



Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2020 May ; 29(5): 999–1008. doi:10.1158/1055-9965.EPI-19-1389.

Genetic and circulating biomarker data improve risk prediction for pancreatic cancer in the general population

Jihye Kim^{1,7}, Chen Yuan², Ana Babic², Ying Bao³, Clary Clish⁴, Michael N. Pollak⁵, Laufey T. Amundadottir⁶, Alison P. Klein^{7,8}, Rachael Z. Stolzenberg-Solomon⁹, Pari V. Pandharipande¹⁰, Lauren K. Brais², Marisa W. Welch², Kimmie Ng², Edward L. Giovannucci^{3,11,12}, Howard D. Sesso^{11,13}, JoAnn E. Manson^{11,13}, Meir J. Stampfer^{3,11,12}, Charles S. Fuchs^{14,15,16}, Brian M. Wolpin^{2,†}, Peter Kraft^{1,11,†}

¹Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. ²Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts. ³Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts. ⁴Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts. ⁵Cancer Prevention Research Unit, Department of Oncology, Faculty of Medicine, McGill University, Montreal, Quebec, Canada. ⁶Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland. ⁷Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins School of Medicine, Baltimore, Maryland. ⁸Department of Pathology, Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins School of Medicine, Baltimore, Maryland. ⁹Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland. ¹⁰Department of Radiology and Institute for Technology Assessment, Massachusetts General Hospital, Boston, Massachusetts. ¹¹Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. ¹²Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. ¹³Division of Prevention Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts. ¹⁴Department of Medical Oncology, Yale Cancer Center, New Haven, Connecticut. ¹⁵Department of Medicine, Yale School of Medicine, New Haven, Connecticut. ¹⁶Department of Medical Oncology, Smilow Cancer Hospital, New Haven, Connecticut.

Abstract

Background: Pancreatic cancer is the third leading cause of cancer death in the U.S., and 80% of patients present with advanced, incurable disease. Risk markers for pancreatic cancer have been characterized, but combined models are not used clinically to identify individuals at high risk for the disease.

Corresponding authors: Brian M. Wolpin. Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215. Phone: 617-632-6942; Fax: 617-632-5370; brian_wolpin@dfci.harvard.edu., Peter Kraft. Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115. Phone: 617-432-4271; pkraft@hsph.harvard.edu.

[†]Co-senior authors

Methods: Within a nested case-control study of 500 pancreatic cancer cases diagnosed after blood collection and 1,091 matched controls enrolled in four U.S. prospective cohorts, we characterized absolute risk models that included clinical factors (e.g., body-mass index, history of diabetes), germline genetic polymorphisms, and circulating biomarkers. Model discrimination showed an area under receiver operating characteristic curve of 0.62 via cross-validation.

Results: Our final integrated model identified 3.7% of men and 2.6% of women who had at least 3 times greater than average risk in the ensuing 10 years. Individuals within the top risk percentile had a 4% risk of developing pancreatic cancer by age 80 years and 2% 10-year risk at age 70 years.

Conclusions: Risk models include established clinical, genetic, as well as circulating biomarker factors improved disease discrimination over models using clinical factors alone.

Impact: Our absolute risk models for pancreatic cancer may help identify individuals in the general population appropriate for disease interception.

INTRODUCTION

Pancreatic cancer is the third leading cause of cancer-related mortality in the United States.¹ Incidence rates of pancreatic cancer continue to rise and 56,770 new cases are expected in 2019, such that pancreatic cancer is projected to become the second leading cause of cancer death in the U.S. within the next ten years.² The high mortality from pancreatic cancer is due in large part to late diagnosis, as nearly 80% of patients present with locally advanced or metastatic disease that is incurable.³ In contrast, patients diagnosed with localized, early stage pancreatic cancer can be cured using a combination of surgery, chemotherapy and radiation.⁴ Thus, identifying individuals at high risk of pancreatic cancer is of great importance, so that appropriate patients can be targeted for cancer prevention and earlier diagnosis.

Epidemiological studies from numerous distinct populations have identified demographic, lifestyle, and clinical factors associated with increased risk of pancreatic cancer. Firmly established risk factors include older age, male gender, African-American race/ethnicity, cigarette smoking, obesity, family history of pancreatic cancer, history of diabetes mellitus, and history of chronic pancreatitis.⁵⁻⁷ In nested prospective studies, future pancreatic cancer risk has been associated with circulating levels of several biomarkers related to insulin resistance (insulin, proinsulin, hemoglobin A1c,⁸⁻¹⁰ insulin-like growth factor binding protein 1,¹¹ 25-hydroxyvitamin D¹²), adipokines (adiponectin,^{13, 14} leptin^{15, 16}), inflammation (interleukin-6 [IL-6]¹⁷), and peripheral tissue catabolism (branched chain amino acids [BCAAs])¹⁸⁻²⁰.

Inherited genetic variants have been identified that predispose to development of pancreatic cancer. Medium to high-penetrance alterations have been found in several genes (e.g., *ATM*, *BRCA1*, *BRCA2*, *CDKN2A*, and *PALB2*) but these alterations are present in only 5 to 10% of patients with pancreatic cancer²¹⁻²³. Therefore, these gene mutations explain only a small fraction of the genetic risk for pancreatic cancer in the general population.²⁴ To identify common susceptibility loci, six large genome-wide association studies (GWAS) have been conducted in populations of European ancestry.²⁵⁻³⁰ To date, 18 susceptibility

loci carrying 22 independent single nucleotide polymorphisms (SNPs) have been identified surpassing the genome-wide significance threshold ($P < 5 \times 10^{-8}$).

Although risk factors have been investigated risk individually,^{31–33} their joint contribution to risk discrimination remain largely unknown. Therefore, we examined absolute risk models for pancreatic cancer that incorporate established clinical factors, common genetic predisposition variants, and circulating biomarkers in four large prospective cohorts. To estimate lifetime risk and 10-year risk, models were evaluated for the full nested case-control population and cases diagnosed within 10 years of blood collection and their matched controls, respectively.

Materials and Methods

Study Population

This study included participants from four large prospective cohort studies, the Health Professionals Follow-up Study (HPFS), Nurses' Health Study (NHS), Physicians' Health Study I (PHS I), and Women's Health Initiative (WHI) Observational Study. HPFS began enrollment of 51,529 male health professionals aged 40–75 years in 1986.³⁴ In NHS cohort, 121,701 female nurses aged 30–55 years began enrollment in 1976.³⁵ PHS I was a randomized clinical trial initiated in 1982 to examine effects of aspirin and B-carotene among 22,071 healthy male physicians aged 40–84 years. After trial completion in 1995, PHS I participants were followed up in an observational cohort.³⁶ In the WHI, 93,726 women aged 50–79 years enrolled between 1994 and 1998 to examine potential risk factors and causes of morbidity and mortality among postmenopausal women.³⁷

In this study, cases were incident patients with primary pancreatic adenocarcinoma ascertained between 1984 and 2010 through self-report, report of next-of-kin, or national death certificates and confirmed by medical record review and tumor registry data. All cases provided blood samples prior to their pancreatic cancer diagnosis and we randomly selected controls with matching on cohort (which also matches on sex), year of birth, smoking status, fasting status, and time of blood collection (month and year) with a matching ratio of 1:2 or 1:3. We excluded non-White participants, as GWAS risk variants were identified in subjects of European ancestry, and the strength of their association with pancreatic cancer in other populations requires further study. We also excluded participants who had complete missing data for questionnaires or blood samples or did not have matched counterparts. This study was approved by Human Research Committee at Brigham and Women's Hospital (Boston, MA), and participants of each cohort provided informed consent.

