HHS Public Access

Author manuscript

Int J Cancer. Author manuscript; available in PMC 2024 September 15.

Published in final edited form as:

Int J Cancer. 2023 September 15; 153(6): 1201–1216. doi:10.1002/ijc.34624.

External validation of genetically predicted protein biomarkers for pancreatic cancer risk using aptamer-based plasma levels: A prospective analysis in the Atherosclerosis Risk in Communities Study

Tanxin Liu[1], **Corinne E. Joshu**[1,2], **Jiayun Lu**[1], **Anna Prizment**[3,4], **Nilanjan Chatterjee**[2,5], **Josef Coresh**[1,5,6], **Lang Wu**[7], **Elizabeth A. Platz**[1,2]

[1]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

[2]Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, Maryland, USA

[3]Division of Hematology, Oncology and Transplantation, University of Minnesota, Minneapolis, Minnesota, USA

[4]University of Minnesota Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota, USA

[5]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

[6]Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins Medicine, Baltimore, Maryland, USA

[7]University of Hawaii Cancer Center, University of Hawaii at Manoa, Honolulu, Hawaii, USA

Abstract

Genetically predicted proteins have been associated with pancreatic cancer risk previously. We aimed to externally validate the associations of 53 candidate proteins with pancreatic cancer risk

Correspondence: Elizabeth A. Platz, Department of Epidemiology Johns Hopkins Bloomberg School of Public Health 615 N. Wolfe St., Room E6132 Baltimore, MD 21205, USA. eplatz1@jhu.edu, Lang Wu, Cancer Epidemiology Division, Population Sciences in the Pacific Program, University of Hawaii Cancer Center, University of Hawaii at Manoa, 701 Ilalo Street, Building B, Room 520, Honolulu, HI 96813, USA. lwu@cc.hawaii.edu.
Lang Wu and Elizabeth A. Platz are co-senior authors.

AUTHOR CONTRIBUTIONS

The work reported in the article has been performed by the authors, unless clearly specified in the text. **Tanxin Liu**: Formal Analysis, Visualization, Writing—original draft. **Corinne Joshu**: Resources, Funding acquisition, Investigation, Writing – review & editing. **Jiayun Lu:** Data curation, Writing – review & editing. **Anna Prizment:** Writing – review & editing. **Nilanjan Chatterjee**: Writing – review & editing. Josef Coresh: Funding acquisition, Investigation, Resources, Writing – review & editing. **Lang Wu**: Conceptualization, Supervision, Writing – review & editing. **Elizabeth Platz:** Conceptualization, Funding acquisition, Investigation, Supervision, Writing – review & editing.

CONFLICT OF INTEREST STATEMENT

Josef Coresh is the scientific advisor for SomaLogic. The other authors do not have a conflict of interest to declare.

ETHICS STATEMENT

The ARIC study protocol was approved by the Institutional Review Boards (IRB) of all participating centers. Informed consent was obtained from each study participant.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

using directly measured, prediagnostic levels. We conducted a prospective cohort study of 10 355 US Black and White men and women in the Atherosclerosis Risk in Communities (ARIC) study. Aptamer-based plasma proteomic profiling was previously performed using blood collected in 1993 to 1995, from which the proteins were selected. By 2015 (median: 20 years), 93 incident pancreatic cancer cases were ascertained. Cox regression was used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for protein tertiles, and adjust for age, race, and known risk factors. Of the 53 proteins, three were statistically significantly, positively associated with risk—GLCE (tertile 3 vs 1: HR = 1.88, 95% CI: 1.12–3.13; *P*-trend = 0.01), GOLM1 (aptamer 1: HR = 1.98, 95% CI: 1.16–3.37; *P*-trend = 0.01; aptamer 2: HR = 1.86, 95% CI: 1.07–3.24; *P*-trend = 0.05), and QSOX2 (HR = 1.96, 95% CI: 1.09–3.58; *P*-trend = 0.05); two were inversely associated—F177A (HR = 0.59, 95% CI: 0.35–1.00; *P*-trend = 0.05) and LIFsR (HR = 0.55, 95% CI: 0.32–0.93; *P*-trend = 0.03); and one showed a statistically significant lower risk in the middle tertile—endoglin (HR = 0.50, 95% CI: 0.29–0.86); by chance, we expected significant associations for 2.65 proteins. FAM3D, IP10, sTie-1 (positive); SEM6A and JAG1 (inverse) were suggestively associated with risk. Of these 11, 10 proteins—endoglin, FAM3D, F177A, GLCE, GOLM1, JAG1, LIFsR, QSOX2, SEM6A and sTie-1—were consistent in direction of association with the discovery studies. This prospective study validated or supports 10 proteins as associated with pancreatic cancer risk.

Keywords

biomarkers; pancreatic cancer; proteins; risk

1 | INTRODUCTION

Pancreatic cancer is the seventh leading cause of cancer death world-wide, accounting for almost as many deaths (466 003) as cases (495 773).[1] It was the third leading cause of cancer death (estimated deaths: 49830; 8%) in the United States in 2022, with a 5-year relative survival of 11%.[2] Most people with pancreatic cancer are diagnosed at an incurable stage.[3] This is mainly due to patients developing late-stage, nonspecific symptoms, lack of effective screening tests and challenges in diagnosis.[3] The burden of pancreatic cancer has brought primary prevention to the forefront. Cigarette smoking, diabetes, physical inactivity, greater body fatness and chronic pancreatitis are established risk factors for pancreatic cancer, which suggest dysregulated inflammatory pathways in pancreatic carcinogenesis.[4] The tissue microenvironment, including inflammatory, plays an important role in tumor development, progression and metastasis.[5] Identifying circulating biomarkers associated with pancreatic cancer risk, whether marking etiologic exposures or an inflamed or injured pancreatic tissue microenvironment, has the potential to improve risk stratification and allow for targeted prevention strategies and surveillance during the preclinical phase. Circulating plasma biomarkers may provide suggestive information regarding causal biological pathways contributing to pancreatic carcinogenesis and shed light on future pathway-specific and personalized interception strategies for precancer.

The only established circulating pancreatic cancer biomarker, carbohydrate antigen (CA 19–9), has several challenges involving false positives in conditions of nonpancreatic cancers

and false negatives in Lewis-negative individuals.[6,7] Thus, CA 19–9 is used for monitoring patients after a pancreatic cancer diagnosis and treatment, but not for early detection and not for risk stratification for prevention. Additional proteins, including CA242, PIVKA-II, EphA2, PAM4 and RBM6, have been identified based on comparing patients with and without pancreatic cancer, but these proteins lack the desired sensitivity and specificity for early detection,[8–12] and due to the study designs used, do not inform risk stratification for prevention. A compendium of potential biomarkers of pancreatic cancer referenced 2325 papers and reported 930 secreted proteins that were overexpressed in pancreatic cancers relative to normal pancreatic tissue; while this approach may inform biomarkers of early detection, it does not inform biomarkers of risk.[12]

With respect to the discovery of novel circulating protein biomarkers of pancreatic cancer risk, a prospective case-cohort study in Japan evaluated plasma concentrations of 62 chemokines, cytokines and growth factors and found no significant associations with pancreatic cancer risk.[5] A case-cohort study measuring 92 proteins using the OLINK panel in 610 cases and a subcohort of 623 individuals in China identified some chemokines, interleukins, growth factors and membrane proteins associated with pancreatic ductal adenocarcinoma.[3] Moreover, differences were identified in the markers associated with short-term (ie, markers of an undetected pancreatic cancer) vs long-term risk (ie, markers of the future development of pancreatic cancer)[3]; here, short vs long-term refers to time since blood draw before cancer diagnosis. To determine whether the protein markers associated with long-term risk are causally related to pancreatic cancer risk or are due to the long natural history of disease (reverse causation), several studies have been conducted employing genetic instruments for protein levels.[3]

Recently, two studies used genetic instruments to predict circulating protein concentrations and future pancreatic cancer risk[13,14] using data, in part, from the INTERVAL study.[15] The INTERVAL study, a genetic atlas of the human plasma proteome using an aptamer-based multiplex protein assay (SomaScan), identified 1927 genotype-protein associations (pQTLs) between 1478 proteins and 764 genomic loci from a total of 3301 healthy European descent individuals.[15] Zhu and colleagues then used the INTERVAL pQTL data[15] combined with GWAS data in the Pancreatic Cancer Cohort Consortium (PanScan) and the Pancreatic Cancer Case-Control (PANC4) Consortium, which included 8280 pancreatic cancer cases and 6728 controls of European ancestry, to investigate the associations between genetically predicted concentrations of protein biomarkers and pancreatic cancer risk.[13] A total of 38 proteins were statistically significantly associated with pancreatic cancer at a false discovery rate of <0.05.[13] Eight of the 38 proteins remained significant after adjusting for known pancreatic cancer risk variants identified in previous GWAS.[13] A follow-up study used genetic and proteomic data from the INTERVAL study to develop comprehensive protein genetic prediction models and identified 40 proteins that were significantly associated with pancreatic cancer risk at a false discovery rate of <0.05, 13 of which were novel.[14]

In total, 53 protein candidates have been identified as potential biomarkers for pancreatic cancer risk using genetic instruments. In terms of causal inference for pancreatic cancer risk, these studies relied on a statistical-based selection of genetically regulated components

of plasma protein levels, not measured levels and thus may be susceptible to pleiotropic effects.[16]

Therefore, we conducted a prospective cohort study to externally validate the 53 proteins with pancreatic cancer risk using prediagnostic protein levels (aptamer-measured) in the Atherosclerosis Risk in Communities (ARIC) study. The study included ~20 years of follow up of more than 10 000 Black and White men and women from four field centers in the United States between 1996 and 2015.

