BMJ Open

Impact of race, sex and age on the probability of pancreatic cancer among patients with newly diagnosed diabetes in a claims-based cohort

Elham Afghani [1], Bryan Lau,[2,3] Laura K Becker,[4] Michael Goggins [1,3,5] Alison P Klein [1,2,3,5]

To cite: Afghani E, Lau B, Becker LK, *et al*. Impact of race, sex and age on the probability of pancreatic cancer among patients with newly diagnosed diabetes in a claims-based cohort. *BMJ Open* 2025;**15**:e099488. doi:10.1136/bmjopen-2025-099488

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2025-099488).

Received 18 January 2025
Accepted 30 April 2025

For numbered affiliations see end of article.

Correspondence to
Dr Alison P Klein;
aklein1@jhmi.edu

ABSTRACT

Objective Pancreatic cancer diagnoses are frequently preceded by a new diabetes diagnosis. Screening individuals newly diagnosed with diabetes could enable the earlier detection of pancreatic cancer. We sought to estimate the risk of pancreatic cancer by age, sex, race and time since diabetes diagnosis.

Design Claims-based cohort study.

Setting Johns Hopkins Medicine conducted this deidentified claims-based cohort study using the Optum Labs Data Warehouse.

Participants Insurance enrollees from 1/2008–9/2018 were identified as non-diabetic or newly diagnosed diabetic. Our risk set included 4 732 313 individuals (424 129 newly diabetic) in 5 844 934 enrolment periods.

Primary outcome measures Time to pancreatic cancer. Diabetes and cancer were defined using *International Classification of Diseases* (ICD)-9/10 codes.

Results Individuals with newly diagnosed diabetes were at an increased HR of pancreatic cancer, but this effect waned over time. The HR of pancreatic cancer following a diabetes diagnosis was higher in younger individuals and varied by race (lower HR in non-White individuals) (p<0.01, main effects and interactions). Thus, the probability of pancreatic cancer following a diabetes diagnosis was dependent on age, race and sex. For example, the 1-year probability of pancreatic cancer in a White male aged 75 was 0.45% (95%CI 0.41% to 0.49%) if they were newly diagnosed with diabetes and 0.090% (95%CI 0.084% to 0.096%) if they were free of diabetes. In contrast, the risk was lower at 0.15% (new-diabetic, 95% CI 0.13% to 0.16%) and 0.022% (diabetes free, 95%CI 0.020% to 0.023%) at age 55. The HR of pancreatic cancer for individuals with newly diagnosed diabetes compared with those free of diabetes was highest in the month following diagnosis (HR=14.7 and 9.6 for a 55 and 75 year old White male, respectively) but decreased in the following months, with a HR of 7.8 and 5.8 at 3 months, 5.6 and 4.1 at 6 months, and 3.9 and 2.8 at 1 year (p<0.01).

Conclusions Consideration of the age–race–sex specific probability of pancreatic cancer and time since diabetes diagnosis is necessary when evaluating the risk of pancreatic cancer following a diabetes diagnosis.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This study uses real world claims data from the Optum Labs Data warehouse which captures a large cross section of the United States, and therefore the timing of diabetes diagnosis and cancer diagnoses represent that observed in routine clinical care.
⇒ The large sample size allows for the examination of interactions between age, race and sex on the impact of a recent diabetes diagnosis on the subsequent risk of pancreatic cancer.
⇒ This study helps quantify the risk of pancreatic cancer after a new diabetes diagnosis and demonstrates the potential yield and window of opportunity of pancreatic cancer screening in individuals with newly diagnosed diabetes.
⇒ We were unable to include laboratory measures (Hemoglobin A1c) and clinical data (Body Mass Index) given our cohort was limited to claims-based data.
⇒ The cohort was developed from commercial insurance claims data and thus does not capture individuals with other types of insurance.

INTRODUCTION

Pancreatic cancer is the third leading cause of cancer-related death in the United States with an estimated 60 430 new cases and 48 220 deaths in 2021.[1] Globally, the number of pancreatic cancer cases diagnosed annually has doubled since 1990, and age-adjusted incidence rates have increased from 5 per 100 000 person-years in 1990 to 5.7 per 100 000 person-years in 2017.[2] Pancreatic cancer is usually clinically silent in its early stages. The advanced stage at diagnosis is a major factor in the low 5-year survival rates, currently 10%, the lowest of any major tumour site.[1 3]

Studies have consistently identified an association between pancreatic cancer and diabetes.[4–6] Long-standing diabetes is associated with an approximately twofold increased risk of pancreatic cancer.[7] However, many patients with pancreatic cancer present with a recent diagnosis of diabetes,[7–18] and up to 85% of patients with pancreatic cancer have

diabetes or impaired fasting glucose.[10] Resection of the tumour frequently leads to resolution of the diabetes, suggesting that this is a manifestation of the cancer[10] and it may be one of the few early manifestations of an otherwise silent disease.