Lifestyle and clinical characteristics

Data on individual characteristics, such as lifestyles and medical history, were obtained by self-report on questionnaires, as previously reported.^{34–37} We used study questionnaires completed at or just prior to the blood draw in HPFS and NHS and the baseline questionnaire in PHS and WHI cohorts to collect data for age, sex, body mass index (BMI; kg/m^2), waist-to-hip ratio (WHR; inch/inch), physical activity (measured by MET-hour per week), and history of diabetes. Because WHR data were completely missing at baseline in

PHS, we incorporated post-baseline questionnaire data obtained at 108 months of follow-up in the cohort. We included WHR data of PHS only for cases (and matched controls) who were diagnosed after 108 months.

Blood collection and plasma assays

Blood samples were collected from 18,225 men in HPFS (1993–1995), 14,916 men in PHS I (1982–1984), 32,826 women in NHS (1989–1990), and 93,676 women in WHI (1994–1998). Details for blood processing and storage have been described previously.¹⁸ All cases and matched controls in our study provided blood samples prior to the case's diagnosis. Circulating levels of proinsulin (pM), adiponectin (ug/mL), IL-6 (pg/mL), and BCAAs (uM) were measured and represent four major categories of circulating markers related to pancreatic cancer risk (insulin resistance, adipokines, inflammation, and peripheral tissue catabolism, respectively). We dichotomized circulating adiponectin with a cutoff of 4.4 ug/ml as done previously.¹³ Details for laboratory assays and coefficients of variance (CVs) have been previously reported.^{8, 11–13, 15, 18} Coefficients of variance for blinded pooled plasma samples for all circulating markers were <11%.

DNA sequencing and SNP selection

Genomic DNA was extracted from peripheral blood leucocytes of cohort participants. Details on genotyping, variant imputation, and quality control procedures have been previously reported.³⁰ From the PanScan and PanC4 consortia GWAS,^{27–30} we included 22 SNPs that were associated with the risk of pancreatic cancer at genome-wide significance level ($P < 5 \times 10^{-8}$): rs13303010, rs10919791, rs2816938, rs1486134, rs9854771, rs2736098, rs31490, rs35226131, rs78417682, rs17688601, rs6971499, rs2941471, rs10094872, rs1561927, rs687289, rs9581943, rs9543325, rs7190458, rs4795218, rs11655237, rs1517037, and rs16986825 (Supplementary Table S1). SNP data were unavailable for participants not included in the consortia GWAS and were predominantly matched controls (Supplementary Table 4). Because of the missing SNP data, we imputed genotypes by randomly sampling from observed genotypes with replacement, conditional on study and case-control status. We calculated a weighted genetic risk score (wGRS) as the weighted sum of risk alleles using weights determined by the log-odds ratios reported in the PanScan and PanC4 consortia GWAS.^{27–30}

Statistical Analysis

To compare risk factor characteristics, we tabulated frequencies and distributions between cases and matched controls in cohort-specific and pooled analyses. We pooled data across the four cohorts; there was no evidence of substantial effect heterogeneity across the cohorts for most of risk factors ($P > 0.05$). Missing proportions of the non-genetic risk factors were ranged from 0.01 to 0.20. To minimize the effects of missing data, we used conditional mean imputation: we replaced missing values with the average value of each variable for each individual from the 25 imputed datasets generated using Multivariate Imputation by Chained Equations. All continuous variables were standardized with a mean of 0 and standard deviation (SD) of 1 in each cohort.

We first examined the associations between risk factors and pancreatic cancer in pooled univariable analyses using conditional logistic regression. Using multivariable conditional logistic regression, we then built three relative risk models for men and women separately including the following covariates: the first (“clinical model”) with BMI, WHR, MET-hour/week of physical activity, and history of diabetes (yes or no); the second (“clinical/genetic model”) added the wGRS to the clinical model; and the third (“clinical/genetic/biomarker model”) added proinsulin, adiponectin, IL-6, and total BCAAs to the clinical/genetic model. We compared goodness of fit of the three models using the likelihood ratio test. Risk models were built for all participants in the full follow-up population (maximum 26 years between data/blood collection and case diagnosis) and limited to cases diagnosed within 10 years of data/blood collection and their matched controls to allow evaluation of “lifetime” and 10-year absolute risks, respectively.

Model discrimination was assessed using the area under Receiver Operating Characteristic (ROC) curve analyses. To validate the discriminative performance of each model, we performed a 5-fold cross-validation leaving out 20% randomly selected data as a validation dataset and all the remaining data as a training dataset in our cohort data. Specifically, we randomly partitioned matched case-control sets into five equally-sized disjoint subsets, withheld each of the partitions in turn as a testing set, trained the models in the remaining data, and evaluated the area under the receiver operating characteristic curve (AUC) of the fit model in the testing set. We repeated this process over 20 different random partitions. We then calculated the average of AUC for each relative risk model over the resulting 100 test sets as a representative AUC of each model. We restricted validation samples to cases diagnosed within 10 years of blood collection and their matched controls because of the differences in the follow-up time across the four cohorts.

To calculate absolute risk for pancreatic cancer, we combined the multivariable relative risk models fit in our data with age- and sex-specific U.S. pancreatic cancer incidence rates, mortality rates, and the joint distribution of risk factors among U.S. non-Hispanic whites.^{40, 41} We included the effects of smoking and family history on pancreatic cancer risk in our absolute risk models using covariate-adjusted relative risks for these factors taken from the literature.^{42, 43}

To estimate the joint distribution of pancreatic cancer risk among U.S. non-Hispanic whites, we simulated 20,000 men and 20,000 women by first sampling smoking status based on the prevalence of smoking among white men and women (20.4% and 15.8%, respectively) in the U.S. general population (age-adjusted distributions for adults aged 18 and over from the National Health Interview Survey data, 2011–2014).⁴² We then sampled remaining clinical, genetic, and biomarker risk factors (except family history) by drawing a risk factor profile at random (with replacement) from male controls and female controls separately, conditional on smoking status. Finally, we sampled family history conditional on polygenic risk score (PRS), the sum of risk alleles of SNPs associated with pancreatic cancer, assuming the population prevalence of positive family history of pancreatic cancer is 3.6%.⁴³

Then we calculated individualized relative risk for each simulated subject on the basis of personal risk profile as follows:

$$RR(X) = \exp\left\{\sum_{i=1\dots k} \beta_i^T X_i\right\}$$

where X_1, X_2, \dots, X_k are an individual's risk factor values and $\beta_1, \beta_2, \dots, \beta_k$ are the log odds ratios (OR) for the risk factors in our risk models and literature estimates for current smoking and family history of pancreatic cancer.^{5, 43}

We calculated absolute risks of pancreatic cancer by combining the estimated relative risk with age- and sex-specific average incidence rates for non-Hispanic whites in U.S. Surveillance, Epidemiology and End Results (SEER) 17 from 2001 to 2005 (<http://seer.cancer.gov/>) and competing mortality risks obtained from U.S. mortality data of white men and women in 2007.⁴⁴ Using these data, we converted relative risks to absolute risks (p) as follows:⁴⁰

$$p(a, s) = \sum_{t=a+1}^{s-1} F(t)RR(x)\lambda_0(t).$$

Here $p(a, s)$ denotes the probability that a subject who is pancreatic-cancer-free at age a will be diagnosed with pancreatic cancer before age s ; where

$F(t) = \exp\left\{-\sum_{x=a}^t [RR(x)\lambda_0(t) + \mu_0(t)]\right\}$ is the probability of survival until age t , RR is relative risk with the given risk factors, λ_0 is baseline incidence of pancreatic cancer at age t from the SEER data, and μ_0 is the competing mortality risk at age t . We calculate the baseline incidence $\lambda_0(t)$ separately in men and women by dividing the age-specific SEER incidence rates $\lambda(t)$ by the average RR in the simulated cohort. We calculated 10-year absolute risks (i.e. $p(a, a+10)$ for different reference age a) and cumulative absolute risks (defined as $p(50, 80)$) by categories of risk percentile (10th to 99th percentile). All P values were 2-sided and statistical analyses were performed using SAS (version 9.4; SAS institute Inc, Cary, NC) and R.