2 | METHODS

2.1 | Study design

Details of the ARIC (RRID: SCR_021769) study design, objectives and baseline response rates are described elsewhere.[17,18] In brief, ARIC is a prospective community-based cohort study that recruited 15 792 adults 45 to 64 years old in 1987 to 1989 from four communities: Forsyth County, North Carolina; Jackson, Mississippi; Minneapolis, Minnesota; and Washington County, Maryland. The baseline and six follow-up visits (in 1990–1992, 1993–1995, 1996–1998, 2011–2013, 2016–2017 and 2018–2019) included interviews, laboratory measurements and clinical examinations. At baseline, participants were interviewed about their demographics, lifestyle factors and medical history, and received an extensive examination, including measurement of anthropometrics by trained personnel.

For the current analyses, participants were excluded if they (a) did not attend Visit 3 (n = 2931); (b) had prevalent cancer prior to Visit 3 (n = 1197); (c) were not Black or White (n = 36) or (d) were Black from Minneapolis or Washington County (n = 35); (e) did not have proteomic profiling at Visit 3 (n = 1225); (f) had missing values for BMI (n = 3) or education level (n = 16). After these exclusions, 10 355 participants comprised the analytic cohort.

2.2 | Protein measurement

Briefly, a panel of ~5000 plasma proteins or protein complexes was measured in plasma collected at ARIC Visit 3 using an aptamer-based capture array (SomaLogic).[19] From that panel, we selected the 53 candidate proteins identified in previous studies as having a significant association with pancreatic cancer risk (Table S1).[13,14] At each field center at each visit, plasma was collected according to a standardized protocol, frozen at −80°C and shipped on dry ice to the ARIC central laboratory, where it was stored frozen. After thawing, the plasma was aliquoted into barcoded microtiter plates.[19] The plates for Visit 3 (N = 12 887) were sent to SomaLogic for quantification.[19] Subsequently, plates for Visit 2 (N = 14 348) were sent to SomaLogic. We used Visit 3 as the start of follow-up for the main analysis and took advantage of the two time points to explore changes over the 3 years from Visit 2 to 3 in relation to pancreatic cancer risk.

A total of 5284 modified aptamers were used to measure relative fluorescence units (RFU) for a standard plasma volume per participant.[19] All candidate proteins in our study were measured using one aptamer, except Golgi membrane protein 1 (GOLM1), which was

measured using two aptamers that targeted overlapping amino acids (aptamer 1: amino acids 36–401, aptamer 2: amino acids 109–214).

The SomaScan assay, SomaLogic quality control and ARIC quality control procedures have been described in detail elsewhere.[19,20] All proteins were $\log_2$ transformed.

2.3 | Covariate assessment

A priori, covariates were selected for this analysis if they are known or suspected pancreatic cancer risk factors. Data on demographics (sex [male/female]), race [Black/White], education (less than high school, high school graduate/vocational school, at least some college) were reported at baseline. Because race and field centers are highly linked, we classified participants as White from Minneapolis, White from Washington County, White from Forsyth County, Black from Jackson, or Black from Forsyth County. All other covariates were assessed at Visit 3 (ie, concurrent with plasma protein measurement), including age (years); body mass index (BMI; kg/m^2); cigarette smoking status (current, former, never); alcohol consumption (current, former, never); diagnosed diabetes (self-reported physician diagnosis or used diabetes medications), undiagnosed diabetes (fasting glucose ≥126 mg/dL or nonfasting 200 mg/dL), at risk for diabetes (fasting glucose 100 to <126 mg/dL or nonfasting glucose of 140-<200 mg/dL; each vs no); and family history of cancer (yes, no/do not know). Pack-years of smoking were calculated based on duration of smoking, current smoking status and number of cigarettes per day reported.

2.4 | Ascertainment of pancreatic cancer incidence and death

Incident pancreatic cancers were ascertained from 1987 through 2015 by linkage with cancer registries in the four ARIC states, supplemented by routine abstraction of medical records and collection of hospital discharge summaries for self-reported cases.[18] We contacted participants who reported at a visit or on an annual or semiannual (since 2012) follow-up telephone call that they had a cancer diagnosis for more information on their diagnosis and requested and reviewed their medical records. Cancer mortality was ascertained through 2015 from death certificates where cancer was the underlying cause of death. First primary pancreatic cancer incidence and mortality were identified using the following ICD-codes: ICD-O-1: 157.0–157.9; ICD-O-2: 25.0–25.9; ICD-O-3: C25.0-C25.9; ICD-8: 157.0–157.9; ICD-9: 157.0–157.9; ICD-10: 25.0–25.9. Among the analytic cohort, from Visit 3 to 2015, 93 first primary pancreatic cancer cases and 98 pancreatic cancer deaths were ascertained. Of the incident cases, 16 were first identified by death certificate, and 8 cases were identified by hospital discharge summary or self-report during annual follow-up calls.

2.5 | Statistical analysis

Statistical analyses were conducted using SAS 9.4 (RRID: SCR_008567). Figures were plotted using R 4.1.1 (RRID: SCR_001905). All statistical tests were 2-sided, and $P < .05$ was considered to be statistically significant. A priori we did not propose to adjust for multiple testing because the goal was external validation of each candidate protein.

We estimated age-adjusted percentages or means of characteristics at Visit 3 by cancer status using linear regression (continuous variables) or logistic regression (categorical variables). We calculated Spearman rank correlation coefficients between the proteins adjusting for age. The Wilcoxon rank-sum test was used to test for differences in protein concentration after $\log_2$ transformation. Multivariable Cox proportional hazards regression with time on study as the timescale was used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) of incident pancreatic cancer or pancreatic cancer death in association with the 53 candidate proteins. Participants contributed person-time at risk from Visit 3 until the diagnosis of a first primary pancreatic cancer (or pancreatic cancer death), a first primary cancer other than pancreatic (incidence only), death due to another cause or administrative censoring on December 31, 2015, whichever came first. We entered into the model tertiles of each protein RFU with cut-points based on the distribution in the whole analytic cohort. We tested for trends across tertiles by modeling a single ordinal variable using the medians of the tertiles, the coefficient for which we evaluated using the Wald test. Model 1 was adjusted for age (continuous), sex, joint terms for race by center and education. Model 2 was additionally adjusted for cigarette smoking status; packyears smoked (continuous); alcohol consumption (status); BMI (continuous); diagnosed diabetes, undiagnosed diabetes, at risk for diabetes status; and family history of cancer. Missing indicator variables were used for missing pack-years smoked (n = 684) and missing family history of cancer (n = 628). Standard tests using Schoenfeld residuals suggested no evidence of violation of the proportional hazards assumption. To assess the shape of the dose-response associations, we modeled the association using restricted cubic splines, with three knots placed at the 10th, 50th and 90th percentiles of the distribution of protein RFU in the analytic cohort. We tested for linearity by performing a likelihood ratio test comparing the linear model (protein entered as a continuous variable) and the spline model.

We stratified by race (Black, White), sex (female, male) and also by Visit 3 (baseline for the analysis) BMI [normal weight (18.5 to <25), overweight (25 to <30), obese (≥30 kg/m^2)] and smoking status (ever, never) to investigate potential effect modification. We used the same cut-points for the proteins as in the full analytic cohort. We used the likelihood ratio test comparing models with and without the cross-product term between a protein and the potential effect modifier to test for interaction.

We conducted several sensitivity analyses. We excluded the first 5 years of follow-up after Visit 3 to assess the possibility of reverse causation. Among participants who had proteins measured at both Visits 2 and 3, we performed the following analyses with follow-up starting at Visit 3: (a) lagged the analysis by 3 years by assigning Visit 2 protein levels to Visit 3 to further assess the possibility of reverse causation; (b) assessed the association of mean of Visit 2 and 3 protein levels and (c) assessed the association of an increase or decrease in $\log_2$ protein levels from Visit 2 to 3 using stable levels as the comparison.

3 | RESULTS

Among 10 355 participants in the analytic cohort, mean age was 60.0 (SD 5.7) years, 54% were female and 22% were Black (Table 1). During the median of 20 years of follow-up (167 142 person-years), 93 first primary pancreatic cancer cases were ascertained.

Of the 53 proteins, unadjusted, prediagnostic levels of 4 proteins—D-glucuronyl C5-epimerase (GLCE; $P = .03$), Golgi membrane protein 1 (GOLM1) (aptamer 1: $P = .04$; aptamer 2: $P = .01$), leukemia inhibitory factor receptor (LIFsR; $P = .04$) and sulfhydryl oxidase 2 (QSOX2; $P = .06$) differed between participants who were subsequently diagnosed with pancreatic cancer and those who were not (Table 2). GOLM1 levels measured by the two aptamers were strongly correlated (Spearman partial $r = .83$). Among the 53 proteins, 29 protein pairs were moderately correlated when taking into account age ($r > |.5|$; Figure S1).