One of the first studies to examine pancreatic cancer risk in individuals with newly diagnosed diabetes was a retrospective study within the Mayo Clinic population, which indicated up to 1% of patients develop pancreatic cancer within 3 years of their diabetes diagnosis.[14] However, subsequent studies conducted with the VA (United States Veterans Adminstration) system found the risk of pancreatic cancer to be lower, with <0.3% developing pancreatic cancer within 3 years of their diabetes diagnosis compared with 0.11% in patients without diabetes.[11 12] A recent study using electronic health record (EHR) data within the Kaiser Permanente Southern California system found that the risk of pancreatic cancer was increased sevenfold in those with diabetes diagnosed within 1 year of cancer diagnosis.[16] Other studies have suggested a~4.5–5 increased odds of having pancreatic cancer in the year following a diabetes diagnosis.[15 19] However, these studies had a limited ability to examine how age, a major predictor of pancreatic cancer risk, as well as race and sex impact the probability of a pancreatic cancer diagnosis following a diabetes diagnosis. A better understanding of the time-dependent probability of a pancreatic cancer diagnosis following a diabetes diagnosis, and how age and race modify this probability, is necessary to understand the potential benefit of focused pancreatic cancer screening in individuals with a new diabetes diagnosis. This includes an understanding of the lead time between a diabetes claim and subsequent cancer diagnosis reflecting the window of opportunity for earlier detection of pancreatic cancer.

There are ongoing efforts to enrol individuals with a new diabetes diagnosis to determine how to effectively identify the subset with underlying pancreatic cancer.[20 21] A better understanding of how the probability of a pancreatic cancer diagnosis following a new diabetes diagnosis varies with age, race, sex and time since diabetes diagnosis can provide insight into the potential role for earlier detection in this population. Furthermore, estimation of these risks using real-world data can provide insights beyond what can be learnt in more controlled observational settings. The goal of this study was to use claims data to quantify the probability of a pancreatic cancer diagnosis as a function of time following the initial diabetes claim while accounting for age, race and sex to demonstrate the window for early detection. This informs the potential yield of pancreatic cancer screening interventions in the real-world settings of individuals from US with health insurance.

METHODS

Our claims-based cohort was defined using de-identified administrative claims data from the Optum Labs Data Warehouse (OLDW). Since this study involved analysis of pre-existing, de-identified data, it was exempt from the Institutional Review Board review. The database contains longitudinal health information on enrollees and patients, representing a mixture of ages and geographies across the United States. The claims data in OLDW includes medical and pharmacy claims, and enrolment records for commercial and Medicare Advantage enrollees.[22] Demographic information, health plan enrolment status, inpatient and outpatient medical encounters were coded using the *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)*, the *International Statistical Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM), Current Procedural Terminology, fourth edition (CPT-4*) and filled prescriptions (including the National Drug Code numbers, quantity dispensed and days' supply) were recorded for each patient (See online supplemental table S1 for codes used). Ethnicity was assigned by an external vendor who uses a rule-based system that combines analysis of first names, middle names, surnames, and surname prefixes and suffixes with geographic criteria. Optum Labs then assigns these ethnicity values into one of five compliance-determined race code values: W (Non-Hispanic White), B (Non-Hispanic Black), H (Hispanic), A (Asian) and U (Unknown).

Our cohort was limited to individuals aged 40–88 years with a minimum of 12 months of continuous enrolment (with both medical and pharmacy benefits) from 2007 to 2018. Using claims data, we identified two groups of individuals: (1) those with an incident diabetes diagnosis and (2) individuals without diabetes claims. *Individuals with newly diagnosed diabetes were defined as individuals* with (1) two non-diagnostic medical claims for diabetes at least 30 days apart, or (2) one non-diagnostic medical claim for diabetes and a prescription filled for an oral anti-diabetic or insulin medication. The earliest claim associated with diabetes was identified as the index date. ICD codes covering all subtypes of diabetes, and pharmaceutical claims for hypoglycaemic medications and insulin were considered evidence of diabetes (see online supplemental table S1 for codes used). To exclude individuals with prevalent diabetes, we required a minimum of 1 year period without evidence of diabetes (none of the diabetes defining claims) prior to the index data. *Non-diabetic comparison group* was defined as individuals without evidence of diabetes claims (no non-diagnostic medical claims nor diabetic prescriptions). For these individuals, if they had no evidence of diabetes at any time before the index date (defined as enrolment start date plus a randomly generated number of days), based on a γ distribution, α=1.0 and β=1100 was used at the start time in a time-to-event analysis to mirror the distribution in the incident diabetes cohort. Individuals may contribute more than one enrolment period (ie, individuals may enter and exit the medical plan, and each time continuous enrolment starts, a wash-out period of 1 year was evaluated for study inclusion/exclusion criteria), as