Results

Our analysis data set included 500 pancreatic cancer cases and 1,091 matched controls from four prospective cohorts (Table 1 and Methods). In univariable analysis among the full population, we found that increased risk of pancreatic cancer was significantly associated ($P<0.05$) with higher body mass index (BMI) and WHR, history of diabetes, higher levels of circulating proinsulin, IL-6, and total BCAAs, lower levels of circulating adiponectin, and higher weighted genetic risk score (wGRS) of 22 known common susceptibility variants for pancreatic cancer (Table 2). When we restricted our population to cases and matched controls who were diagnosed with pancreatic cancer in the 0–10 years after blood collection, physical activity became a significant risk factor and BMI and WHR were no longer significant (Table 2).

We evaluated three pre-specified multivariable-adjusted risk models that included clinical variables only, clinical variables plus the wGRS, and clinical variables plus the wGRS and

circulating biomarkers (Table 3). We could not include smoking status or family history of pancreatic cancer, two important pancreatic cancer risk factors, as our cases and controls were matched on smoking status and family history information was missing in 58% of subjects. We included external risk estimates for smoking and family history in our final absolute risk model.

In the full population, model fit was improved with addition of the wGRS ($P=3.24\times 10^{-8}$) to the clinical model and with the addition of circulating biomarkers ($P=6.03\times 10^{-5}$) to the model with clinical variables and the wGRS (Table 3). Also, we found a significant improvement of model fit by adding circulating biomarkers only to the clinical model ($P=2.10\times 10^{-5}$) (Supplementary Table 5). For the cases diagnosed within 10 years of covariate data and blood collection and their matched controls, model fit was improved with addition of the wGRS ($P=2.91\times 10^{-7}$) and the circulating biomarkers ($P=2.92\times 10^{-3}$) to the clinical model (Table 3). We also observed that model fit was improved by addition of the circulating biomarkers only to the clinical model ($P=1.05\times 10^{-3}$) (Supplementary Table 5).

Model discrimination was evaluated before and after cross-validation among the 10-year follow-up population (Figure 1). The average AUC estimated by cross-validation was 0.55 for the clinical model, 0.61 for the clinical/genetic model, and 0.62 for the clinical/genetic/biomarker model. Figure 2 shows the population distribution of pancreatic cancer relative risk among US non-Hispanic white men and women by plotting the relative risk (y-axis) as a function of risk percentile based on three risk models (x-axis). These models also incorporate the effects and prevalence of smoking and family history of pancreatic cancer using external risk estimates.^{5, 43} The risk models identified a subset of men and women at 3-fold higher risk for pancreatic cancer than the average risk of men and women in the general population. For instance, the clinical model identified 0.2% of men and 1.5% of women at 3-fold risk of pancreatic cancer during the full follow-up period and the clinical/genetic/biomarker model additionally identified 1.8% of men and 0.7% of women (i.e. 2.0% of men and 2.3% women at 3-fold risk of pancreatic cancer during the full follow-up period). When restricting the follow-up time to 0 – 10 years, the clinical/genetic/biomarker model identified 3.7% of men and 2.6% of women at 3-fold risk for pancreatic cancer over the ensuing 10 years.

We estimated cumulative absolute risk and 10-year risk of pancreatic cancer using the clinical/genetic/biomarker model. We plotted absolute risks (y-axis) by the range of age (x-axis) between 50 and 80 years for cumulative absolute risk and between 50 and 70 years for 10-year absolute risk, stratified by risk percentiles (Figure 3). For cumulative absolute risk, the 10th and 99th risk percentiles showed 0.4% and 3.8% probabilities of developing pancreatic cancer by age 80 years among men. Among women, the corresponding probabilities were 0.4% and 3.6% by age 80 years. The probability of developing pancreatic cancer in the next ten years among cancer-free 70-year-old individuals was 0.2% at the 10th percentile in both men and women and 2.0% and 1.7% at the 99th percentile in men and women, respectively.

Discussion

We developed absolute risk models for pancreatic cancer in the general population, integrating established risk markers for pancreatic cancer, including lifestyle factors, medical comorbidities, common germline variants, and circulating biomarkers. We found that the addition of genetic variants and circulating markers added discriminatory ability beyond clinical factors that could be solicited in a physician's office. The final integrated model identified a subset of approximately 2% of individuals who had 3-fold higher risk than the average in the general U.S. population. Furthermore, the individuals in the top 1% of pancreatic cancer risk as determined by the final integrated model carried a 4% lifetime risk of pancreatic cancer and a 2% 10-year risk at age 70 years.

Screening programs for pancreatic cancer remain early in their development, and recently updated US Preventive Services Task Force (USPSTF) for screening of pancreatic cancer reaffirms that potential benefits of screening do not outweigh the potential harm in asymptomatic, average-risk individuals⁴⁵. However, the USPSTF also confirms that persons with inherited genetic syndromes or family history are at high risk of the disease and their recommendation against screening does not apply to the high-risk populations. In the current study, the high-risks defined here (i.e. 3-fold increased RR) are within a range similar to those for patients with germline mutations in genes such as *BRCA1*, *BRCA2*, or *CDKN2A* (e.g., OR=2.6, 6.2, and 12.3)^{23, 46} or patients with affected family members where the disease screening for these specific populations is being studied^{47, 48}.

We previously used participant data from case-control studies and prospective cohorts in the PanScan consortium to generate a pancreatic cancer risk model based on a small subset of the risk factors included in the current study.³¹ The available risk factors for the prior model included smoking status, alcohol use, BMI, diabetes history, family history of pancreatic cancer, three common genetic susceptibility variants (at 1q32, 5p15, and 13q22) and *ABO* genotype. The full model from this prior work had an in-sample AUC of 0.61 (95% CI = 0.58 – 0.63) and identified 2.9% of men and 2.6% of women who had more than twice the average lifetime risk for pancreatic cancer. In the current study, we improved upon this model by including 18 additional genetic risk variants discovered in subsequent GWAS and several circulating biomarkers, validating our models using cross-validation. Importantly, because all our subjects were enrolled in prospective cohorts, all risk factor data and circulating markers were measured before the cases' diagnosis of pancreatic cancer. This design faithfully recapitulates the situation faced by primary care physicians, where decisions related to disease screening are made in the pre-diagnostic setting using data collected in the several years prior to cancer diagnosis.

A prior case-control study developed a risk prediction model for pancreatic cancer that included current smoking, recent diagnosis of diabetes or pancreatitis, ABO blood type, Jewish ancestry, and use of a proton pump inhibitor.³² Considering these factors, the investigators identified 0.87% of controls that had 5-year absolute risks of 5% or higher. Although risk estimates were based on a single retrospective case-control study from a limited geographic region and with a small number of pancreatic cancer cases, this work highlights the potential utility of including recent development of conditions such as

diabetes and pancreatitis in risk models. Another risk modeling effort has focused specifically on developing prediction models for pancreatic cancer in patients with recently diagnosed diabetes.^{33, 49, 50} In the general population, 0.5% to 0.85% of patients aged 50 years with new-onset diabetes are diagnosed with pancreatic cancer within the ensuing 3 years.⁵¹ With further enrichment, this population may constitute a high-risk group worthy of disease screening. Nevertheless, the majority of patients with pancreatic cancer do not develop diabetes in the three years before diagnosis, so risk models for the general population will remain necessary to diagnosis this disease earlier in most individuals.