3.1 | Prediagnostic plasma protein levels and pancreatic cancer risk

Associations between levels of the 53 plasma proteins and pancreatic cancer risk were similar before (Model 1) and after further adjusting for pancreatic cancer risk factors (Model 2; Table S2). Of the 53 proteins, in the fully-adjusted model (Model 2; Table 3), 3 were statistically significantly positively associated—GLCE (tertile 3 vs 1: HR = 1.88, 95% CI: 1.12–3.13; *P*-trend = 0.01), GOLM1 (aptamer 1: HR = 1.98, 95% CI: 1.16–3.37; *P*-trend = 0.01; aptamer 2 HR = 1.86, 95% CI: 1.07–3.24; *P*-trend = 0.05) and QSOX2 (HR = 1.96, 95% CI:[1.09–3.58]; *P*-trend = 0.05). Two of the proteins were inversely associated with pancreatic cancer risk—F177A (HR = 0.59, 95% CI: 0.35–1.00; *P*-trend = 0.05) and LIFsR (HR = 0.55, 95% CI: 0.32–0.93; *P*-trend = 0.03). For 1 protein—endoglin—the HR for tertile 2 (vs tertile 1) was statistically significant and inverse (0.50, 95% CI: 0.29–0.86), while the HR for tertile 3 was not significant but was <1.0 (HR = 0.85, 95% CI: 0.52–1.38). We expected associations for 2.65 proteins to be statistically significant by chance alone. Five other proteins were suggestively associated with pancreatic cancer risk (based on the *P*-trend or 95% CI; Table 3) in the positive direction—FAM3D, C-X-C motif chemokine 10 (IP10) and tyrosine-protein kinase receptor Tie-1, soluble (sTie-1) or inverse direction—protein jagged-1 (JAG1) and semaphorin-6A (SEM6A). Ten of the 11 proteins that were significantly or suggestively associated with risk—endoglin, FAM3D, F177A, GLCE, GOLM1, JAG1, LIFsR, QSOX2, SEM6A and sTie-1—were consistent in the direction of the association with the discovery studies.

Of the 11 proteins that were significantly or suggestively associated with risk, endoglin was moderately to strongly correlated (Spearman partial $r > .6$) with both JAG1 and LIFsR, and SEM6A was moderately to strongly correlated with LIFsR (Spearman partial $r > .6$). When adjusted for endoglin, the inverse associations for JAG1 and LIFsR remained, while for endoglin, the HR for the middle tertile became less inverse and the top tertile shifted from below to above 1. We also mutually adjusted LIFsR and JAG1 (Spearman partial $r > .55$) and the inverse association for LIFsR remained, while the association for JAG1 remained inverse but was attenuated. When mutually adjusted, the inverse associations for both LIFsR and SEM6A remained inverse, but were attenuated.

We investigated the shape of the dose-response for the associations (natural logarithm scale) for the 11 proteins with pancreatic cancer risk using restricted cubic splines (Figure 1). QSOX2 showed an inverted U-shaped relationship with pancreatic cancer ($P = .02$). While the tests for linearity were all $P > .15$ for the other 10 proteins, most showed nonlinear associations with pancreatic cancer risk (on the ln scale).

3.2 | Effect modification

SEM6A level did not differ between White and Black participants ($P = .64$). Levels of 42 of the 53 proteins statistically significantly differed between male and female participants. Of the 11 proteins, compared to male participants, female participants had higher levels of IP10 and sTie1, and lower levels of endoglin, GLCE, GOLM1 (both aptamers), JAG1, LIFsR, QSOX2 and SEM6A ($P < .01$). F177A and FAM3D level did not differ between female and male participants (both $P > .50$).

Given the difference in risk of pancreatic cancer by race and sex, we explored effect modification by race and sex for the 11 proteins that were significantly or suggestively associated with pancreatic cancer risk (Table 4). The only statistically significant interaction at the $P < .05$ level was for sex and FAM3D (P-interaction <.01). Possible interactions ($P \leq .10$) were also observed for race and F177A, GOLM1-aptamer 2 and IP10; and sex and JAG1; smoking and sTie-1.

3.3 | Sensitivity analyses

The sensitivity analysis (a), in which we excluded the first 5 years of follow-up (77 pancreatic cancer cases, 164 618 person-years), yielded HRs that were consistent in direction and not substantially attenuated relative to the main analysis (Table S4). In the sensitivity analyses among participants with proteins at both Visits 2 and 3 (9125 participants, 85 pancreatic cancer cases, 148 589 person-years), we observed the following results. (b) When lagging Visit 2 protein level to Visit 3 (Table S5) and (c) when using the mean of Visit 2 and 3 protein levels (Table S6), the HRs were consistent in direction with the overall analysis except for endoglin. When modeling the percent increase or decrease in $\log_2$ level from Visit 2 to 3 (compared to stable level), participants with >1% change (increase or decrease) in IP10 level between visits had a lower risk of pancreatic cancer, and those with >1% decrease in LIFsR had a higher risk of pancreatic cancer (Table S7).

3.4 | Prediagnostic protein levels and pancreatic cancer mortality

During a median of 21 years of follow-up (186 914 person-years), 98 pancreatic cancer deaths were documented. In the fully adjusted model, patterns of associations with pancreatic cancer mortality were largely consistent with the incidence analysis (Table S8). Statistically significant associations with pancreatic cancer mortality were observed for 5 of the 11 proteins that were significantly or suggestively associated with risk—GLCE, GOLM1 (aptamer 1), LIFsR, QSOX2 and sTie-1.

4 | DISCUSSION

This prospective cohort analysis in the ARIC study aimed to externally validate the associations of 53 candidate genetically predicted proteins using aptamer-measured levels with pancreatic cancer. Of the 53 proteins, 11 proteins were significantly or suggestively associated with pancreatic cancer risk. Associations were in the same direction for pancreatic cancer incidence and mortality. GOLM1 levels detected by two aptamers that target overlapping amino acids were highly correlated and showed similar positive associations with pancreatic cancer. Ten of the 11 proteins that were significantly or

suggestively associated with risk—endoglin, FAM3D, F177A, GLCE, GOLM1, JAG1, LIFsR, QSOX2, SEM6A and sTie-1 were consistent in the direction of association with the discovery studies. Given that we studied proteins that were the top hits from studies of the associations of genetically predicted proteins with pancreatic cancer risk, our findings support that the validated or supported proteins are mediators on the pathway from SNPs to pancreatic cancer risk. If confirmed, these 10 proteins could be investigated for etiology and for use in risk stratification for surveillance and precancer interception.

Our study provides information about possible protein biomarkers of pancreatic cancer risk. With respect to etiology, some proteins may be made in a different level as a response to known, suspected or even unknown risk factors for pancreatic cancer. Etiologic proteins may be useful in risk stratification if they have sufficient prediction. Also, with respect to risk stratification, some proteins may be made by the pancreas if it is injured or made in response to an injured pancreas (eg, inflammation and wound healing). If that injury increases the risk of subsequent pancreatic cancer, those proteins could be useful biomarkers for risk stratification for surveillance. With respect to early detection, some proteins may be made or made in different amounts than usual by pancreatic cancer tumors or precursors before they are diagnosed; some of these proteins could be detectable in circulation.[3] In our study, we used various strategies to inform whether the confirmed proteins most likely reflect exposures or steps in pancreatic carcinogenesis. To address etiology, we adjusted for known or suspected risk factors to determine whether the proteins were marking the risk factors but not themselves associated with pancreatic cancer. All 11 proteins were still statistically significantly or suggestively associated with pancreatic cancer risk after adjustment, suggesting that (a) the proteins are etiologically linked with pancreatic cancer risk or (b) that the proteins are marking unknown risk factors but themselves are not associated with pancreatic cancer. Related to (a), the proteins may be mediating the causal pathway from the known factors to pancreatic cancer. To address etiology vs proteins produced by a tumor or in response to a tumor, in a sensitivity analysis, we excluded the first 5 years after the blood draw to determine the possibility that pancreas damage or an undiagnosed pancreatic cancer is influencing baseline levels of the proteins. In that sensitivity analysis, associations for 11 of the proteins were not attenuated, suggesting that reverse causation is unlikely. Thus, these proteins may be useful for investigating etiology, and for determining whether they may contribute to risk prediction, and thus may be useful for risk stratification.

Table S9 compares the findings from the ARIC study using aptamer-measured protein levels to those from the two discovery studies using genetically predicted protein levels.[13,14] Of the 10 proteins that were consistent in the direction of association with one or both of the discovery studies and were significant or suggestive in ARIC, FAM3D, GLCE, GOLM1 and sTie-1 were positively associated, endoglin, F177A, LIFsR, JAG1 and SEM6A were inversely associated, and QSOX2 showed a U-shaped relationship with pancreatic cancer risk. While conventionally showing null associations (ie, HRs modest and not significant), an additional 12 proteins were consistent in the direction of the observed HRs in one or both of the discovery studies; and of these 8 had HRs >1, suggesting a pattern different from chance. IP10 was positively associated with risk in ARIC, but genetically predicted levels were inversely associated in one or both of the discovery studies.

In Zhu et al associations for the 5 proteins—endoglin, F177A, GLCE, JAG1 and LIFsR—of the 10 proteins that were statistically significant or suggestive and consistent in direction between ARIC and Zhu et al, were each substantially attenuated when they adjusted for previously identified pancreatic cancer risk variants.[13] This suggests that these proteins may mediate the association between the previously identified risk SNPs and pancreatic cancer risk; other explanations are possible, including genetic confounding and chance.[13]

Of the 10 proteins that were significantly or suggestively associated with pancreatic cancer risk in ARIC and in the same direction as in one or both of the discovery genetically predicted protein studies, LIFsR is the only protein that was consistent in the direction in all three studies (inverse) and was statistically significant in all three studies (Table S9).[13,14] LIFsR was strongly correlated with endoglin and SEM6A (Spearman partial $r >$ .6), but after their mutual adjustment, LIFsR remained inversely associated. LIFsR was also moderately correlated with a number of proteins, but those proteins, other than JAG1, were not statistically significantly or suggestively associated with pancreatic cancer risk. Because endoglin was strongly associated with both LIFsR and JAG1 (Spearman partial $r >$ .6), and JAG1 was suggestively associated with pancreatic cancer risk, we also mutually adjusted them, and LIFsR remained inverse while JAG1 was attenuated. These analyses support that LIFsR is the protein explaining the inverse association rather than a correlated protein among the candidates.

