

Genetic and Circulating Biomarker Data Improve Risk Prediction for Pancreatic Cancer in the General Population

Jihye Kim^{1,2}, Chen Yuan³, Ana Babic³, Ying Bao⁴, Clary B. Clish⁵, Michael N. Pollak⁶, Laufey T. Amundadottir⁷, Alison P. Klein^{8,9}, Rachael Z. Stolzenberg-Solomon¹⁰, Pari V. Pandharipande¹¹, Lauren K. Brais³, Marisa W. Welch³, Kimmie Ng³, Edward L. Giovannucci^{2,4,12}, Howard D. Sesso^{2,13}, JoAnn E. Manson^{2,13}, Meir J. Stampfer^{2,4,12}, Charles S. Fuchs^{14,15,16}, Brian M. Wolpin³, and Peter Kraft^{1,2}

ABSTRACT

Background: Pancreatic cancer is the third leading cause of cancer death in the United States, 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.

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.

Results: Model discrimination showed an area under ROC curve of 0.62 via cross-validation. 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 that include established clinical, genetic, and circulating factors improved disease discrimination over models using clinical factors alone.

Impact: 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 United States within the next 10 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 radiotherapy (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.

Epidemiologic 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), and 25-hydroxyvitamin D (12)], adipokines [adiponectin (13, 14) and leptin (15, 16)], inflammation (IL6; ref. 17), and peripheral tissue catabolism [branched chain amino acids (BCAA; refs. 18–20)].

Inherited genetic variants have been identified that predispose to development of pancreatic cancer. Medium- to high-penetrance

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

B.M. Wolpin and P. Kraft are co-senior authors of this article.

Corresponding Authors: Peter Kraft, Harvard T.H. Chan School of Public Health, 655 Huntington Avenue, Boston, MA 02115. Phone: 617-432-4271; Fax: 617-432-1722; E-mail: pkraft@hsph.harvard.edu; and Brian M. Wolpin, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215. Phone: 617-632-6942; Fax: 617-632-5370; E-mail: brian_wolpin@dfci.harvard.edu

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 cancer (21–23). 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 SNPs have been identified surpassing the genome-wide significance threshold ($P < 5 \times 10^{-8}$).

Although risk factors have been investigated individually (31–33), their joint contribution to risk discrimination remains 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 ages 40 to 75 years in 1986 (34). In NHS cohort, 121,701 female nurses ages 30 to 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 ages 40 to 84 years. After trial completion in 1995, PHS I participants were followed up in an observational cohort (36). In the WHI, 93,726 women ages 50 to 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 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 nonwhite 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 postbaseline 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 (pmol/L), adiponectin ($\mu\text{g}/\text{mL}$), IL6 (pg/mL; ref. 38), and BCAAs ($\mu\text{mol}/\text{L}$) 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 $\mu\text{g}/\text{mL}$ as done previously (13). Details for laboratory assays and coefficients of variance (CV) have been previously reported (8, 11–13, 15, 18). CVs 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 S2). 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 ORs 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 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, IL6, 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 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 ROC 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 (39, 40). 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 (41, 42).

To estimate the joint distribution of pancreatic cancer risk factors 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 ages 18 and over from the National Health Interview Survey data, 2011–2014; ref. 41). 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, 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% (42).

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

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

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 OR for the risk factors in our risk models and literature estimates for current smoking and family history of pancreatic cancer (5, 42).

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 (43). Using these data, we converted relative risks to absolute risks (p) as follows (39):

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

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 $F(t) = \exp\{-\sum_{x=a}^t [RR(x)\lambda_0(t) + \mu_0(t)]\}$ 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 percentiles). All P values were two-sided, and statistical analyses were performed using SAS (version 9.4; SAS institute Inc.) and R.

Results

Our analysis dataset included 500 pancreatic cancer cases and 1,091 matched controls from four prospective cohorts (Table 1; Supplementary Table S3; and Materials 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 BMI and WHR, history of diabetes, higher levels of circulating proinsulin, IL6, and total BCAAs, lower levels of circulating adiponectin, and higher wGRS of 22 known common susceptibility variants for pancreatic cancer (Table 2; Supplementary Table S4). When we restricted our population to cases and matched controls who were diagnosed with pancreatic cancer in the 0 to 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 prespecified 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 S5). 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 S5).