The present study has limitations that should be considered. Family history of pancreatic cancer was not collected from most study participants, so the relative risk for family history could not be estimated from our nested case-control data. Additionally, because smoking status was a matching factor at study design stage, so we could not estimate the risk of current smoking in our population. However, we used risk estimates for these factors based on the large PanScan consortium dataset to allow for their inclusion in absolute risk models. For some genetic variants, the proportion of controls missing genotype data was larger than for cases. We imputed genotypes of risk SNPs conditional on case-control status to account for the different missing patterns and allele frequencies between cases and controls. Since cohort data were collected prospectively from study participants using mailed questionnaires every 1 to 2 years, we may have missed some recent-onset diabetes diagnoses. As shown in other risk modeling efforts, recent onset diabetes has predictive ability for pancreatic cancer and therefore our models might underestimate the risk discrimination capabilities of models that incorporate this risk factor. Although we included participants from four separate large U.S. cohorts and performed cross-validation, we could not examine our risk models (absolute or relative risk models) in an independent prospective dataset, which would further validate our models and provide evidence regarding the generalizability of these models in other populations and settings. So, future work for our risk models will be external validation and calibration in independent samples. In particular, since this study did not include non-White participants in the current analyses, further studies that include more racially diverse participant populations will be needed to explore the performance of these models in subjects of other racial and ethnic groups.

Our study has multiple important strengths. The evaluation of participants from large prospective cohorts allowed data and blood samples to be collected pre-diagnostically, minimizing recall bias and the impact of current disease on circulating biomarkers. Our spectrum of pancreatic cancer cases was also less likely to be influenced by survival bias, as participants were identified years before their cancer diagnosis. Our participants were enrolled from across the U.S., enhancing the generalizability of our results to the general population, beyond those who sought care at a specific center or within a particular health care system. We used three types of data to build our risk models, including clinical data that could be queried or measured in the doctor's office, genetic data that could be assessed with sequencing of a germline DNA sample (e.g. peripheral blood white cells or buccal swab), and circulating biomarkers that could be measured from peripheral blood in commonly collected plasma tubes. Overall, these design features are extremely well suited to simulate the data available to providers seeing patients in general medicine clinics. If such a risk stratification tool were available to primary care providers, excess pancreatic cancer risk

could trigger further biomarker testing (e.g., specialized blood tests) or imaging-based screening tests (e.g., magnetic resonance imaging or endoscopic ultrasound) to detect an early pancreatic cancer that could be treated for cure. Such risk stratification tools will become increasingly important as novel early detection biomarkers become available and imaging tests are improved for detection of small tumors.^{52–55}

In summary, we have examined absolute risk models of pancreatic cancer that combine established clinical factors, germline genetic variants and circulating biomarkers. The final integrated model has improved risk discrimination over those that include clinical factors alone and successfully identify a small segment of the general population at elevated risk of pancreatic cancer. Further refinement and validation in independent samples will be necessary to make these models clinically actionable and impact survival of patients with pancreatic cancer. Given the late stage at presentation for most patients with pancreatic cancer, earlier detection approaches are worthy of significant investment as a critical means to reduce mortality from pancreatic cancer, soon to be the second leading cause of cancer death in the United States.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors would like to thank the participants and staff of the Health Professionals Follow-up Study, Nurses' Health Study, Physicians' Health Study, and Women's Health Initiative for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, and WY.

This project was supported by cohort grants (UM1CA167552 (W. Willett) and U01CA167552 (W. Willett) for the Health Professionals Follow-up Study; UM1CA186107 (M. Stampfer), P01CA87969 (R. Tamimi), and R01CA49449 (S. Hankinson) for the Nurses' Health Study; R01CA97193 (JM. Gaziano), R01CA34944 (C. Hennekens), R01CA40360 (J. Buring), R01HL26490 (C. Hennekens), and R01HL34595 (C. Hennekens) for the Physicians' Health Study; N01WH22110 (R. Prentice), N01WH24152 (N. Lasser), N01WH32100 (S. Beresford), N01WH32101 (R. Grimm), N01WH32102 (R. Wallace), N01WH32105 (A. Oberman), N01WH32106 (E. Paskett), N01WH32108 (P. Greenland), N01WH32109 (J. Manson), N01WH32111 (N. Watts), N01WH32112 (L. Kuller), N01WH32113 (J. Robbins), N01WH32115 (T. Bassford), N01WH32118 (K. Johnson), N01WH32119 (A. Assaf), N01WH32122 (M. Trivisan), N01WH42107 (A. Hubbell), N01WH42108 (J. Hsia), N01WH42109 (M. Stefanick), N01WH42110 (J. Hays), N01WH42111 (R. Schenken), N01WH42112 (R. Jackson), N01WH42113 (S. Daugherty), N01WH42114 (C. Ritenbaugh), N01WH42115 (D. Lane), N01WH42116 (J. Ockene), N01WH42117 (G. Heiss), N01WH42118 (S. Hendrix), N01WH42119 (S. Wassertheil-Smoller), N01WH42120 (R. Chiebowski), N01WH42121 (B. Canne), N01WH42122 (J. Kotchen), N01WH42123 (B. Howard), N01WH42124 (H. Black), N01WH42125 (H. Judd), N01WH42126 (J. Liu), N01WH42129 (M. Limacher), N01WH42130 (J. Curb), N01WH42131 (M. O'Sullivan), N01WH42132 (C. Allen), and N01WH44221 (S. Shumaker) for the Women's Health Initiative program) from the U.S. National Institutes of Health (NIH).

B.M. Wolpin acknowledges primary research support from Dana-Farber Cancer Institute Hale Family Center for Pancreatic Cancer Research, NIH/NCI U01CA210171, Lustgarten Foundation and Stand Up to Cancer, with additional support from Pancreatic Cancer Action Network, Noble Effort Fund, and Promises for Purple. K. Ng acknowledges research funding from the Broman Fund for Pancreatic Cancer Research.

Disclosure of conflict of interest: KN declares research funding from Gilead, Celgene, Pharmavite, and Tarrex, and consulting or advisory board participation for Genentech, Bayer, Lilly, Seattle Genetics, and Tarrex. BMW declares research funding from Celgene and Eli Lilly & Company, and consulting for G1 Therapeutics, BioLineRx, Celgene, and GRAIL.

Abbreviations used in this paper:

AIC	akaike information criterion
AUC	area under the receiver operating characteristic curve
BCAAs	branched chain amino acids
BMI	body mass index
CV	coefficient of variance
GWAS	genome-wide association study
HPFS	health professionals follow-up study
IL-6	interleukin-6
MET	metabolic equivalent of task
MICE	multivariate imputation by chained equations
NHS	nurses' health study
OR	odds ratio
PHS	physicians' health study
PRS	polygenic risk score
ROC	receiver operating characteristic
RR	relative risk
SD	standard deviation
SEER	Surveillance, Epidemiology, and End Results
SNP	single nucleotide polymorphism
wGRS	weighted genetic risk score
WHI	women's health initiative
WHR	waist-to-hip ratio

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin* 2018;68:7–30. [PubMed: 29313949]
2. Rahib L, Smith BD, Aizenberg R, et al. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* 2014;74:2913–21. [PubMed: 24840647]
3. Vincent A, Herman J, Schulick R, et al. Pancreatic cancer. *Lancet* 2011;378:607–20. [PubMed: 21620466]