LIFsR is the soluble form of the leukemia inhibitory factor receptor (LIF-R). LIFRβ forms a heterodimer with gp130 to form a membrane-bound receptor. Its ligand, LIF is a member of the interleukin-6 family. This ligand-receptor complex signals through the JAK/STAT pathway, and has roles in human embryogenesis and carcinogenesis.[21,22] For pancreatic cancer, studies support that signaling by this complex is involved in pancreatic tumor growth, but suppression of pancreatic metastases.[12] Like other soluble cytokine receptors, LIFsR binds LIF and prevents this cell signaling.[22] While we do not know the cellular source of LIFsR, we sought to determine whether our findings are due to reverse causation (eg, level is influenced by an undiagnosed tumor and thus level appears to be associated with risk). LIFsR remained inversely associated with risk when using multiple statistical approaches to assess reverse causation within a 3- to 5-year window. When addressing change in level over 3 years before the start of follow-up, participants whose LIFsR level decreased had a higher risk of pancreatic cancer compared to those whose level remained stable. Within the context of what is known about LIFsR and given the mediation findings from Zhu et al, our findings suggest a possible role for LIFsR in pancreatic cancer development, and thus risk. Nevertheless, we cannot rule out that LIFsR is marking later moments in the long preclinical phase of pancreatic cancer.

In addition to LIFsR, some of the other proteins that were statistically significantly or suggestively associated have been investigated in relation to pancreatic neoplasia in the context of early detection, diagnosis and prognosis.[12,23–27] More research is needed to understand the potential biological consequences of the associations between the other 10 proteins and the risk of the development of pancreatic cancer.

There are several possible explanations for why some proteins identified in the discovery studies by Zhu et al[13] and Ghoneim et al[14] were not confirmed in ARIC (Table S9). First, circulating protein concentrations are influenced by both inherent (genetic) and environmental (extrinsic) factors, and thus, the use of the genetically regulated components of circulating protein concentrations, by using genetically predicted levels, may not have captured extrinsic factors.[28] For example, in Ghoneim et al, they were not able to incorporate some established covariates (eg, smoking, BMI, diabetes) for adjustment during protein-genetic prediction model construction.[14] In contrast, we focused on measured levels of proteins in circulation (genetic and extrinsic). Second, Zhu et al relied on protein quantitative trait loci (pQTLs) in GWAS, and thus the explained variation in protein concentration may be limited by the number of known pQTLs.[13] In contrast, our use of measured proteins likely captures all genetic regulators as well as extrinsic regulators. Third, genetic instruments for a protein of interest as in Zhu et al[13] and Ghoneim et al[14] could be associated with other proteins (pleiotropy), and those other proteins rather than the proteins of interest may be involved in pancreatic carcinogenesis.[29] Our use of measured levels of the 53 candidates precludes the possibility that some genetically associated proteins lie up- or downstream of the candidate protein levels and mediate in the pancreatic cancer risk as part of the same causal pathway.[30]

4.1 | Strengths and limitations

Strengths of our study include the use of a high-throughput proteomic technology,[19] a large community cohort, inclusion of Black and White women and men, prospective design, approximate 20-year follow-up and validity of cancer outcome and covariates data, which were ascertained without respect to protein measurements.

Nevertheless, our results should be interpreted within the context of several possible limitations. First, we have not confirmed whether the proteins or their levels measured using the SomaLogic aptamer-based method in this cohort would be consistent with those measured using ELISA or other conventional methods. Additional work using mass spectrometry for accurate mass determination, protein characterization and plasma concentration (rather than RFU) is needed. Second, given the number of pancreatic cancer cases in our study, we had low power to detect modest to moderate associations and effect modification, and were unable to investigate associations by tumor stage and histologic type. Third, we are unable to determine whether the proteins may mediate the causal pathway from known pancreatic cancer risk factors to pancreatic cancer risk vs whether the risk factors are confounders of the association between the proteins and pancreatic cancer risk. Fourth, if the risk factors are confounders, even though we adjusted for BMI, smoking status, pack-years, alcohol consumption and diabetes status, we cannot rule out residual confounding of the association between proteins and pancreatic cancer risk by factors that are already known or strongly suspected risk factors for pancreatic cancer. Finally, we cannot rule out the role of chance, although by chance, we expected significant associations for 2.65 of the 53 proteins, and we observed that 6 were statistically significant in the primary analysis (Table 3).

In summary, we validated or identified support for 10 proteins, including LIFsR, that were previously identified in studies using genetic instruments. These proteins were associated with pancreatic cancer risk when using aptamer-based levels in a prospective cohort study in ARIC. Findings from our study support previous evidence for prediagnostic protein biomarkers of pancreatic cancer risk. If confirmed, these proteins could be investigated for etiology and use in risk stratification for surveillance and interception for precancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors thank the staff and participants of the ARIC study for their important contributions. Cancer data was provided by the Maryland Cancer Registry, Center for Cancer Prevention and Control, Maryland Department of Health, with funding from the State of Maryland and the Maryland Cigarette Restitution Fund. The collection and availability of cancer registry data are also supported by the Cooperative Agreement NU58DP006333, funded by the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the Department of Health and Human Services.

FUNDING INFORMATION

The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung and Blood Institute, National Institutes of Health, Department of Health and Human Services, under Contract nos. (75N92022D00001, 75N92022D00002, 75N92022D00003, 75N92022D00004, 75N92022D00005). Studies on cancer in ARIC are also supported by the National Cancer Institute (U01 CA164975). SomaLogic Inc. conducted the SomaScan assays in exchange for use of ARIC data. This work was also supported in part by NIH/NHLBI grant R01 HL134320. The content of this work is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

DATA AVAILABILITY STATEMENT

ARIC data can be accessed with the permission of the ARIC committee (https://sites.cscc.unc.edu/aric/description). ARIC data also are available via BioLINCC (https://biolincc.nhlbi.nih.gov/studies/aric/). Further information is available from the corresponding author upon request.

Abbreviations:

ARIC	atherosclerosis risk in communities
BMI	body mass index
CA 19–9	carbohydrate antigen 19–9
CI	confidence interval
EphA2	ephrin type-A receptor 2
F177A	family with sequence similarity 177, member A1
FAM3D	family with sequence similarity 3, member D
GLCE	D-glucuronyl C5-epimerase

GOLM1	golgi membrane protein 1
GWAS	genome-wide association study
HR	hazard ratio
IP10	C-X-C motif chemokine 10
JAG1	protein jagged-1
LIFsR	leukemia inhibitory factor receptor
PAM4	clivatuzumab
PIVKA-II	protein induced by vitamin K absence-II
pQTL	protein quantitative trait loci
QSOX2	sulfhydryl oxidase 2
RBM6	RNA-binding protein 6
RFU	relative fluorescence unit
SEM6A	semaphorin-6A
SNP	single nucleotide polymorphism
sTie-1	tyrosine-protein kinase receptor Tie-1, soluble

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–249. [PubMed: 33538338]
2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022 Jan. CA Cancer J Clin. 2022;72(1):7–33. [PubMed: 35020204]
3. Kartsonaki C, Pang Y, Millwood I, et al. Circulating proteins and risk of pancreatic cancer: a case-subcohort study among Chinese adults. Int J Epidemiol. 2022;51(3):817–829. [PubMed: 35064782]
4. World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Expert Report 2018. Diet, Nutrition, Physical Activity and Pancreatic Cancer. dietandcancerreport.org
5. Ma E, Shimazu T, Song M, et al. Circulating inflammation markers and pancreatic cancer risk: a prospective case-cohort study in Japan. Cancer Epidemiol Biomarkers Prev. 2022;31(1):236–241. [PubMed: 34697062]
6. Fahrmann JF, Schmidt CM, Mao X, et al. Lead-time trajectory of CA19–9 as an anchor marker for pancreatic cancer early detection. Gastroenterology. 2021;160(4):1373–1383. [PubMed: 33333055]
7. Young MR, Wagner PD, Ghosh S, et al. Validation of biomarkers for early detection of pancreatic cancer: summary of the Alliance of pancreatic cancer consortia for biomarkers for early detection workshop. Pancreas. 2018;47(2):135–141. [PubMed: 29346214]
8. Tartaglione S, Pecorella I, Zarrillo SR, et al. Protein induced by vitamin K absence II (PIVKA-II) as a potential serological biomarker in pancreatic cancer: a pilot study. Biochem Med (Zagreb). 2019;29(2):020707. [PubMed: 31223261]
9. Duan B, Hu X, Fan M, et al. RNA-binding motif protein 6 is a candidate serum biomarker for pancreatic cancer. Proteomics Clin Appl. 2019;13(5):e1900048. [PubMed: 31207145]