Model discrimination was evaluated before and after cross-validation among the 10-year follow-up population (Fig. 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 U.S. non-Hispanic white men

Table 1. Characteristics of pancreatic cancer cases and matched controls.

Variables	Full population (n = 1,591)		0-10-year population ^a (n = 956)	
	Cases (n = 500)	Controls (n = 1,091)	Cases (n = 304)	Controls (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)
Noncurrent 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				
BMI, mean (SD), kg/m ²	26.30 (5.03)	25.70 (4.33)	26.60 (5.50)	26.00 (4.63)
WHR, 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), pmol/L	16.10 (18.70)	12.90 (19.30)	15.70 (19.00)	12.90 (19.30)
Adiponectin (≥4.4 μg/mL), n (%)	301 (71.84)	743 (81.20)	219 (74.74)	524 (83.71)
IL6, mean (SD), pg/mL	2.38 (4.20)	1.96 (3.36)	2.60 (4.72)	2.00 (3.03)
Total BCAAs, mean (SD), μmol/L	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)

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

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

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

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

Table 2. Univariable ORs and 95% CIs for susceptibility factors and future pancreatic cancer risk.

Variables	Full population (n = 1,591)	0-10-year population (n = 956)
	OR (95% CI)	OR (95% CI)
Lifestyle and clinical factors		
BMI ^a	1.14 (1.03-1.27)	1.12 (0.98-1.27)
WHR ^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)
Circulating biomarkers		
Proinsulin ^a	1.27 (1.14-1.42)	1.21 (1.07-1.38)
Adiponectin (≥4.4 μg/mL)	0.62 (0.48-0.80)	0.57 (0.40-0.80)
IL6 ^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)
Genetic risk factor		
wGRS ^a	1.37 (1.23-1.53)	1.46 (1.27-1.68)

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

pancreatic cancer using external risk estimates (5, 42). 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 to 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 (Fig. 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 10 years among cancer-free 70-year-old individuals was 0.2% at the 10th percentile in both men and women

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)		3.24×10^{-8}	6.03×10^{-5}
Model AUC	0.61	0.65	0.67
OR (95% CI)			
BMI ^c	1.08 (0.97–1.21)	1.07 (0.95–1.20)	0.98 (0.86–1.10)
WHR ^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 $\mu\text{g/mL}$)			0.76 (0.58–0.99)
IL6 ^c			1.10 (0.99–1.23)
Total BCAAs ^c			1.25 (1.04–1.51)
0–10 years of follow-up period			
Model comparison (<i>P</i> value ^b)		2.91×10^{-7}	2.92×10^{-3}
Model AUC	0.61	0.67	0.69
OR (95% CI)			
BMI ^c	1.05 (0.91–1.22)	1.04 (0.90–1.21)	0.96 (0.82–1.12)
WHR ^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 $\mu\text{g/mL}$)			0.70 (0.48–1.02)
IL6 ^c			1.13 (0.98–1.30)
Total BCAAs ^c			1.24 (1.00–1.54)

^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.

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

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) recommendation for screening of pancreatic cancer reaffirms that potential benefits of screening do not outweigh the potential harm in asymptomatic, average-risk individuals (44). 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 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; refs. 23, 45) or patients with affected family members where the disease screening for these specific populations is being studied (46, 47).

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% confidence interval (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 prediagnostic 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