4. Paniccia A, Hosokawa P, Henderson W, et al. Characteristics of 10-Year Survivors of Pancreatic Ductal Adenocarcinoma. *JAMA Surg* 2015;150:701–10. [PubMed: 26062046]
5. Lynch SM, Vrieling A, Lubin JH, et al. Cigarette smoking and pancreatic cancer: a pooled analysis from the pancreatic cancer cohort consortium. *Am J Epidemiol* 2009;170:403–13. [PubMed: 19561064]
6. Michaud DS, Giovannucci E, Willett WC, et al. Physical activity, obesity, height, and the risk of pancreatic cancer. *JAMA* 2001;286:921–9. [PubMed: 11509056]
7. Silverman DT, Schiffman M, Everhart J, et al. Diabetes mellitus, other medical conditions and familial history of cancer as risk factors for pancreatic cancer. *Br J Cancer* 1999;80:1830–7. [PubMed: 10468306]
8. Wolpin BM, Bao Y, Qian ZR, et al. Hyperglycemia, Insulin Resistance, Impaired Pancreatic beta-Cell Function, and Risk of Pancreatic Cancer. *J Natl Cancer Inst* 2013;105:1027–1035. [PubMed: 23847240]
9. Stolzenberg-Solomon RZ, Graubard BI, Chari S, et al. Insulin, glucose, insulin resistance, and pancreatic cancer in male smokers. *JAMA* 2005;294:2872–8. [PubMed: 16352795]
10. Sadr-Azodi O, Gudbjornsdottir S, Ljung R. Pattern of increasing HbA1c levels in patients with diabetes mellitus before clinical detection of pancreatic cancer - a population-based nationwide case-control study. *Acta Oncol* 2015;54:986–92. [PubMed: 25734801]
11. Wolpin BM, Michaud DS, Giovannucci EL, et al. Circulating insulin-like growth factor binding protein-1 and the risk of pancreatic cancer. *Cancer Res* 2007;67:7923–8. [PubMed: 17699799]
12. Wolpin BM, Ng K, Bao Y, et al. Plasma 25-hydroxyvitamin D and risk of pancreatic cancer. *Cancer Epidemiol Biomarkers Prev* 2012;21:82–91. [PubMed: 22086883]
13. Bao Y, Giovannucci EL, Kraft P, et al. A prospective study of plasma adiponectin and pancreatic cancer risk in five US cohorts. *J Natl Cancer Inst* 2013;105:95–103. [PubMed: 23243202]
14. White DL, Hoogeveen RC, Chen L, et al. A prospective study of soluble receptor for advanced glycation end products and adipokines in association with pancreatic cancer in postmenopausal women. *Cancer Med* 2018;7:2180–2191. [PubMed: 29573228]
15. Babic A, Bao Y, Qian ZR, et al. Pancreatic Cancer Risk Associated with Prediagnostic Plasma Levels of Leptin and Leptin Receptor Genetic Polymorphisms. *Cancer Res* 2016;76:7160–7167. [PubMed: 27780823]
16. Stolzenberg-Solomon RZ, Newton CC, Silverman DT, et al. Circulating Leptin and Risk of Pancreatic Cancer: A Pooled Analysis From 3 Cohorts. *Am J Epidemiol* 2015;182:187–97. [PubMed: 26085045]
17. Vainer N, Dehlendorff C, Johansen JS. Systematic literature review of IL-6 as a biomarker or treatment target in patients with gastric, bile duct, pancreatic and colorectal cancer. *Oncotarget* 2018;9:29820–29841. [PubMed: 30038723]
18. Mayers JR, Wu C, Clish CB, et al. Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development. *Nat Med* 2014;20:1193–1198. [PubMed: 25261994]
19. Katagiri R, Goto A, Nakagawa T, et al. Increased Levels of Branched-Chain Amino Acid Associated With Increased Risk of Pancreatic Cancer in a Prospective Case-Control Study of a Large Cohort. *Gastroenterology* 2018;155:1474–1482 e1. [PubMed: 30076838]
20. Yip-Schneider MT, Simpson R, Carr RA, et al. Circulating Leptin and Branched Chain Amino Acids—Correlation with Intraductal Papillary Mucinous Neoplasm Dysplastic Grade. *J Gastrointest Surg* 2019;23:966–974. [PubMed: 30215202]
21. Shindo K, Yu J, Suenaga M, et al. Deleterious Germline Mutations in Patients With Apparently Sporadic Pancreatic Adenocarcinoma. *J Clin Oncol* 2017;35:3382–3390. [PubMed: 28767289]
22. Yurgelun MB, Chittenden AB, Morales-Oyarvide V, et al. Germline cancer susceptibility gene variants, somatic second hits, and survival outcomes in patients with resected pancreatic cancer. *Genet Med* 2018.
23. Hu C, Hart SN, Polley EC, et al. Association Between Inherited Germline Mutations in Cancer Predisposition Genes and Risk of Pancreatic Cancer. *JAMA* 2018;319:2401–2409. [PubMed: 29922827]

24. Lu Y, Ek WE, Whiteman D, et al. Most common 'sporadic' cancers have a significant germline genetic component. *Hum Mol Genet* 2014;23:6112–8. [PubMed: 24943595]
25. Amundadottir L, Kraft P, Stolzenberg-Solomon RZ, et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nat Genet* 2009;41:986–90. [PubMed: 19648918]
26. Petersen GM, Amundadottir L, Fuchs CS, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. *Nat Genet* 2010;42:224–8. [PubMed: 20101243]
27. Wolpin BM, Rizzato C, Kraft P, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. *Nat Genet* 2014;46:994–1000. [PubMed: 25086665]
28. Childs EJ, Mocchi E, Campa D, et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. *Nat Genet* 2015;47:911–6. [PubMed: 26098869]
29. Zhang M, Wang Z, Obazee O, et al. Three new pancreatic cancer susceptibility signals identified on chromosomes 1q32.1, 5p15.33 and 8q24.21. *Oncotarget* 2016;7:66328–66343. [PubMed: 27579533]
30. Klein AP, Wolpin BM, Risch HA, et al. Genome-wide meta-analysis identifies five new susceptibility loci for pancreatic cancer. *Nat Commun* 2018;9:556. [PubMed: 29422604]
31. Klein AP, Lindstrom S, Mendelsohn JB, et al. An absolute risk model to identify individuals at elevated risk for pancreatic cancer in the general population. *PLoS One* 2013;8:e72311.
32. Risch HA, Yu H, Lu L, et al. Detectable Symptomatology Preceding the Diagnosis of Pancreatic Cancer and Absolute Risk of Pancreatic Cancer Diagnosis. *Am J Epidemiol* 2015;182:26–34. [PubMed: 26049860]
33. Boursi B, Finkelman B, Giantonio BJ, et al. A Clinical Prediction Model to Assess Risk for Pancreatic Cancer Among Patients With New-Onset Diabetes. *Gastroenterology* 2017;152:840–850 e3. [PubMed: 27923728]
34. Giovannucci E, Ascherio A, Rimm EB, et al. Physical activity, obesity, and risk for colon cancer and adenoma in men. *Ann Intern Med* 1995;122:327–34. [PubMed: 7847643]
35. Colditz GA, Hankinson SE. The Nurses' Health Study: lifestyle and health among women. *Nat Rev Cancer* 2005;5:388–96. [PubMed: 15864280]
36. Steering Committee of the Physicians' Health Study Research G. Final report on the aspirin component of the ongoing Physicians' Health Study. *N Engl J Med* 1989;321:129–35. [PubMed: 2664509]
37. Langer RD, White E, Lewis CE, et al. The Women's Health Initiative Observational Study: baseline characteristics of participants and reliability of baseline measures. *Ann Epidemiol* 2003;13:S107–21. [PubMed: 14575943]
38. Bao Y, Giovannucci EL, Kraft P, et al. Inflammatory plasma markers and pancreatic cancer risk: a prospective study of five U.S. cohorts. *Cancer Epidemiol Biomarkers Prev* 2013;22:855–61. [PubMed: 23462920]
39. Wolpin BM, Bao Y, Qian ZR, et al. Hyperglycemia, insulin resistance, impaired pancreatic beta-cell function, and risk of pancreatic cancer. *J Natl Cancer Inst* 2013;105:1027–35. [PubMed: 23847240]
40. Dupont WD. Converting relative risks to absolute risks: a graphical approach. *Stat Med* 1989;8:641–51. [PubMed: 2749072]
41. Gail MH, Pfeiffer RM. On criteria for evaluating models of absolute risk. *Biostatistics* 2005;6:227–39. [PubMed: 15772102]
42. NCHS. Tables of Summary Health Statistics. Volume 2018, 2009.
43. Jacobs EJ, Chanock SJ, Fuchs CS, et al. Family history of cancer and risk of pancreatic cancer: a pooled analysis from the Pancreatic Cancer Cohort Consortium (PanScan). *Int J Cancer* 2010;127:1421–8. [PubMed: 20049842]
44. NCHS. Detailed technical notes to the United States 2007 data—mortality. Volume 2018, 2010.
45. Force USPST, Owens DK, Davidson KW, et al. Screening for Pancreatic Cancer: US Preventive Services Task Force Reaffirmation Recommendation Statement. *JAMA* 2019;322:438–444. [PubMed: 31386141]