10. Loosen SH, Neumann UP, Trautwein C, Roderburg C, Luedde T. Current and future biomarkers for pancreatic adenocarcinoma. Tumour Biol. 2017;39(6):1010428317692231. doi:10.1177/1010428317692231 [PubMed: 28618958]
11. Koshikawa N, Minegishi T, Kiyokawa H, Seiki M. Specific detection of soluble EphA2 fragments in blood as a new biomarker for pancreatic cancer. Cell Death Dis. 2017;8(10):e3134. [PubMed: 29072678]
12. Harsha HC, Kandasamy K, Ranganathan P, et al. A compendium of potential biomarkers of pancreatic cancer. PLoS Med. 2009;6(4):e1000046. [PubMed: 19360088]
13. Zhu J, Shu X, Guo X, et al. Associations between genetically predicted blood protein biomarkers and pancreatic cancer risk. Cancer Epidemiol Biomarkers Prev. 2020;29(7):1501–1508. [PubMed: 32439797]
14. Ghoneim D, Zhu JJ, Ha JW, et al. Associations of genetically predicted blood protein biomarkers with pancreatic ductal adenocarcinoma risk: a study using comprehensive protein genetic prediction models. Cancer Res. 2021;2021:81.
15. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. Nature. 2018;558(7708):73–79. [PubMed: 29875488]
16. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey SG. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008;27(8):1133–1163. [PubMed: 17886233]
17. Wright JD, Folsom AR, Coresh J, et al. The ARIC (atherosclerosis risk In communities) study: JACC focus seminar 3/8. J Am Coll Cardiol. 2021;77(23):2939–2959. [PubMed: 34112321]
18. Joshu CE, Barber JR, Coresh J, et al. Enhancing the infrastructure of the atherosclerosis risk in communities (ARIC) study for cancer epidemiology research: ARIC cancer. Cancer Epidemiol Biomarkers Prev. 2018;27(3):295–305. [PubMed: 29263187]
19. Walker KA, Chen J, Zhang J, et al. Large-scale plasma proteomic analysis identifies proteins and pathways associated with dementia risk. Nat Aging. 2021;1:473–489. [PubMed: 37118015]
20. Kim CH, Tworoger SS, Stampfer MJ, et al. Stability and reproducibility of proteomic profiles measured with an aptamer-based platform. Sci Rep. 2018;8(1):8382. [PubMed: 29849057]
21. Halder S, Parte S, Kshirsagar P, et al. The pleiotropic role, functions and targeted therapies of LIF/LIFR axis in cancer: old spectacles with new insights. Biochim Biophys Acta Rev Cancer. 2022;1877(4):188737. [PubMed: 35680099]
22. Christianson J, Oxford JT, Jorcyk CL. Emerging perspectives on leukemia inhibitory factor and its receptor in cancer. Front Oncol. 2021;29(11):693724.
23. Prasad NB, Biankin AV, Fukushima N, et al. Gene expression profiles in pancreatic intraepithelial neoplasia reflect the effects of Hedgehog signaling on pancreatic ductal epithelial cells. Cancer Res. 2005;65(5):1619–1626. [PubMed: 15753353]
24. Pandey V, Fleming-Martinez A, Bastea L, et al. CXCL10/CXCR3 signaling contributes to an inflammatory microenvironment and its blockade enhances progression of murine pancreatic precancerous lesions. Elife. 2021;30(10):e60646.
25. Chung WC, Challagundla L, Zhou Y, Li M, Atfi A, Xu K. Loss of *Jag1* cooperates with oncogenic *Kras* to induce pancreatic cystic neoplasms. Life Sci Alliance. 2020;4(2):e201900503. [PubMed: 33268505]
26. Fukushima N, Sato N, Prasad N, Leach SD, Hruban RH, Goggins M. Characterization of gene expression in mucinous cystic neoplasms of the pancreas using oligonucleotide microarrays. Oncogene. 2004;23(56):9042–9051. [PubMed: 15489895]
27. Adjuto-Saccone M, Soubeyran P, Garcia J, et al. TNF-α induces endothelial-mesenchymal transition promoting stromal development of pancreatic adenocarcinoma. Cell Death Dis. 2021;12(7):649. [PubMed: 34172716]
28. Wu S, Zhu W, Thompson P, Hannun YA. Evaluating intrinsic and non-intrinsic cancer risk factors. Nat Commun. 2018;9(1):3490. [PubMed: 30154431]
29. Shu X, Bao J, Wu L, et al. Evaluation of associations between genetically predicted circulating protein biomarkers and breast cancer risk. Int J Cancer. 2020;146(8):2130–2138. [PubMed: 31265136]

30. Considine DPC, Jia G, Shu X, et al. Genetically predicted circulating protein biomarkers and ovarian cancer risk. Gynecol Oncol. 2021;160(2):506–513. [PubMed: 33246661]

What's new?

The use of genetic tools to predict proteins associated with pancreatic cancer risk may be influenced by pleiotropy, which could lead to the misidentification of proteins involved in pancreatic carcinogenesis. Here, the authors aimed to validate associations for 53 candidate proteins using directly measured, prediagnostic levels in a prospective cohort. Ten proteins were significantly or suggestively associated with pancreatic cancer risk. Positive and inverse associations between the 10 proteins and risk of pancreatic malignancy matched findings of previous studies. The findings validate previous work and suggest that the identified proteins are candidate markers for pancreatic cancer prevention and surveillance.

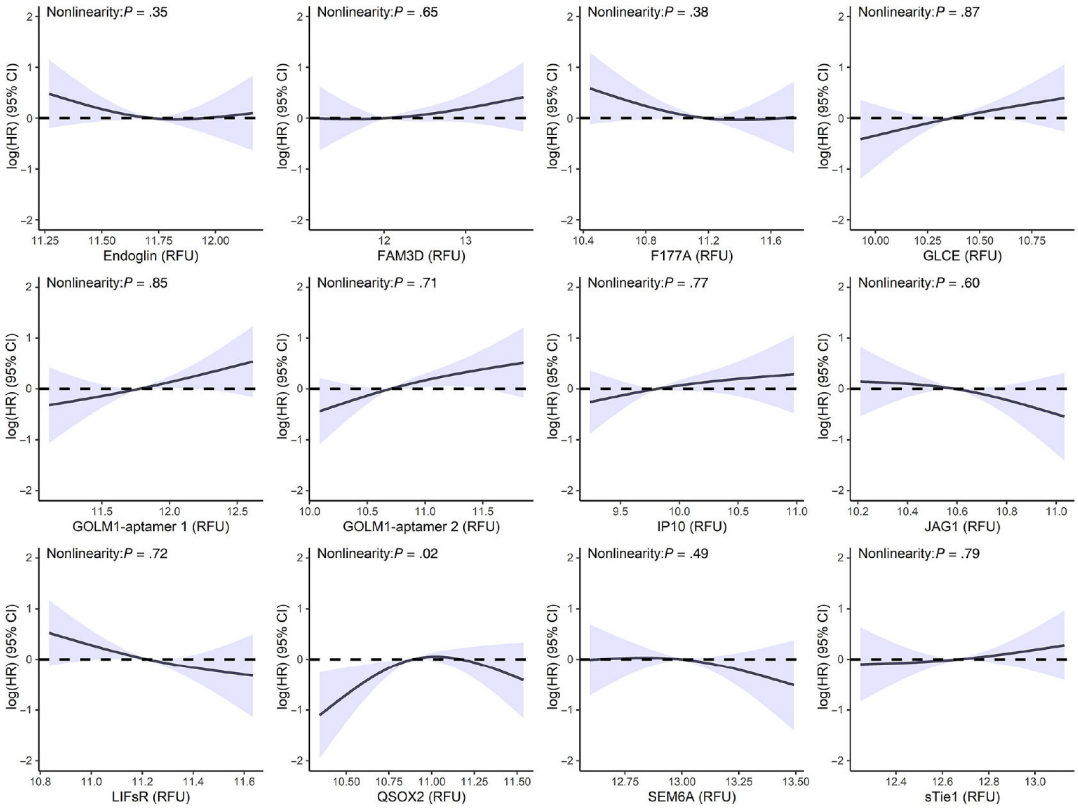

FIGURE 1.

Plots for the association between levels (RFU) of 11 plasma proteins* and pancreatic cancer risk, 10 355 participants in ARIC, Visit 3 (1993–1995) through 2015. The figure shows plots from Cox proportional hazards regression models using restricted cubic splines with three knots placed at the 10th, 50th and 90th percentiles of the distribution of the 11 plasma proteins RFU in the analytic cohort adjusted for age at Visit 3, sex, joint race by center and education, cigarette smoking status, packyears smoked at Visit 3, alcohol consumption, BMI, diagnosed diabetes, undiagnosed diabetes, at risk for diabetes status and family history of cancer. The 11 proteins were selected from the 53 candidate proteins identified based on References 13, 14 and were statistically significantly or suggestively associated with pancreatic cancer risk overall (Table 3). Note that the SomaScan panel includes two aptamers for GOLM1. Aptamer 1 targets amino acids 36 to 401. Aptamer 2 targets amino acids 109 to 214. Their Spearman partial correlation is .83 in the analytic cohort. The *P*-value for nonlinearity is from a likelihood ratio test comparing the fit of the linear model and the fit of the spline model; a high *P*-value supports that the association is linear on the natural logarithm scale of the HR. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 1

Age-adjusted characteristics[a] by subsequent incident pancreatic cancer status, ARIC, 1993 to 1995 (Visit 3).

		Subsequent pancreatic cancer status		
	Analytic cohort	**Case**	**Noncase**	***P*-value**
N	10 355	93	10 262	
Age, mean (SE, years)	60.0 (0.1)	60.0 (0.5)	60.0 (0.1)	.44
Female (%)	54.0	46.5	53.9	.15
Black (%)	21.9	32.5	21.6	.02
Field center (%)				.03
Forsyth County, NC	26.2	17.2	26.4	
Jackson, MS	19.3	30.4	18.9	
Minneapolis, MN	27.4	24.8	27.6	
Washington County, MD	27.1	27.6	27.1	
Education level (%)				.65
Less than high school	20.6	23.7	19.9	
High school graduate, vocational school	42.3	40.3	42.8	
College graduate and beyond	37.1	36.0	37.3	
Body mass index, mean (SE, kg/m^2)	28.6 (0.1)	28.8 (0.6)	28.6 (0.1)	.71
Cigarette smoking (%)				.90
Current smoker	17.9	19.4	17.7	
Former smoker	44.4	42.8	44.4	
Never smoker	37.7	37.8	37.9	
Packyears smoked among ever smokers[b], mean (SE)	28.7 (0.3)	26.0 (3.1)	28.7 (0.3)	.31
Alcohol consumption (%)				.63
Current drinker	52.7	55.2	52.6	
Former drinker	31.0	32.1	31.0	
Never drinker	16.4	12.7	16.4	
Diabetes status (%)				<.01
Diabetes, diagnosed diabetes	11.3	10.5	11.1	
Diabetes, undiagnosed	7.8	7.5	7.8	
At risk for diabetes	31.3	46.2	31.1	
Normoglycemic	49.6	35.8	50.0	
Family history of cancer[c] (%)	56.2	59.0	56.2	.58

[a] Linear or multinomial logistic regression models were used to estimate predicted probabilities adjusted for age as appropriate. Probabilities were multiplied by 100.