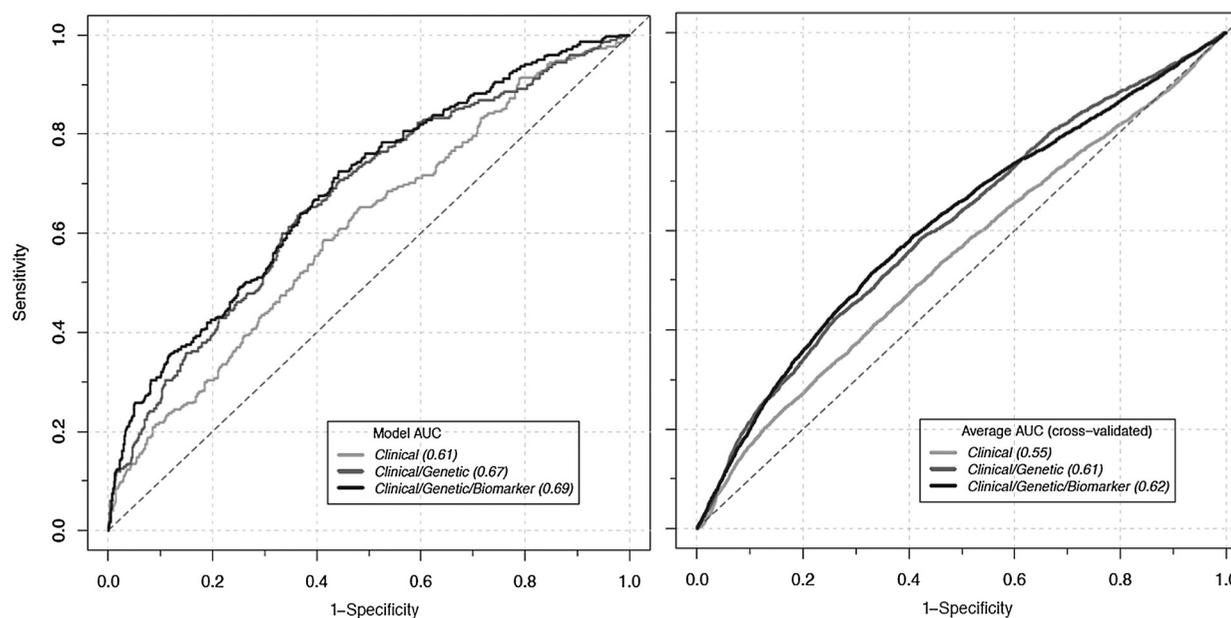


Figure 1.

ROC curves from before (left) and after 5-fold cross-validation (right) in the 0–10-year follow-up population. Each line represents the clinical model (light gray), the clinical/genetic model (gray), and the clinical/genetic/biomarker model (dark gray). 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.

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, 48, 49). 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 (50). 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 3 years before diagnosis, so risk models for the general population will remain necessary to diagnosis of 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. In addition, 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. Because 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, because this study did not include nonwhite 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 prediagnostically, 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 United States, 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