46. Petersen GM. Familial Pancreatic Adenocarcinoma. *Hematol Oncol Clin North Am* 2015;29:641–53. [PubMed: 26226902]
47. Lucas AL, Kastrinos F. Screening for Pancreatic Cancer. *JAMA* 2019;322:407–408. [PubMed: 31386115]
48. Canto MI, Harinck F, Hruban RH, et al. International Cancer of the Pancreas Screening (CAPS) Consortium summit on the management of patients with increased risk for familial pancreatic cancer. *Gut* 2013;62:339–47. [PubMed: 23135763]
49. Munigala S, Singh A, Gelrud A, et al. Predictors for Pancreatic Cancer Diagnosis Following New-Onset Diabetes Mellitus. *Clin Transl Gastroenterol* 2015;6:e118. [PubMed: 26492440]
50. Sharma A, Kandlakunta H, Nagpal SJS, et al. Model to Determine Risk of Pancreatic Cancer in Patients With New-Onset Diabetes. *Gastroenterology* 2018.
51. Chari ST, Leibson CL, Rabe KG, et al. Probability of pancreatic cancer following diabetes: a population-based study. *Gastroenterology* 2005;129:504–11. [PubMed: 16083707]
52. Cohen JD, Javed AA, Thoburn C, et al. Combined circulating tumor DNA and protein biomarker-based liquid biopsy for the earlier detection of pancreatic cancers. *Proc Natl Acad Sci U S A* 2017;114:10202–10207. [PubMed: 28874546]
53. Fahrman JF, Bantis LE, Capello M, et al. A Plasma-Derived Protein-Metabolite Multiplexed Panel for Early-Stage Pancreatic Cancer. *J Natl Cancer Inst* 2018.
54. Koay EJ, Lee Y, Cristini V, et al. A visually apparent and quantifiable CT imaging feature identifies biophysical subtypes of pancreatic ductal adenocarcinoma. *Clin Cancer Res* 2018.
55. Abou-Elkacem L, Wang H, Chowdhury SM, et al. Thy1-Targeted Microbubbles for Ultrasound Molecular Imaging of Pancreatic Ductal Adenocarcinoma. *Clin Cancer Res* 2018;24:1574–1585. [PubMed: 29301827]

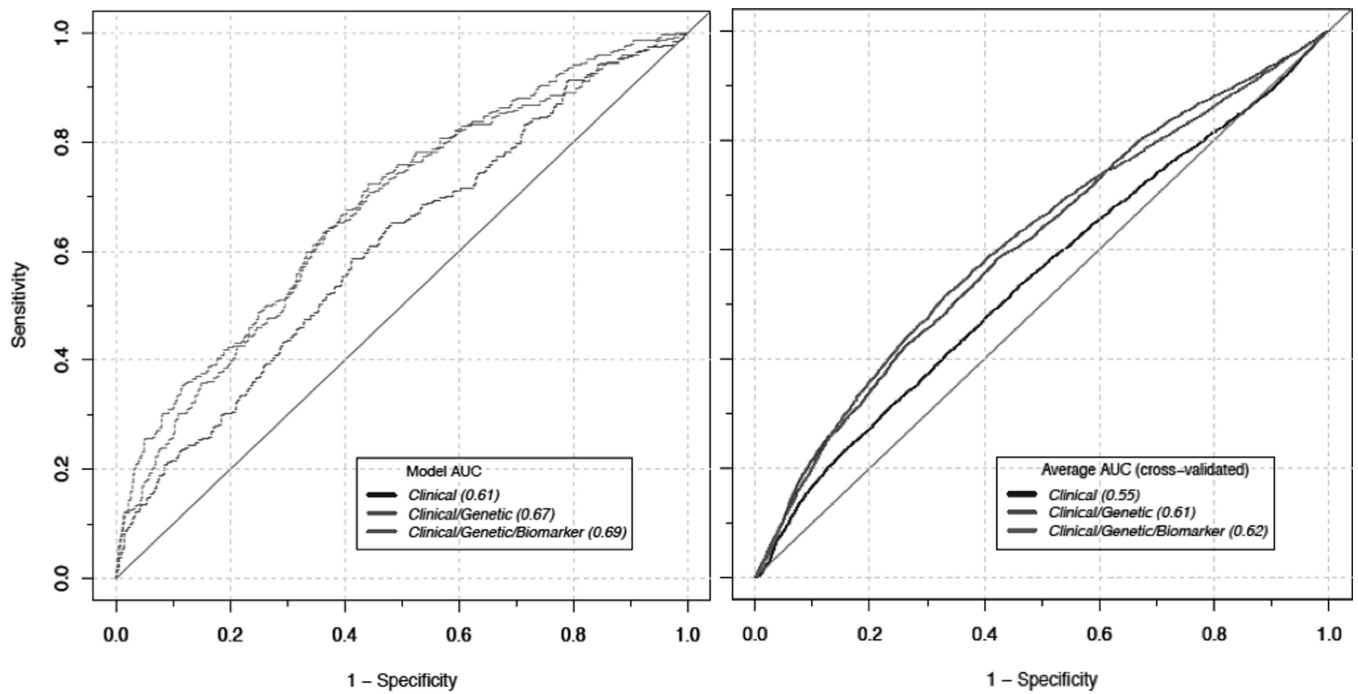


Figure 1. ROC curves from before (left) and after 5-fold cross validation (right) in the 0–10 years follow-up population. Each line represents the clinical model (navy), the clinical/genetic model (green), and the clinical/genetic/biomarker model (red). The 5-fold cross validation leaving out 20% randomly selected dataset as a test set at a time was performed 20 times. The average AUC was calculated as a mean of 100 AUC values estimated in the test datasets.