[b] Missing information for 684 participants.

[c] Missing information for 628 participants.

TABLE 2

Median (IQR) $\log_2$ transformed plasma levels (RFU) of 53 proteins[a] by subsequent pancreatic cancer status, ARIC, 1993 to 1995 (Visit 3).

		Subsequent pancreatic cancer status		
Protein	**Analytic cohort**	**Case**	**Noncase**	***P*-value**
N	10 355	93	10 262	
ADH1B	8.71 (0.24)	8.70 (0.20)	8.71 (0.24)	.32
Alkaline phosphatase, intestine	8.29 (0.48)	8.38 (0.39)	8.29 (0.48)	.22
B3GN2	11.56 (0.26)	11.54 (0.25)	11.56 (0.26)	.21
B4GT1	14.45 (0.25)	14.43 (0.22)	14.45 (0.25)	.10
BGAT	13.53 (5.47)	13.91 (5.66)	13.52 (5.47)	.11
C1GLC	9.78 (0.24)	9.75 (0.25)	9.78 (0.24)	.88
Cadherin-5	13.4 (0.32)	13.43 (0.39)	13.40 (0.32)	.29
Carboxypeptidase B1	11.22 (0.77)	11.24 (0.70)	11.22 (0.77)	.86
CD36 antigen	13.64 (0.70)	13.74 (0.58)	13.64 (0.70)	.26
CHSTB	12.04 (0.34)	12.07 (0.42)	12.04 (0.34)	.49
Chymotrypsin	8.88 (0.61)	8.86 (0.64)	8.88 (0.61)	.37
CRBB2	8.93 (0.27)	8.94 (0.29)	8.93 (0.27)	.77
DC-SIGN	10.89 (0.73)	10.90 (0.68)	10.89 (0.73)	.60
Desmoglein2	11.32 (0.40)	11.28 (0.41)	11.32 (0.40)	.54
DOCK9	8.77 (0.13)	8.76 (0.13)	8.77 (0.13)	.34
DPEP2	13.46 (0.38)	13.46 (0.43)	13.46 (0.37)	.90
Endoglin	11.72 (0.36)	11.65 (0.38)	11.72 (0.36)	.16
F177A	11.18 (0.46)	11.14 (0.49)	11.18 (0.46)	.34
FA20B	9.67 (0.24)	9.65 (0.23)	9.67 (0.24)	.69
FAM3B	11.64 (0.99)	11.55 (1.03)	11.64 (0.99)	.59
FAM3D	12.02 (1.25)	12.13 (1.22)	12.02 (1.25)	.32
GFRAL	9.70 (0.40)	9.67 (0.37)	9.70 (0.40)	.63
GLCE	10.36 (0.39)	10.44 (0.42)	10.36 (0.39)	.03
GOLM1–aptamer 1[b]	11.77 (0.61)	11.92 (0.59)	11.77 (0.61)	.04
GOLM1–aptamer 2[b]	10.70 (0.75)	10.81 (0.75)	10.70 (0.75)	.01

Protein	Analytic cohort	Subsequent pancreatic cancer status		*P*-value
		Case	Noncase	
GP116	10.40 (0.68)	10.30 (0.68)	10.40 (0.68)	.41
gp130, soluble	12.86 (0.22)	12.88 (0.16)	12.86 (0.22)	.69
IGFI sR	12.05 (0.35)	12.03 (0.33)	12.05 (0.35)	.66
IL23R	10.34 (0.31)	10.32 (0.32)	10.34 (0.31)	.49
IL-3Ra	12.61 (0.64)	12.70 (0.67)	12.61 (0.64)	.15
IP10	9.81 (0.57)	9.81 (0.50)	9.81 (0.57)	.41
IR	13.20 (0.49)	13.16 (0.49)	13.20 (0.49)	.48
ISLR2	12.02 (0.39)	12.06 (0.42)	12.02 (0.39)	.40
JAG1	10.59 (0.32)	10.56 (0.28)	10.59 (0.32)	.11
KIN17	8.16 (0.22)	8.17 (0.28)	8.16 (0.22)	.62
LIF-sR	11.22 (0.32)	11.15 (0.31)	11.22 (0.32)	.04
LMA2L	10.38 (0.51)	10.37 (0.63)	10.38 (0.50)	.70
Met	12.92 (0.32)	12.88 (0.31)	12.92 (0.32)	.30
Notch 1	13.96 (0.17)	13.94 (0.18)	13.96 (0.17)	.33
PSelectin	13.58 (0.48)	13.60 (0.48)	13.58 (0.48)	.38
QSOX2	10.89 (0.54)	10.96 (0.37)	10.89 (0.54)	.06
RAP	11.71 (0.34)	11.69 (0.31)	11.71 (0.34)	.44
sESelectin	14.48 (0.67)	14.41 (0.74)	14.48 (0.67)	.80
SEM6A	13.00 (0.33)	12.98 (0.25)	13.00 (0.33)	.37
ST4S6	12.02 (0.28)	11.99 (0.26)	12.02 (0.28)	.60
sTie-1	12.66 (0.35)	12.71 (0.32)	12.66 (0.35)	.29
sTie-2	10.61 (0.30)	10.60 (0.34)	10.61 (0.30)	.33
STOM	9.99 (0.18)	10.00 (0.17)	9.99 (0.18)	.90
TENC1	8.15 (0.16)	8.15 (0.16)	8.15 (0.16)	.61
THSD1	9.69 (0.30)	9.67 (0.27)	9.69 (0.30)	.75
TLL1	13.57 (0.39)	13.63 (0.42)	13.57 (0.39)	.28
TM11D	8.86 (0.27)	8.90 (0.32)	8.86 (0.27)	.30
TPST2	9.92 (0.20)	9.95 (0.22)	9.92 (0.20)	.10
VEGF sR2	12.58 (0.32)	12.56 (0.33)	12.58 (0.32)	.92

Abbreviation: RFU, relative fluorescence unit.

[a] Fifty-three candidate proteins were identified based on References [13, 14].

[b] The SomaScan® panel includes two aptamers for GOLM1. Aptamer 1 targets amino acids 36 to 401. Aptamer 2 targets amino acids 109 to 214. The Spearman partial correlation coefficient for these two measurements in the analytic cohort is .83.

TABLE 3

Association between levels (RFU) of 11 plasma protein[a] and pancreatic cancer risk, 10 355 participants in ARIC, Visit 3 (1993–1995) through 2015.

					Model 2[c]	
HR (95% CI)	**Pancreatic cancer cases**	**Person-years at risk (95% CI)**	**Median protein level in RFU (IQR)**	**Model 1[b], HR**	***P*-trend[d]**	
Endoglin						
Tertile 1	42	55 000	11.46 (0.19)	Ref	Ref	.43
Tertile 2	20	56 578	11.72 (0.11)	0.50 (0.29–0.85)	0.50 (0.29–0.86)	
Tertile 3	31	55 564	11.98 (0.19)	0.83 (0.51–1.35)	0.85 (0.52–1.38)	
Per 1 IQR					0.91 (0.69–1.21)	
F177A						
Tertile 1	37	55 060	10.83 (0.33)	Ref	Ref	.05
Tertile 2	29	56 038	11.18 (0.14)	0.71 (0.42–1.20)	0.71 (0.44–1.16)	
Tertile 3	27	56 044	11.50 (0.25)	0.59 (0.35–0.99)	0.59 (0.35–1.00)	
Per 1 IQR					0.78 (0.63–0.96)	
FAM3D						
Tertile 1	28	56 972	11.50 (0.31)	Ref	Ref	.21
Tertile 2	31	55 367	12.02 (0.33)	1.14 (0.68–1.91)	1.14 (0.68–1.92)	
Tertile 3	34	54 803	13.17 (0.63)	1.36 (0.82–2.26)	1.38 (0.83–2.29)	
Per 1 IQR					1.34 (0.98–1.83)	
GLCE						
Tertile 1	24	55 287	10.11 (0.18)	Ref	Ref	.01
Tertile 2	28	56 322	10.36 (0.12)	1.21 (0.70–2.08)	1.21 (0.70–2.09)	
Tertile 3	41	55 533	10.68 (0.24)	1.86 (1.11–3.10)	1.88 (1.12–3.13)	
Per 1 IQR					1.24 (1.00–1.53)	
GOLM1–aptamer 1[e]						
Tertile 1	22	58 825	11.36 (0.27)	Ref	Ref	.01
Tertile 2	30	55 645	11.77 (0.19)	1.38 (0.79–2.39)	1.42 (0.82–2.47)	
Tertile 3	41	52 671	12.24 (0.37)	1.92 (1.14–3.23)	1.98 (1.16–3.37)	
Per 1 IQR					1.27 (0.98–1.65)	
GOLM1–aptamer 2[e]						