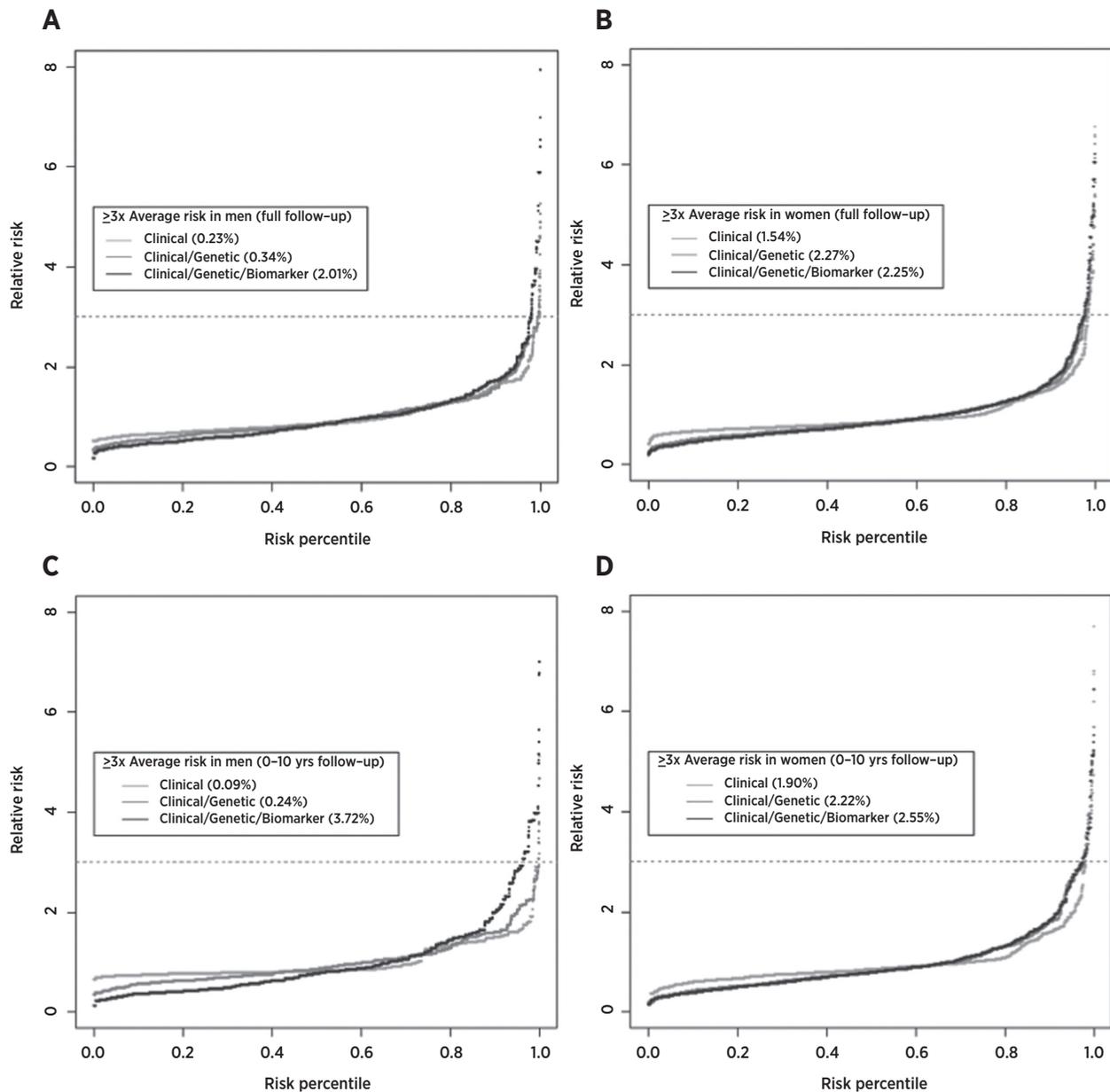


Figure 2. Pancreatic cancer risk in the general population. The data were simulated with a total of 20,000 men and 20,000 women based on the average of our imputed datasets using external risk estimates for smoking status and family history of pancreatic cancer. The relative risk of pancreatic cancer 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 to 10 years of follow-up, and (D) women in 0 to 10 years of follow-up. The lines represent three risk models, including clinical factors only (light gray), clinical and genetic factors (gray), and clinical and genetic factors as well as circulating biomarkers (dark gray).

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., MRI 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 (51–54).

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

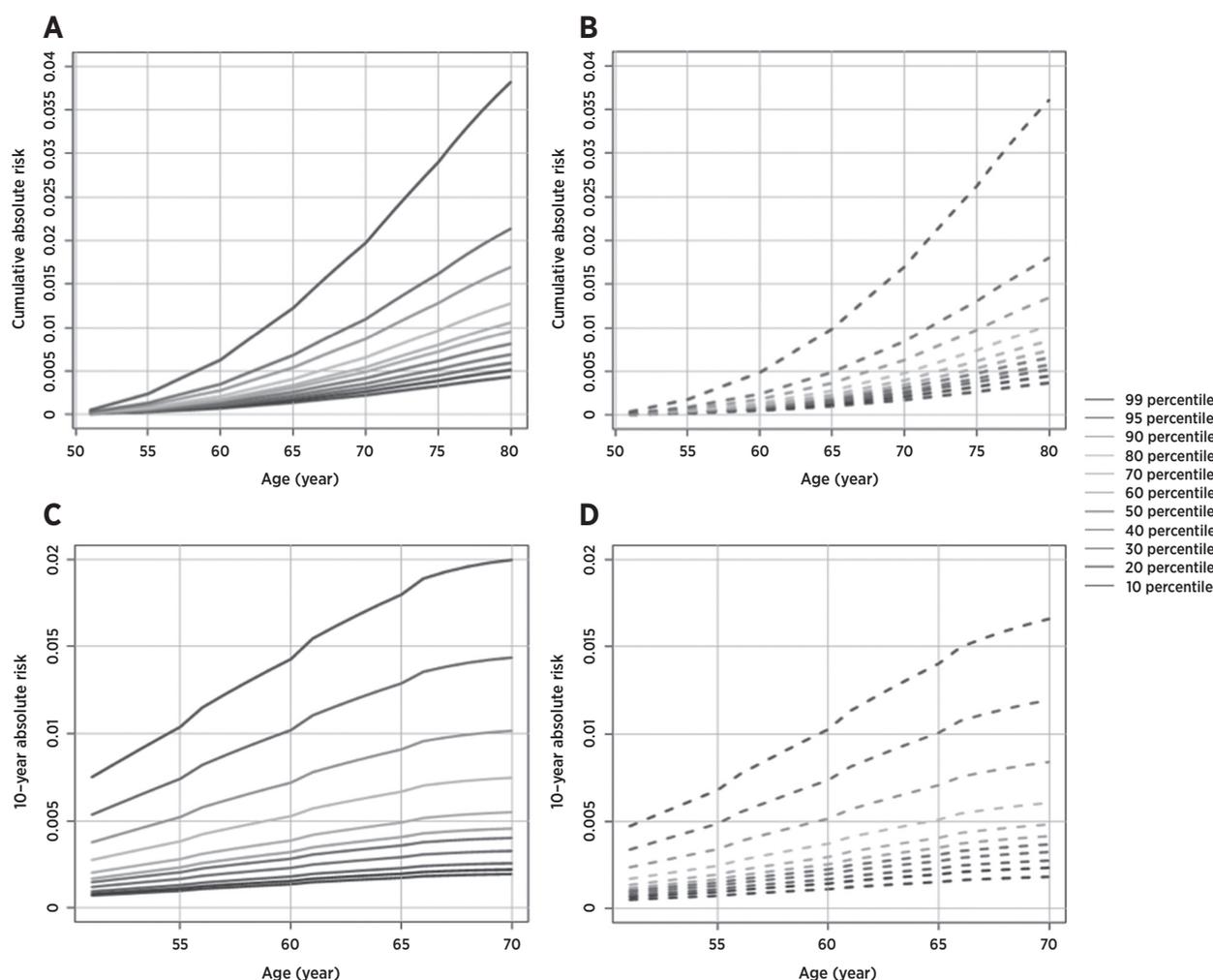


Figure 3.

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, WHR, physical activity, diagnosed diabetes, wGRS, proinsulin, adiponectin, IL6, and total BCAAs). **A**, Men in the full follow-up population. **B**, Women in the full follow-up population. **C**, Men in the 0 to 10 years of follow-up population. **D**, Women in the 0 to 10 years of follow-up population.

pancreatic cancer, soon to be the second leading cause of cancer death in the United States.

Disclosure of Potential Conflicts of Interest

C.S. Fuchs is a consultant for Agios, Bain Capital, Taiho, Unum Therapeutics, Daiichi Sankyo, Bayer, Celgene, Eli Lilly, Entrinsic Health, Genentech, Merck, Merrimack Pharma, and Sanofi, and has ownership interest (including patents) in CytomX Therapeutics and Entrinsic Health. B.M. Wolpin is a consultant for Celgene, GRAIL, and BioLineRx, and reports receiving commercial research grants from Celgene and Eli Lilly. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: A.P. Klein, E.L. Giovannucci, J.E. Manson, C.S. Fuchs, B.M. Wolpin, P. Kraft

Development of methodology: M.N. Pollak, B.M. Wolpin, P. Kraft

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): C.B. Clish, M.N. Pollak, L.T. Amundadottir, R.Z. Stolzenberg-Solomon, L.K. Brais, K. Ng, E.L. Giovannucci, H.D. Sesso, J.E. Manson, C.S. Fuchs

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): J. Kim, C. Yuan, A. Babic, Y. Bao, M.N. Pollak, L.T. Amundadottir, R.Z. Stolzenberg-Solomon, P.V. Pandharipande, K. Ng, B.M. Wolpin, P. Kraft

Writing, review, and/or revision of the manuscript: J. Kim, C. Yuan, A. Babic, Y. Bao, C.B. Clish, M.N. Pollak, L.T. Amundadottir, A.P. Klein, R.Z. Stolzenberg-Solomon, P.V. Pandharipande, L.K. Brais, K. Ng, E.L. Giovannucci, H.D. Sesso, J.E. Manson, M.J. Stampfer, C.S. Fuchs, B.M. Wolpin, P. Kraft

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): J. Kim, Y. Bao, L.K. Brais, M.W. Welch, J.E. Manson, M.J. Stampfer, P. Kraft

Study supervision: B.M. Wolpin, P. Kraft

Acknowledgments

The authors would like to thank the participants and staff of the HPFS, NHS, PHS, and WHI 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 HPFS; UM1CA186107 (M.J. Stampfer), P01CA87969 (R. Tamimi), and R01CA49449 (S. Hankinson) for the Nurses' Health Study; R01CA97193 (J.M. Gaziano), R01CA34944 (C. Hennekens), R01CA40360 (J. Buring), R01HL26490 (C. Hennekens), and R01HL34595 (C. Hennekens) for the PHS; 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. Chiebowksi), 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 WHI program] from the 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, including a Stand Up To Cancer-Lustgarten Foundation Pancreatic Cancer Interception Translational Cancer Research Grant (Grant Number: SU2C-AACR-DT25-17), with additional support from Pancreatic Action Network, Noble Effort Fund, and Promises for Purple. Stand Up To Cancer is a division of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the scientific partner of SU2C. K. Ng acknowledges research funding from the Broman Fund for Pancreatic Cancer Research.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received November 7, 2019; revised January 31, 2020; accepted February 7, 2020; published first April 22, 2020.