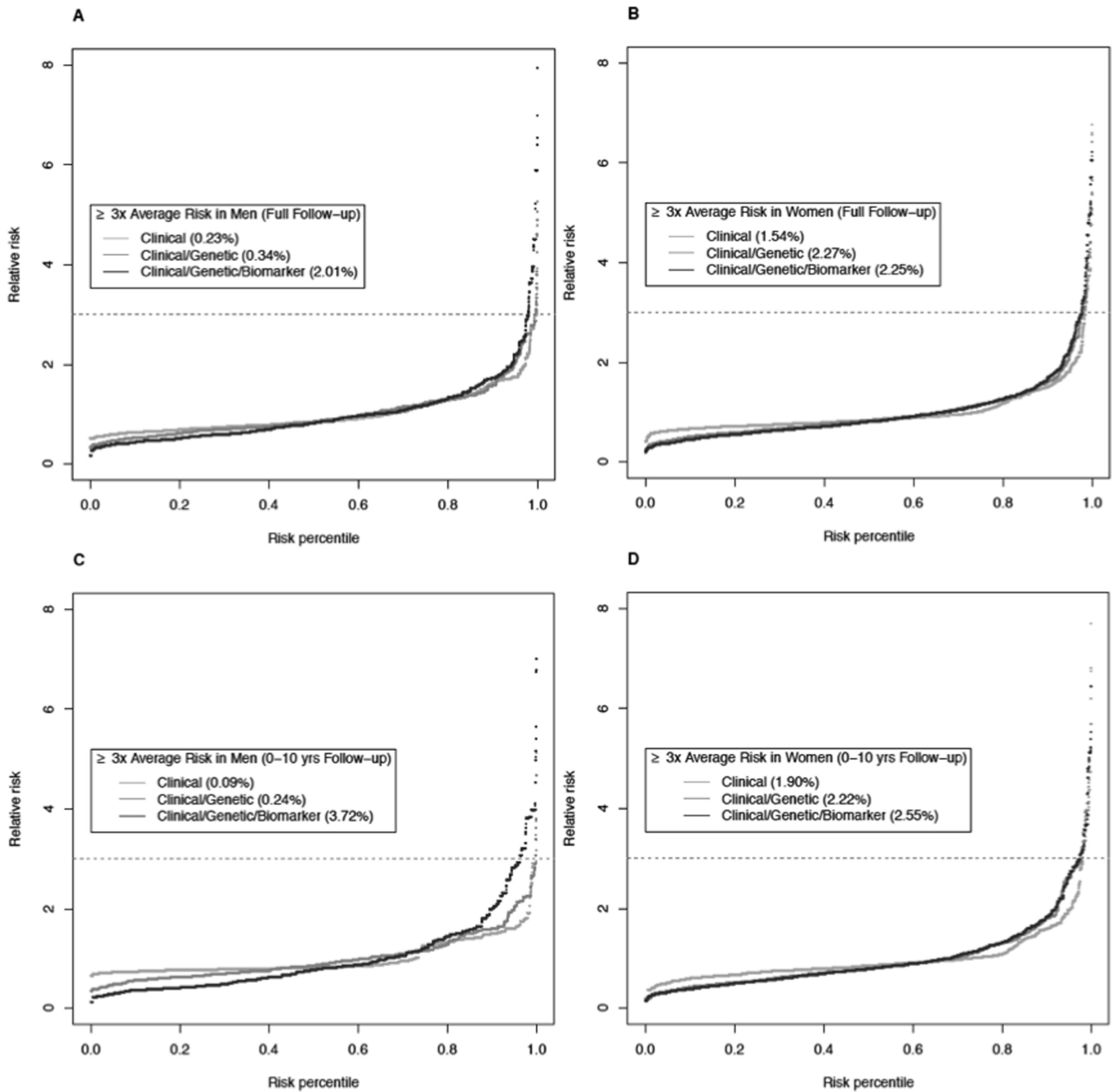


Figure 2.

Pancreatic cancer risk in the general population. The data were simulated with a total of 20,000 men and 20,000 women based on the average of our imputed datasets using external risk estimates for smoking status and family history of pancreatic cancer. The relative risk of pancreatic cancer was plotted with a function of the risk percentile for (A) men in the full years of follow-up, (B) women in the full years of follow-up, (C) men in 0–10 years of follow-up, and (D) women in 0–10 years of follow-up. The lines represent three risk models including clinical factors only (yellow), clinical and genetic factors (orange), and clinical and genetic factors as well as circulating biomarkers (red).

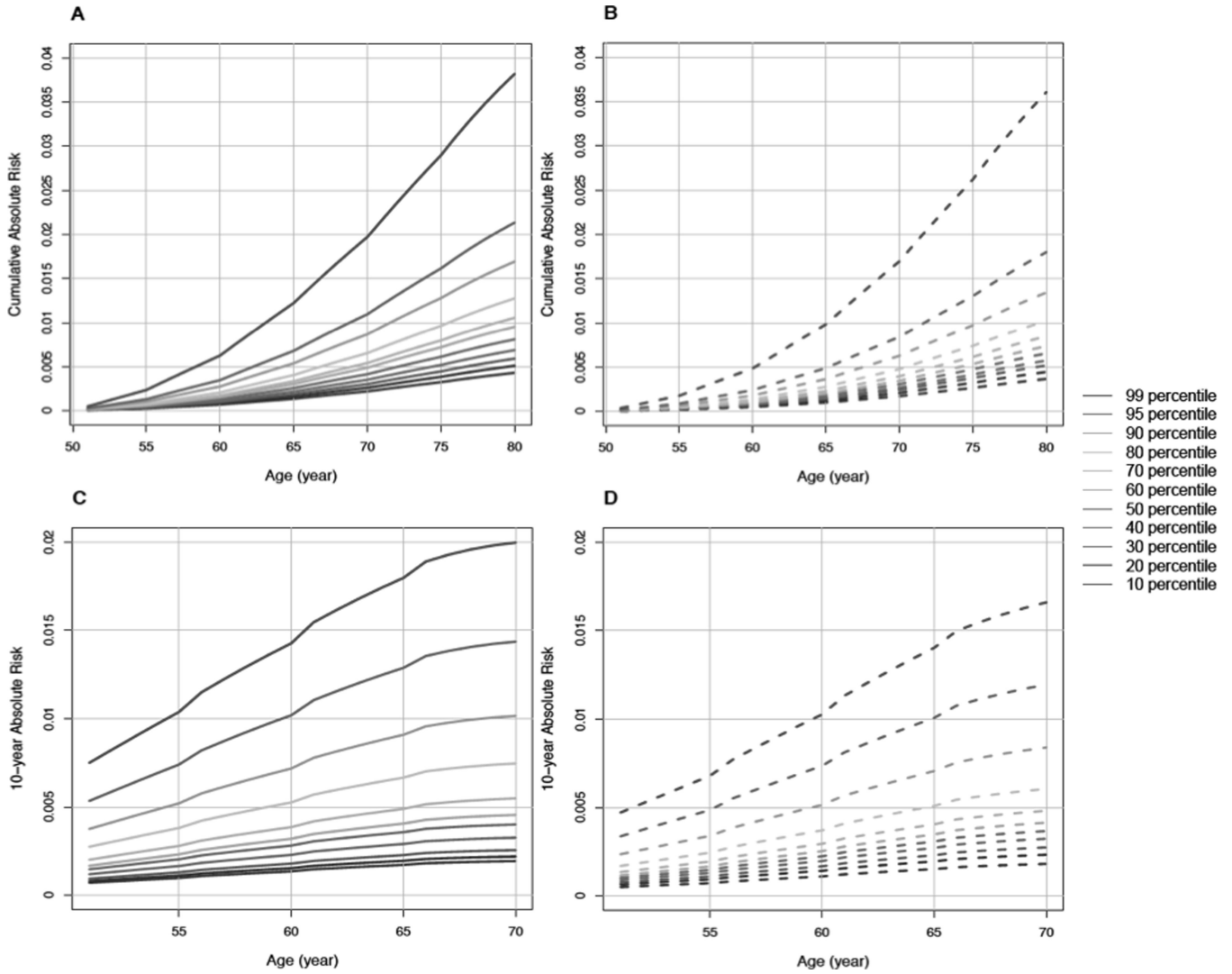


Figure 3. Cumulative absolute risk and 10-year absolute risks of pancreatic cancer estimated using simulated data of 20,000 men and 20,000 women with smoking status and family history status based on the average of imputed datasets. Each color line represents different relative risk percentiles in each gender group and the percentiles were estimated based on the clinical/genetic/biomarker model (including BMI, waist-to-hip ratio, physical activity, diagnosed diabetes, wGRS, proinsulin, adiponectin, interleukin-6, and total BCAAs). (A) Men in the full follow-up population, (B) women in in the full follow-up population, (C) men in the 0–10 years follow-up population, and (D) women in the 0–10 years follow-up population.

Table 1.