HR (95% CI)	Pancreatic cancer cases	Person-years at risk (95% CI)	Median protein level in RFU (IQR)	Model 1[b], HR	Model 2[c]	
					P-trend[d]	
Tertile 1	20	58 909	10.28 (0.22)	Ref	Ref	.05
Tertile 2	37	54 826	10.70 (0.23)	1.89 (1.09–3.26)	1.90 (1.10–3.30)	
Tertile 3	36	53 407	11.32 (0.51)	1.85 (1.07–3.20)	1.86 (1.07–3.24)	
Per 1 IQR					1.22 (1.00–1.48)	
IP10						
Tertile 1	27	56 068	9.45 (0.23)	Ref	Ref	.10
Tertile 2	31	56 131	9.81 (0.18)	1.18 (0.70–1.98)	1.26 (0.75–2.13)	
Tertile 3	35	54 942	10.32 (0.56)	1.44 (0.86–2.39)	1.55 (0.92–2.60)	
Per 1 IQR					1.14 (0.94–1.39)	
JAG1						
Tertile 1	37	56 522	10.37 (0.16)	Ref	Ref	.08
Tertile 2	34	55 686	10.59 (0.10)	0.92 (0.58–1.47)	0.94 (0.59–1.50)	
Tertile 3	22	54 934	10.84 (0.19)	0.59 (0.35–1.01)	0.61 (0.36–1.05)	
Per 1 IQR					0.86 (0.65–1.13)	
LIF-sR						
Tertile 1	41	55 464	11.00 (0.17)	Ref	Ref	.03
Tertile 2	30	56 588	11.22 (0.10)	0.74 (0.46–1.19)	0.73 (0.46–1.18)	
Tertile 3	22	55 090	11.46 (0.18)	0.56 (0.33–0.95)	0.55 (0.32–0.93)	
Per 1 IQR					0.78 (0.60–1.02)	
QSOX2						
Tertile 1	17	55 866	10.55 (0.22)	Ref	Ref	.05
Tertile 2	42	55 809	10.89 (0.17)	2.34 (1.33–4.11)	2.35 (1.34–4.16)	
Tertile 3	34	55 467	11.28 (0.26)	1.92 (1.07–3.46)	1.96 (1.09–3.58)	
Per 1 IQR					1.24 (0.97–1.66)	
SEM6A						
Tertile 1	35	57 162	12.77 (0.17)	Ref	Ref	.08
Tertile 2	36	56 148	13.00 (0.10)	1.00 (0.62–1.59)	0.97 (0.61–1.56)	
Tertile 3	22	53 832	13.25 (0.22)	0.62 (0.35–1.07)	0.60 (0.34–1.05)	
Per 1 IQR					0.89 (0.68–1.16)	

Int J Cancer. Author manuscript; available in PMC 2024 September 15.

HR (95% CI)	Pancreatic cancer cases	Person-years at risk (95% CI)	Median protein level in RFU (IQR)	Model 1[b], HR	Model 2[c]	
					P-trend[d]	
sTie-1						
Tertile 1	25	55 633	12.42 (0.17)	Ref	Ref	.10
Tertile 2	32	56 265	12.66 (0.11)	1.34 (0.80–2.27)	1.35 (0.80–2.28)	
Tertile 3	36	55 244	12.93 (0.21)	1.51 (0.90–2.51)	1.54 (0.92–2.56)	
Per 1 IQR					1.18 (0.92–1.52)	

[a] Eleven of the 53 candidate proteins identified based on References 13, 14 were statistically significantly or suggestively associated with pancreatic cancer risk. See Table S2 for the associations for all 53 proteins.

[b] Model 1 was adjusted for age at Visit 3 (continuous), sex (female, male), joint race by center (White from Minneapolis [ref.], White from Washington Co., White from Forsyth Co., Black from Jackson, Black from Forsyth Co.) and education (less than high school, high school graduate/vocational school, college graduate/some graduate school/graduate degree).

[c] Model 2 was additionally adjusted for cigarette smoking status (current, former, never), packyears smoked at Visit 3 (continuous), alcohol consumption (current, former, never), BMI (continuous), diagnosed diabetes (yes, no), undiagnosed diabetes (yes, no), at risk for diabetes status (yes, no), family history of cancer (yes, no).

[d] We tested for trend across tertiles of protein level by assigning the median protein level in the tertile into which each participant fell.

[e] The SomaScan® panel includes two aptamers for GOLM1. Aptamer 1 targets amino acids 36 to 401. Aptamer 2 targets amino acids 109 to 214. The Spearman partial correlation coefficient for these two measurements in the analytic cohort is .83.

TABLE 4

Association between levels (RFU) of 11 plasma proteins[a] and pancreatic cancer risk by race, sex, smoking status and BMI, 10 355 participants in ARIC, Visit 3 (1993–1995) through 2015.

		Race					Sex				
		Black		White			Female		Male		
	Overall HR[b]	Cases	HR[b]	Cases	HR[b]	*P*-interaction	Cases	HR[b]	Cases	HR[b]	*P*-interaction
Endoglin											
Tertile 1	Ref	19	Ref	23	Ref	.90	22	Ref	20	Ref	.65
Tertile 2	0.50 (0.29–0.86)	6	0.58 (0.23–1.47)	14	0.47 (0.24–0.91)		7	0.39 (0.17–0.93)	13	0.62 (0.30–1.26)	
Tertile 3	0.85 (0.52–1.38)	5	0.78 (0.29–2.10)	26	0.86 (0.49–1.52)		14	0.87 (0.43–1.78)	17	0.88 (0.45–1.73)	
FAM3D											
Tertile 1	Ref	8	Ref	20	Ref	.21	6	Ref	22	Ref	<.01
Tertile 2	1.14 (0.68–1.92)	14	1.93 (0.80–4.68)	17	0.84 (0.44–1.60)		16	2.78 (1.08–7.17)	15	0.69 (0.35–1.34)	
Tertile 3	1.38 (0.83–2.29)	8	1.69 (0.62–4.60)	26	1.21 (0.67–2.17)		21	4.22 (1.68–10.62)	13	0.64 (0.32–1.28)	
F177A											
Tertile 1	Ref	12	Ref	25	Ref	.10	18	Ref	19	Ref	.97
Tertile 2	0.71 (0.44–1.16)	5	0.30 (0.11–0.86)	24	0.96 (0.55–1.69)		12	0.66 (0.31–1.37)	17	0.76 (0.39–1.47)	
Tertile 3	0.59 (0.35–1.00)	13	0.38 (0.17–0.84)	14	0.76 (0.39–1.47)		13	0.56 (0.27–1.18)	14	0.64 (0.31–1.32)	
GLCE											
Tertile 1	Ref	12	Ref	12	Ref	.76	11	Ref	13	Ref	.42
Tertile 2	1.21 (0.70–2.09)	9	0.99 (0.42–2.36)	19	1.39 (0.67–2.87)		11	1.02 (0.44–2.36)	17	1.40 (0.68–2.90)	

Tertile 3	1.88 (1.12–3.13)	9	1.27 (0.52–3.06)	32	2.21 (1.14–4.29)		21	2.25 (1.07–4.74)	20	1.56 (0.77–3.19)	
GOLMl-aptamer 1											
Tertile 1	Ref	6	Ref	16	Ref	.89	10	Ref	12	Ref	.87
Tertile 2	1.42 (0.82–2.47)	10	1.33 (0.48–3.70)	20	1.41 (0.73–2.74)		13	1.32 (0.58–3.05)	17	1.48 (0.70–3.12)	
Tertile 3	1.98 (1.16–3.37)	14	1.67 (0.62–4.49)	27	2.11 (1.12–3.98)		20	2.09 (0.95–4.61)	21	1.83 (0.89–3.79)	
GOLMl-aptamer 2											
Tertile 1	Ref	8	Ref	12	Ref	.07	7	Ref	13	Ref	.32
Tertile 2	1.90 (1.10–3.30)	8	0.74 (0.28–2.00)	29	2.85 (1.45–5.62)		17	2.46 (1.01–5.99)	20	1.58 (0.78–3.21)	
Tertile 3	1.86 (1.07–3.24)	14	1.40 (0.57–3.45)	22	2.10 (1.03–4.28)		19	2.95 (1.22–7.14)	17	1.26 (0.60–2.61)	
IP10											
Tertile 1	Ref	14	Ref	13	Ref	.10	10	Ref	17	Ref	.49
Tertile 2	1.26 (0.75–2.13)	6	0.56 (0.21–1.48)	25	1.95 (0.99–3.85)		12	1.12 (0.48–2.60)	19	1.41 (0.72–2.74)	
Tertile 3	1.55 (0.92–2.60)	10	1.10 (0.48–2.52)	25	2.09 (1.04–4.17)		21	1.90 (0.88–4.12)	14	1.24 (0.60–2.58)	
JAG1											
Tertile 1	Ref	10	Ref	27	Ref	.15	14	Ref	23	Ref	.10
Tertile 2	0.93 (0.58–1.49)	15	1.72 (0.77–3.85)	19	0.68 (0.38–1.22)		16	1.28 (0.62–2.65)	18	0.76 (0.41–1.43)	
Tertile 3	0.61 (0.36–1.05)	5	0.60 (0.20–1.78)	17	0.60 (0.32–1.12)		13	1.15 (0.53–2.49)	9	0.37 (0.17–0.81)	
LIF-sR											
Tertile 1	Ref	16	Ref	25	Ref	.55	23	Ref	18	Ref	.76
Tertile 2	0.73 (0.46–1.18)	10	0.91 (0.41–2.04)	20	0.67 (0.37–1.20)		12	0.63 (0.31–1.29)	18	0.87 (0.45–1.68)	