References

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin* 2018;68:7–30.
- Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* 2014;74:2913–21.
- Vincent A, Herman J, Schulick R, Hruban RH, Goggins M. Pancreatic cancer. *Lancet* 2011;378:607–20.
- Paniccia A, Hosokawa P, Henderson W, Schulick RD, Edil BH, McCarter MD, et al. Characteristics of 10-year survivors of pancreatic ductal adenocarcinoma. *JAMA Surg* 2015;150:701–10.
- Lynch SM, Vrieling A, Lubin JH, Kraft P, Mendelsohn JB, Hartge P, et al. Cigarette smoking and pancreatic cancer: a pooled analysis from the pancreatic cancer cohort consortium. *Am J Epidemiol* 2009;170:403–13.
- Michaud DS, Giovannucci E, Willett WC, Colditz GA, Stampfer MJ, Fuchs CS. Physical activity, obesity, height, and the risk of pancreatic cancer. *JAMA* 2001;286:921–9.
- Silverman DT, Schiffman M, Everhart J, Goldstein A, Lillemoe KD, Swanson GM, 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.
- Wolpin BM, Bao Y, Qian ZR, Wu C, Kraft P, Ogino S, et al. Hyperglycemia, insulin resistance, impaired pancreatic beta-cell function, and risk of pancreatic cancer. *J Natl Cancer Inst* 2013;105:1027–35.
- Stolzenberg-Solomon RZ, Graubard BI, Chari S, Limburg P, Taylor PR, Virtamo J, et al. Insulin, glucose, insulin resistance, and pancreatic cancer in male smokers. *JAMA* 2005;294:2872–8.
- 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.
- Wolpin BM, Michaud DS, Giovannucci EL, Schernhammer ES, Stampfer MJ, Manson JE, et al. Circulating insulin-like growth factor binding protein-1 and the risk of pancreatic cancer. *Cancer Res* 2007;67:7923–8.
- Wolpin BM, Ng K, Bao Y, Kraft P, Stampfer MJ, Michaud DS, et al. Plasma 25-hydroxyvitamin D and risk of pancreatic cancer. *Cancer Epidemiol Biomarkers Prev* 2012;21:82–91.
- Bao Y, Giovannucci EL, Kraft P, Stampfer MJ, Ogino S, Ma J, et al. A prospective study of plasma adiponectin and pancreatic cancer risk in five US cohorts. *J Natl Cancer Inst* 2013;105:95–103.
- White DL, Hoogeveen RC, Chen L, Richardson P, Ravishankar M, Shah P, 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;7:2180–91.
- Babic A, Bao Y, Qian ZR, Yuan C, Giovannucci EL, Aschard H, et al. Pancreatic cancer risk associated with prediagnostic plasma levels of leptin and leptin receptor genetic polymorphisms. *Cancer Res* 2016;76:7160–7.
- Stolzenberg-Solomon RZ, Newton CC, Silverman DT, Pollak M, Nogueira LM, Weinstein SJ, et al. Circulating leptin and risk of pancreatic cancer: a pooled analysis from 3 cohorts. *Am J Epidemiol* 2015;182:187–97.
- Vainer N, Dehendorff 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;9:29820–41.
- Mayers JR, Wu C, Clish CB, Kraft P, Torrence ME, Fiske BP, et al. Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development. *Nat Med* 2014;20:1193–8.
- Katagiri R, Goto A, Nakagawa T, Nishiumi S, Kobayashi T, Hidaka A, 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;155:1474–82.e1.
- Yip-Schneider MT, Simpson R, Carr RA, Wu H, Fan H, Liu Z, et al. Circulating leptin and branched chain amino acids-correlation with intraductal papillary mucinous neoplasm dysplastic grade. *J Gastrointest Surg* 2019;23:966–74.
- Shindo K, Yu J, Suenaga M, Fesharakizadeh S, Cho C, Macgregor-Das A, et al. Deleterious germline mutations in patients with apparently sporadic pancreatic adenocarcinoma. *J Clin Oncol* 2017;35:3382–90.
- Yurgelun MB, Chittenden AB, Morales-Oyarvide V, Rubinson DA, Dunne RF, Kozak MM, et al. Germline cancer susceptibility gene variants, somatic second hits, and survival outcomes in patients with resected pancreatic cancer. *Genet Med* 2019;21:213–23.
- Hu C, Hart SN, Polley EC, Gnanaolivu R, Shimelis H, Lee KY, et al. Association between inherited germline mutations in cancer predisposition genes and risk of pancreatic cancer. *JAMA* 2018;319:2401–9.
- Lu Y, Ek WE, Whiteman D, Vaughan TL, Spurdle AB, Easton DF, et al. Most common 'sporadic' cancers have a significant germline genetic component. *Hum Mol Genet* 2014;23:6112–8.
- Amundadottir L, Kraft P, Stolzenberg-Solomon RZ, Fuchs CS, Petersen GM, Arslan AA, et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nat Genet* 2009;41:986–90.
- Petersen GM, Amundadottir L, Fuchs CS, Kraft P, Stolzenberg-Solomon RZ, Jacobs KB, 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.
- Wolpin BM, Rizzato C, Kraft P, Kooperberg C, Petersen GM, Wang Z, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. *Nat Genet* 2014;46:994–1000.
- Childs EJ, Mucci E, Campa D, Bracci PM, Gallinger S, Goggins M, et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. *Nat Genet* 2015;47:911–6.
- Zhang M, Wang Z, Obazee O, Jia J, Childs EJ, Hoskins J, et al. Three new pancreatic cancer susceptibility signals identified on chromosomes 1q32.1, 5p15.33 and 8q24.21. *Oncotarget* 2016;7:66328–43.