Characteristics of pancreatic cancer cases and matched controls

Variables	Full Population (n = 1,591)		0–10 years Population ^a (n = 956)	
	Cases	Controls	Cases	Controls
	n = 500	n = 1,091	n = 304	n = 652
Matching factors				
Age, mean (SD), year	63.19 (8.30)	62.67 (8.31)	65.93 (7.59)	65.54 (7.55)
Gender, n (%)				
Male	173 (34.60)	358 (32.81)	82 (26.97)	187 (28.68)
Female	327 (65.40)	733 (67.19)	222 (73.03)	465 (71.32)
Cohort, n (%)				
HPFS	83 (16.60)	195 (17.87)	58 (19.08)	145 (22.24)
NHS	147 (29.40)	396 (36.30)	48 (15.79)	140 (21.47)
PHS	90 (18.00)	163 (14.94)	24 (7.89)	42 (6.44)
WHI	180 (36.00)	337 (30.89)	174 (57.24)	325 (49.85)
Smoking, n (%)				
Current smoker	64 (12.90)	135 (12.45)	37 (12.29)	76 (11.76)
Non-current smoker	432 (87.10)	949 (87.55)	264 (87.71)	570 (88.24)
Fasting status at blood collection, n (%)				
Fasted < 8 hours	142 (28.40)	290 (26.58)	48 (15.79)	118 (18.10)
Fasted ≥ 8 hours	358 (71.60)	801 (73.42)	256 (84.21)	534 (81.90)
Lifestyle and clinical factors				
Body mass index, mean (SD), kg/m ²	26.30 (5.03)	25.70 (4.33)	26.60 (5.50)	26.00 (4.63)
Waist-to-hip ratio, mean (SD), inch/inch	0.85 (0.11)	0.84 (0.10)	0.84 (0.10)	0.83 (0.09)
Physical activity, mean (SD), MET-hour/week	20.10 (32.80)	20.40 (25.80)	17.70 (24.10)	21.50 (29.00)
Diagnosed diabetes (yes), n (%)	29 (5.80)	33 (3.02)	21 (6.91)	24 (3.68)
Circulating biomarkers				
Proinsulin, mean (SD), pM	16.10 (18.70)	12.90 (19.30)	15.70 (19.00)	12.90 (19.30)
Adiponectin (4.4 ug/ml), n (%)	301 (71.84)	743 (81.20)	219 (74.74)	524 (83.71)
Interleukin-6, mean (SD), pg/mL	2.38 (4.20)	1.96 (3.36)	2.60 (4.72)	2.00 (3.03)
Total BCAAs, mean (SD), μM	430.10 (169.89)	359.05 (200.66)	437.16 (141.29)	368.17 (179.37)
Genetic risk factors				
GRS, mean (SD)	23.60 (2.75)	22.90 (2.64)	23.80 (2.75)	22.80 (2.68)
wGRS ^b , mean (SD)	0.21 (1.01)	-0.10 (0.98)	0.26 (1.00)	-0.13 (0.99)

HPFS, Health Professionals Follow-up Study; NHS, Nurses' Health Study; PHS, Physicians' Health Study; WHI, Women's Health Initiative; BCAAs, branched-chain amino acids; GRS, genetic risk score summing the number of risk alleles; wGRS, weighted genetic risk score.

^a0–10 years population refers to cases (and their matched controls) diagnosed within 10 years of blood draw.

^bStandardized wGRS with mean = 0 and standard deviation = 1 within each cohort

Table 2.

Univariable odds ratios and 95% CIs for susceptibility factors and future pancreatic cancer risk

Variables	Full Population	0–10 years Population
	(n = 1,591)	(n = 956)
	OR (95% CI)	OR (95% CI)
<i>Lifestyle and clinical factors</i>		
Body mass index ^a	1.14 (1.03, 1.27)	1.12 (0.98, 1.27)
Waist-to-hip ratio ^a	1.18 (1.06, 1.31)	1.13 (0.99, 1.29)
Physical activity ^a	0.94 (0.85, 1.05)	0.85 (0.73, 0.99)
Diagnosed diabetes (yes)	2.36 (1.32, 4.21)	2.42 (1.20, 4.89)
<i>Circulating biomarkers</i>		
Proinsulin ^a	1.27 (1.14, 1.42)	1.21 (1.07, 1.38)
Adiponectin (4.4 ug/ml)	0.62 (0.48, 0.80)	0.57 (0.40, 0.80)
Interleukin-6 ^a	1.13 (1.02, 1.25)	1.16 (1.02, 1.33)
Total BCAAs ^a	1.46 (1.23, 1.74)	1.43 (1.18, 1.74)
<i>Genetic risk score</i>		
wGRS ^a	1.37 (1.23, 1.53)	1.46 (1.27, 1.68)

OR, odds ratio; CI, confidence interval; BCAAs, branched-chain amino acids; wGRS, weighted genetic risk score.

^bStandardized variables with mean = 0 and SD= 1 within each cohort

Table 3.Estimated ORs and 95% CIs from multivariable^a risk models for pancreatic cancer

	Clinical model	Clinical/genetic model	Clinical/genetic/biomarker model
Full follow-up period			
Model comparison (<i>P</i> value ^b)		3.24e-08	6.03e-05
Model AUC	0.61	0.65	0.67
	OR (95% CI)	OR (95% CI)	OR (95% CI)
Body mass index ^c	1.08 (0.97, 1.21)	1.07 (0.95, 1.20)	0.98 (0.86, 1.10)
Waist-to-hip ratio ^c	1.13 (1.01, 1.26)	1.12 (1.00, 1.26)	1.08 (0.96, 1.21)
Physical activity ^c	0.96 (0.86, 1.06)	0.95 (0.85, 1.06)	0.97 (0.86, 1.08)
Diagnosed diabetes (yes)	2.10 (1.16, 3.79)	2.19 (1.19, 4.02)	1.70 (0.91, 3.19)
wGRS ^c		1.37 (1.22, 1.53)	1.36 (1.21, 1.52)
Proinsulin ^c			1.16 (1.03, 1.31)
Adiponectin (< 4.4 ug/ml)			0.76 (0.58, 0.99)
Interleukin-6 ^c			1.10 (0.99, 1.23)
Total BCAAs ^c			1.25 (1.04, 1.51)
0–10 years follow-up period			
Model comparison (<i>P</i> value ^b)		2.91e-07	2.92e-03
Model AUC	0.61	0.67	0.69
	OR (95% CI)	OR (95% CI)	OR (95% CI)
Body mass index ^c	1.05 (0.91, 1.22)	1.04 (0.90, 1.21)	0.96 (0.82, 1.12)
Waist-to-hip ratio ^c	1.08 (0.93, 1.25)	1.06 (0.91, 1.23)	1.00 (0.86, 1.17)
Physical activity ^c	0.86 (0.74, 1.00)	0.86 (0.74, 1.01)	0.88 (0.75, 1.03)
Diagnosed diabetes (yes)	2.22 (1.09, 4.54)	2.14 (1.02, 4.50)	1.65 (0.77, 3.56)
wGRS ^c		1.44 (1.25, 1.67)	1.43 (1.23, 1.65)
Proinsulin ^c			1.10 (0.94, 1.27)
Adiponectin (< 4.4 ug/ml)			0.70 (0.48, 1.02)
Interleukin-6 ^c			1.13 (0.98, 1.30)
Total BCAAs ^c			1.24 (1.00, 1.54)

AUC, Area under the ROC curve; OR, odds ratio; CI, confidence interval; BCAAs, branched-chain amino acids; wGRS, weighted genetic risk score.

^aAdjusted for matching factors, age, cohort (also gender), race/ethnicity, smoking status, fasting status, and month/year of blood collection.

^b *P* value was estimated from the likelihood ratio test comparing the clinical/genetic model to the clinical model and the clinical/genetic/biomarker model to the clinical/genetic model.

^c Standardized variables with mean = 0 and SD= 1 within each cohort

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript