016**, *151*, 1011–1024.e7. [CrossRef]
- 36. Schoeps, B.; Eckfeld, C.; Prokopchuk, O.; Bottcher, J.; Haussler, D.; Steiger, K.; Demir, I.E.; Knolle, P.; Soehnlein, O.; Jenne, D.E.; et al. TIMP1 Triggers Neutrophil Extracellular Trap Formation in Pancreatic Cancer. *Cancer Res.* 2021, *81*, 3568–3579. [CrossRef]
- Liou, G.Y.; Doppler, H.; Necela, B.; Edenfield, B.; Zhang, L.; Dawson, D.W.; Storz, P. Mutant KRAS-induced expression of ICAM-1 in pancreatic acinar cells causes attraction of macrophages to expedite the formation of precancerous lesions. *Cancer Discov.* 2015, 5, 52–63. [CrossRef] [PubMed]
- Nomura, T.; Katunuma, N. Involvement of cathepsins in the invasion, metastasis and proliferation of cancer cells. *J. Med. Investig.* 2005, 52, 1–9. [CrossRef] [PubMed]
- Dumartin, L.; Whiteman, H.J.; Weeks, M.E.; Hariharan, D.; Dmitrovic, B.; Iacobuzio-Donahue, C.A.; Brentnall, T.A.; Bronner, M.P.; Feakins, R.M.; Timms, J.F.; et al. AGR2 is a novel surface antigen that promotes the dissemination of pancreatic cancer cells through regulation of cathepsins B and D. *Cancer Res.* 2011, *71*, 7091–7102. [CrossRef] [PubMed]
- 40. Albo, D.; Berger, D.H.; Tuszynski, G.P. The effect of thrombospondin-1 and TGF-beta 1 on pancreatic cancer cell invasion. *J. Surg. Res.* **1998**, *76*, 86–90. [CrossRef]
- Bocci, G.; Fioravanti, A.; Orlandi, P.; Di Desidero, T.; Natale, G.; Fanelli, G.; Viacava, P.; Naccarato, A.G.; Francia, G.; Danesi, R. Metronomic ceramide analogs inhibit angiogenesis in pancreatic cancer through up-regulation of caveolin-1 and thrombospondin-1 and down-regulation of cyclin D1. *Neoplasia* 2012, *14*, 833–845. [CrossRef]
- 42. McElroy, M.K.; Kaushal, S.; Tran Cao, H.S.; Moossa, A.R.; Talamini, M.A.; Hoffman, R.M.; Bouvet, M. Upregulation of thrombospondin-1 and angiogenesis in an aggressive human pancreatic cancer cell line selected for high metastasis. *Mol. Cancer Ther.* **2009**, *8*, 1779–1786. [CrossRef]
- 43. Qian, X.; Rothman, V.L.; Nicosia, R.F.; Tuszynski, G.P. Expression of thrombospondin-1 in human pancreatic adenocarcinomas: Role in matrix metalloproteinase-9 production. *Pathol. Oncol. Res.* **2001**, *7*, 251–259. [CrossRef]
- 44. Matsumura, K.; Hayashi, H.; Uemura, N.; Ogata, Y.; Zhao, L.; Sato, H.; Shiraishi, Y.; Kuroki, H.; Kitamura, F.; Kaida, T.; et al. Thrombospondin-1 overexpression stimulates loss of Smad4 and accelerates malignant behavior via TGF-beta signal activation in pancreatic ductal adenocarcinoma. *Transl. Oncol.* **2022**, *26*, 101533. [CrossRef]
- 45. mohamed, A.; Saad, Y.; Saleh, D.; Elawady, R.; Eletreby, R.; Kharalla, A.S.; Badr, E. Can Serum ICAM 1 distinguish pancreatic cancer from chronic pancreatitis? *Asian Pac. J. Cancer Prev.* **2016**, *17*, 4671–4675.
- Park, H.D.; Kang, E.S.; Kim, J.W.; Lee, K.T.; Lee, K.H.; Park, Y.S.; Park, J.O.; Lee, J.; Heo, J.S.; Choi, S.H.; et al. Serum CA19-9, cathepsin D, and matrix metalloproteinase-7 as a diagnostic panel for pancreatic ductal adenocarcinoma. *Proteomics* 2012, 12, 3590–3597. [CrossRef]
- 47. Jenkinson, C.; Elliott, V.L.; Evans, A.; Oldfield, L.; Jenkins, R.E.; O'Brien, D.P.; Apostolidou, S.; Gentry-Maharaj, A.; Fourkala, E.O.; Jacobs, I.J.; et al. 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. [CrossRef] [PubMed]
- 48. Luo, G.; Liu, C.; Guo, M.; Cheng, H.; Lu, Y.; Jin, K.; Liu, L.; Long, J.; Xu, J.; Lu, R.; et al. Potential Biomarkers in Lewis Negative Patients With Pancreatic Cancer. *Ann. Surg.* **2017**, *265*, 800–805. [CrossRef] [PubMed]
- Ando, Y.; Dbouk, M.; Yoshida, T.; Saba, H.; Abou Diwan, E.; Yoshida, K.; Dbouk, A.; Blackford, A.L.; Lin, M.T.; Lennon, A.M.; et al. Using Tumor Marker Gene Variants to Improve the Diagnostic Accuracy of DUPAN-2 and Carbohydrate Antigen 19-9 for Pancreatic Cancer. J. Clin. Oncol. 2024, 42, 2196–2206. [CrossRef] [PubMed]

- 50. Luo, G.; Guo, M.; Jin, K.; Liu, Z.; Liu, C.; Cheng, H.; Lu, Y.; Long, J.; Liu, L.; Xu, J.; et al. Optimize CA19-9 in detecting pancreatic cancer by Lewis and Secretor genotyping. *Pancreatology* **2016**, *16*, 1057–1062. [CrossRef]
- 51. Katona, B.W.; Klute, K.; Brand, R.E.; Everett, J.N.; Farrell, J.J.; Hawthorne, K.; Kaul, V.; Kupfer, S.S.; Paiella, S.; Simeone, D.M.; et al. Racial, Ethnic, and Sex-based Disparities among High-risk Individuals Undergoing Pancreatic Cancer Surveillance. *Cancer Prev. Res.* **2023**, *16*, 343–352. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.