30. Klein AP, Wolpin BM, Risch HA, Stolzenberg-Solomon RZ, Mocci E, Zhang M, et al. Genome-wide meta-analysis identifies five new susceptibility loci for pancreatic cancer. *Nat Commun* 2018;9:556.
31. Klein AP, Lindstrom S, Mendelsohn JB, Stepilowski E, Arslan AA, Bueno-de-Mesquita HB, 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, Kidd MS. Detectable symptomatology preceding the diagnosis of pancreatic cancer and absolute risk of pancreatic cancer diagnosis. *Am J Epidemiol* 2015;182:26–34.
33. Boursi B, Finkelman B, Giantonio BJ, Haynes K, Rustgi AK, Rhim AD, et al. A clinical prediction model to assess risk for pancreatic cancer among patients with new-onset diabetes. *Gastroenterology* 2017;152:840–50.e3.
34. Giovannucci E, Ascherio A, Rimm EB, Colditz GA, Stampfer MJ, Willett WC. Physical activity, obesity, and risk for colon cancer and adenoma in men. *Ann Intern Med* 1995;122:327–34.
35. Colditz GA, Hankinson SE. The Nurses' Health Study: lifestyle and health among women. *Nat Rev Cancer* 2005;5:388–96.
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.
37. Langer RD, White E, Lewis CE, Kotchen JM, Hendrix SL, Trevisan M. The Women's Health Initiative Observational Study: baseline characteristics of participants and reliability of baseline measures. *Ann Epidemiol* 2003;13:S107–21.
38. Bao Y, Giovannucci EL, Kraft P, Qian ZR, Wu C, Ogino S, 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.
39. Dupont WD. Converting relative risks to absolute risks: a graphical approach. *Stat Med* 1989;8:641–51.
40. Gail MH, Pfeiffer RM. On criteria for evaluating models of absolute risk. *Biostatistics* 2005;6:227–39.
41. Centers for Disease Control and Prevention, National Center for Health Statistics. Tables of summary health statistics. Volume 2018, 2009. Available from: <https://www.cdc.gov/nchs/nhis/SHS/tables.htm>.
42. Jacobs EJ, Chanock SJ, Fuchs CS, LaCroix A, McWilliams RR, Stepilowski E, 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.
43. Centers for Disease Control and Prevention, National Center for Health Statistics. Detailed technical notes to the United States 2007 data—mortality. Volume 2018, 2010. Available from: https://www.cdc.gov/nchs/data/dvs/MortFinal2007_Worktable12.pdf.
44. Owens DK, Davidson KW, Krist AH, Barry MJ, Cabana M, Caughey AB, et al. Screening for pancreatic cancer: US Preventive Services Task Force Reaffirmation Recommendation Statement. *JAMA* 2019;322:438–444.
45. Petersen GM. Familial pancreatic adenocarcinoma. *Hematol Oncol Clin North Am* 2015;29:641–53.
46. Lucas AL, Kastrinos F. Screening for pancreatic cancer. *JAMA* 2019;322:407–8.
47. Canto MI, Harinck F, Hruban RH, Offerhaus GJ, Poley J-W, Kamel I, 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.
48. Munigala S, Singh A, Gelrud A, Agarwal B. Predictors for pancreatic cancer diagnosis following new-onset diabetes mellitus. *Clin Transl Gastroenterol* 2015;6:e118.
49. Sharma A, Kandlakunta H, Nagpal SJS, Feng Z, Hoos W, Petersen GM, et al. Model to determine risk of pancreatic cancer in patients with new-onset diabetes. *Gastroenterology* 2018;155:730–9.e3.
50. Chari S, Leibson C, Rabe K, Ransom J, Deandrade M, Petersen G. Probability of pancreatic cancer following diabetes: a population-based study. *Gastroenterology* 2005;129:504–11.
51. Cohen JD, Javed AA, Thoburn C, Wong F, Tie J, Gibbs P, 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;114:10202–7.
52. Fahrman JF, Bantis LE, Capello M, Scelo G, Dennison JB, Patel N, et al. A plasma-derived protein-metabolite multiplexed panel for early-stage pancreatic cancer. *J Natl Cancer Inst* 2019;111:372–9.
53. Koay EJ, Lee Y, Cristini V, Lowengrub JS, Kang Y, Lucas FAS, et al. A visually apparent and quantifiable CT imaging feature identifies biophysical subtypes of pancreatic ductal adenocarcinoma. *Clin Cancer Res* 2018;24:5883–94.
54. Abou-Elkacem L, Wang H, Chowdhury SM, Kimura RH, Bachawal SV, Gambhir SS, et al. Thy1-targeted microbubbles for ultrasound molecular imaging of pancreatic ductal adenocarcinoma. *Clin Cancer Res* 2018;24:1574–85.