Tertile 3	0.55 (0.32–0.93)	4	0.39 (0.13–1.20)	18	0.60 (0.32–1.11)		8	0.46 (0.20–1.05)	14	0.65 (0.32–1.34)		
QSOX2												
Tertile 1	Ref	5	Ref	12	Ref	.93	10	Ref	7	Ref	.27	
Tertile 2	2.35 (1.34–4.16)	14	2.52 (0.90–7.06)	18	2.24 (1.13–4.43)		16	1.91 (0.86–4.22)	26	2.77 (1.20–6.42)		
Tertile 3	1.96 (1.09–3.58)	11	2.30 (0.79–6.74)	23	1.77 (0.88–3.59)		17	2.47 (1.13–5.43)	17	1.65 (0.68–3.99)		
SEM6A												
Tertile 1	Ref	13	Ref	22	Ref	.32	17	Ref	18	Ref	.44	
Tertile 2	0.97 (0.61–1.56)	8	0.67 (0.27–1.63)	28	1.11 (0.63–1.96)		17	1.08 (0.55–2.12)	19	0.86 (0.45–1.66)		
Tertile 3	0.60 (0.34–1.05)	9	0.78 (0.31–1.94)	13	0.53 (0.26–1.07)		9	0.69 (0.30–1.59)	13	0.54 (0.26–1.14)		
sTie-1												
Tertile 1	Ref	11	Ref	14	Ref	.17	9	Ref	16	Ref	.67	
Tertile 2	1.35 (0.80–2.28)	5	0.57 (0.20–1.67)	27	1.81 (0.95–3.45)		16	1.76 (0.78–4.00)	16	1.11 (0.55–2.23)		
Tertile 3	1.54 (0.92–2.56)	14	1.24 (0.56–2.75)	22	1.74 (0.89–3.41)		18	1.94 (0.87–4.34)	18	1.34 (0.68–2.64)		

	Smoking status					**Body mass index[c]**						
	Ever		**Never**			**Normal**		**Overweight**		**Obese**		
	Cases	**HR[b]**	**Cases**	**HR[b]**	***P*-interaction**	**Cases**	**HR[b]**	**Cases**	**HR[b]**	**Cases**	**HR[b]**	***P*-interaction**
Endoglin												
Tertile 1	25	Ref	17	Ref	.76	7	Ref	20	Ref	15	Ref	.42
Tertile 2	14	0.59 (0.31–1.15)	6	0.35 (0.14–0.89)		7	0.88 (0.30–2.57)	5	0.27 (0.10–0.73)	7	0.49 (0.20–1.21)	
Tertile 3	19	0.91 (0.49–1.60)	12	0.74 (0.34–1.61)		7	0.75 (0.25–2.21)	15	0.94 (0.47–1.90)	9	0.82 (0.35–1.93)	
FAM3D												

Tertile 1	16	Ref	12	Ref	.57	5	Ref	14	Ref	8	Ref	.36
Tertile 2	21	1.38 (0.71–2.67)	10	0.82 (0.35–1.92)		5	0.82 (0.23–2.91)	15	1.20 (0.58–2.52)	11	1.50 (0.60–3.78)	
Tertile 3	21	1.47 (0.76–2.84)	13	1.21 (0.54–2.68)		11	1.93 (0.65–5.66)	11	0.88 (0.39–1.95)	12	1.83 (0.73–4.55)	
F177A												
Tertile 1	22	Ref	15	Ref)	.92	9	Ref	15	Ref	13	Ref	.60
Tertile 2	19	0.76 (0.41–1.40)	10	0.65 (0.29–1.46)		7	0.64 (0.24–1.74)	11	0.67 (0.31–1.47)	11	0.90 (0.39–2.03)	
Tertile 3	17	0.61 (0.32–1.19)	10	0.58 (0.25–1.35)		5	0.35 (0.11–1.08)	14	0.82 (0.39–1.76)	7	0.50 (0.19–1.32)	
GLCE												
Tertile 1	16	Ref	8	Ref	.82	8	Ref	9	Ref	6	Ref	.31
Tertile 2	16	1.08 (0.54–2.18)	12	1.51 (0.61–3.74)		4	0.47 (0.14–1.59)	14	1.44 (0.62–3.34)	10	1.98 (0.71–5.47)	
Tertile 3	26	1.80 (0.95–3.39)	15	2.16 (0.90–5.17)		9	1.02 (0.39–2.69)	17	1.95 (0.85–4.45)	15	3.18 (1.22–8.31)	
GOLMl-aptamer 1												
Tertile 1	14	Ref	8	Ref	.94	7	Ref	11	Ref	4	Ref	.94
Tertile 2	18	1.30 (0.64–2.62)	12	1.66 (0.67–4.08)		6	1.15 (0.38–3.44)	14	1.32 (0.59–2.92)	9	1.61 (0.49–5.26)	
Tertile 3	26	1.84 (0.94–3.61)	15	2.16 (0.90–5.23)		18	1.98 (0.69–5.65)	15	1.54 (0.70–3.41)	18	2.58 (0.86–7.73)	
GOLMl-aptamer 2												
Tertile 1	6	Ref	14	Ref	.32	3	Ref	12	Ref	4	Ref	.34
Tertile 2	17	3.46 (1.35–8.87)	20	1.37 (0.69–2.73)		10	4.18 (1.14–15.42)	16	1.43 (0.67–3.06)	11	1.98 (0.63–6.26)	
Tertile 3	12	2.35 (0.87–6.34)	24	1.66 (0.85–3.24)		8	3.91 (1.02–14.93)	12	1.09 (0.48–2.45)	16	2.61 (0.86–7.89)	
IP10												

Tertile 1	18	Ref	9	Ref	.99	7	Ref	12	Ref	8	Ref	.97
Tertile 2	19	1.27 (0.66–2.43)	12	1.18 (0.49–2.82)		7	1.38 (0.47–4.07)	14	1.29 (0.59–2.84)	9	1.02 (0.39–2.68)	
Tertile 3	21	1.53 (0.81–2.91)	14	1.46 (0.62–3.44)		7	1.35 (0.45–4.04)	14	1.55 (0.70–3.42)	14	1.67 (0.68–4.09)	
JAG1												
Tertile 1	23	Ref	14	Ref	.35	5	Ref	17	Ref	14	Ref	.14
Tertile 2	24	1.10 (0.62–1.96)	10	0.62 (0.28–1.41)		11	1.97 (0.67–4.82)	16	0.96 (0.48–1.91)	7	0.51 (0.21–1.28)	
Tertile 3	11	0.51 (0.25–1.06)	11	0.69 (0.31–1.55)		5	0.76 (0.21–2.73)	7	0.43 (0.18–1.05)	10	0.86 (0.38–1.99)	
LIF-sR												
Tertile 1	24	Ref	17	Ref	.54	9	Ref	18	Ref	13	Ref	.24
Tertile 2	21	0.89 (0.49–1.61)	9	0.52 (0.23–1.18)		8	0.85 (0.32–2.25)	9	0.49 (0.22–1.10)	13	1.04 (0.48–2.26)	
Tertile 3	13	0.54 (0.27–1.09)	9	0.53 (0.23–1.22)		4	0.36 (0.11–1.21)	13	0.78 (0.37–1.62)	5	0.40 (0.14–1.16)	
QSOX2												
Tertile 1	9	Ref	8	Ref	.41	3	Ref	9	Ref	5	Ref	.26
Tertile 2	29	2.83 (1.33–6.02)	13	1.70 (0.70–4.13)		10	2.95 (0.80–10.93)	13	1.41 (0.60–3.32)	19	3.88 (1.44–10.51)	
Tertile 3	20	1.94 (0.88–4.29)	14	2.11 (0.88–5.10)		8	2.40 (0.63–9.17)	18	1.82 (0.80–4.11)	7	1.54 (0.48–4.90)	
SEM6A												
Tertile 1	23	Ref	12	Ref	.54	7	Ref	18	Ref	9	Ref	.33
Tertile 2	20	0.82 (0.44–1.49)	16	3.25 (1.19–8.93)		9	1.40 (0.51–3.87)	14	0.75 (0.37–1.53)	13	1.13 (0.48–2.67)	0.87
Tertile 3	15	0.59 (0.30–1.16)	7	2.94 (1.06–8.20)		5	0.89 (0.27–2.94)	8	0.44 (0.19–1.04)	9	0.72 (0.27–1.92)	
sTie-1												
Tertile 1	20	Ref	5	Ref	.06	6	Ref	7	Ref	11	Ref	.33

Tertile 2	16	0.85 (0.44–1.64)	16	3.25 (1.19–8.93)	7	1.17 (0.39–3.52)	16	2.57 (1.05–6.27)	9	0.82 (0.34–1.99)
Tertile 3	22	1.17 (0.63–2.14)	14	2.94 (1.06–8.20)	8	1.35 (0.46–3.96)	17	3.18 (1.31–7.71)	11	0.90 (0.39–2.10)

[a] Eleven of the 53 candidate proteins identified based on References 13, 14 were statistically significantly or suggestively associated with pancreatic cancer risk in the overall analysis (Table 3).

[b] Adjusted for age at Visit 3 (continuous), sex (female, male), joint race by center (White from Minneapolis [ref.], White from Washington Co., White from Forsyth Co., Black from Jackson, Black from Forsyth Co.) and education (less than high school, high school graduate/vocational school, college graduate/some graduate school/graduate degree), cigarette smoking status (current, former, never), packyears smoked at Visit 3 (continuous), alcohol consumption (current, former, never), BMI (continuous), diagnosed diabetes (yes, no), undiagnosed diabetes (yes, no), at risk for diabetes status (yes, no), family history of cancer (yes, no).

[c] Normal 18.5 ≤ BMI < 25 kg/m^2; overweight: 25 ≤ BMI < 30 kg/m^2; obese: BMI ≥ 30 kg/m^2. Excluded 78 participants with BMI < 18.5 kg/m^2.