such, a non-diabetic individual may later become diabetic and thus contribute person–time accordingly. For both groups, follow-up time was censored at the first non-diagnostic claim for pancreatic cancer based on ICD-9/10 codes (online supplemental table S1) and end of health plan enrolment or 6 years after index date. (See online supplemental table S2 for details). Additional exclusions include pregnancy in the first year of baseline period (due to the possibility of gestational diabetes), diagnosis of cancer in the year prior to the index date and those diagnosed with pancreatic neuroendocrine tumour.

Statistical analysis

Time-to-event analysis was conducted using a parametric spline-based survival model which allowed for a better fit to the data over a traditional Cox survival model. Flexible parametric spline models with one to four knots[23] were fit to the data; covariates were plotted individually to verify fit, and multivariable model fit was determined using the Akaike Information Criterion. The best-fit model had two internal knots and two ancillary spline parameters modelled as linear functions of time (diabetes diagnosis or start of follow-up) and age, where the effects of diabetes and age are arbitrarily flexible functions of time, with interaction effects of diabetes on age and race. Hazard rates and survival estimates derived from the multivariable model were used to estimate race, sex and age-specific hazard ratio ($h_{\text{diabetes}}/h_{\text{no-diabetes}}$), and probability estimates ($(1\,s(t))*100$), where $S(t)$ is the survival function estimate at time t for either individuals with diabetes or individuals without diabetes as appropriate at specified time points.

Proportional hazards assumptions and Kaplan–Meier curves were examined in RStudio[24] using the survival package.[25] Time-to-event analyses were conducted in R using the flexsurv and flexsurvspline packages.[26] Plots were produced in R using the ggplot2 package.[27]

RESULTS

Of 5 844 934 enrolment periods across 4 732 313 individuals identified in the study population, there were 424 129 enrolment periods after a new diabetes diagnosis (7.3%) and 5 420 805 periods for non-diabetic individuals. Table 1 shows the characteristics of the enrolment periods. The mean age of enrolment periods was 55 (range 40–88) years, with an older mean age for the enrolment periods after a new diabetes diagnosis (59; range 40–88) than enrolment periods among individuals without diabetes (54, range 40–88 years). The median length of an enrolment period was 1.7 years (range 0–6 years), excluding the wash-out

Table 1 Cohort demographics

	Entire cohort		Enrolment periods after new diabetes diagnosis		Person-enrolment periods without diabetes	
	n	(%)	N	(%)	n	(%)
Total	5 844 934		424 129	7.3%	5 420 805	92.7%
Median days of follow-up (range)	632	(1–2192 days)	772	(1–2192 days)	623	(1–2192 days)
Median age (range)	55	(40–88 years)	59	(40–88 years)	54	(40–88 years)
Age						
40–49	1 971 208	33.7%	86 262	20.3%	1 884 946	34.8%
50–59	1 771 670	30.3%	129 139	30.4%	1 642 531	30.3%
60–69	1 167 946	20.0%	109 406	25.8%	1 058 540	19.5%
70–79+	670 155	11.5%	72 606	17.1%	597 549	11.0%
80+	263 955	4.5%	26 716	6.3%	237 239	4.4%
Gender						
Male	2 811 600	48.1%	226 388	53.4%	2 585 212	47.7%
Female	3 033 334	51.9%	197 741	46.6%	2 835 593	52.3%
Race						
White	4 257 138	72.8%	279 872	66.0%	3 977 266	73.4%
Black	594 666	10.2%	64 920	15.3%	529 746	9.8%
Hispanic	551 801	9.4%	50 470	11.9%	501 331	9.2%
Asian	222 040	3.8%	16 351	3.9%	205 689	3.8%
Unknown	219 289	3.8%	12 516	3.0%	206 773	3.8%
Pancreatic cancer						
No	5 839 221	99.9%	422 680	99.7%	5 416 541	99.9%
Yes	5713	0.1%	1449	0.3%	4264	0.1%

Table 2 HR for pancreatic cancer in individuals with newly diagnosed diabetes compared with individuals without diabetes by age, race, and days since diabetes diagnosis/cohort entry

Age	3 months	6 months	9 months	12 months	15 months	18 months	24 months	30 months
Asian								
45	6.893	4.942	4.015	3.468	3.110	2.856	2.517	2.292
55	5.893	4.211	3.431	2.952	2.628	2.415	2.140	1.943
65	5.076	3.612	2.921	2.512	2.240	2.055	1.806	1.647
75	4.351	3.097	2.494	2.134	1.896	1.735	1.531	1.393
Black								
45	7.049	5.038	4.118	3.546	3.181	2.916	2.572	2.346
55	6.047	4.317	3.516	3.010	2.703	2.479	2.188	1.988
65	5.202	3.704	2.992	2.580	2.287	2.096	1.848	1.693
75	4.459	3.170	2.555	2.178	1.944	1.776	1.560	1.428
Hispanic								
45	6.548	4.692	3.831	3.302	2.953	2.708	2.397	2.181
55	5.614	4.027	3.268	2.817	2.516	2.301	2.027	1.843
65	4.829	3.441	2.785	2.383	2.122	1.960	1.717	1.566
75	4.140	2.948	2.376	2.025	1.806	1.654	1.456	1.325
White								
45	9.067	6.522	5.302	4.570	4.101	3.774	3.320	3.029
55	7.794	5.576	4.541	3.895	3.483	3.195	2.827	2.568
65	6.717	4.780	3.867	3.316	2.952	2.711	2.394	2.171
75	5.769	4.084	3.294	2.815	2.505	2.299	2.014	1.839

year. In total, 5713 individuals were diagnosed with pancreatic cancer during follow-up.

We sought to determine how age, race, sex and time since diabetes diagnosis impacted the probability of a pancreatic cancer diagnosis over time. To estimate the time to pancreatic cancer diagnosis in relation to age, sex and race, we included each of these as predictors along with their interactions in a flexible Weibull survival model (HRs are presented in table 2 and model parameters in table 3). The timescale was days following the index date, which for individuals with incident diabetes reflects the date of the first diabetes claim. The hazard of pancreatic cancer in individuals remaining diabetes-free was relatively constant over the study period (table 3) and was lower than individuals with diabetes. Among individuals without diabetes, the HR for pancreatic cancer was higher (HR 1.29 (95%CI 1.22 to 1.36)) in males compared with females and was higher in Black individuals compared with White individuals (HR=1.27 (95%CI 1.15 to 1.39)), whereas hazard ratios were similar in Asian, Hispanic and White individuals (HR 0.92 (95%CI 0.76 to 1.10)) and (HR=1.02 (95%CI 0.91 to 1.15)) for Asian versus White and Hispanic versus White, respectively. In addition, the hazard rate of pancreatic cancer in individuals without diabetes increased with age but was not constant over all ages (see table 3 for model estimates). The HR of pancreatic cancer in individuals with newly diagnosed diabetes compared with those remaining diabetes free was highest immediately following the diabetes diagnosis and decreased as time following diabetes diagnosis increased (see table 3 and figure 1 for more estimates). On average, the HR comparing individuals with new diabetes compared with those remaining diabetes free decreased by ~39% between 1 and 3 months, ~17% from 3 months to 6 months and ~14% from 6 months to 1 year. In addition, the hazard of pancreatic cancer due to a diabetes diagnosis varied by age and race (figure 1, tables 2 and 3). The HR of pancreatic cancer due to diabetes decreased with age (figure 1, table 2) and was highest in White individuals compared with Black, Asian or Hispanic individuals.

Figure 2 shows the predicted cumulative probability of pancreatic cancer per 10 000 by time since selection/diabetes diagnosis. The curves represent specific ages by diabetes status and panels by race. For clarity, cumulative probability estimates at 3, 6, 12, 24 and 36 months timepoints are also shown in table 4 and online supplemental table S5. The cumulative probability was highest in older individuals with a new diabetes diagnosis, with a cumulative probability of 0.450% (95%CI 0.413 to 0.489) at 1 year after diabetes diagnosis in a 75-year-old White male. The HR of pancreatic cancer in individuals with diabetes was highest in White individuals. However, cumulative pancreatic cancer probability in Black individuals with a new diabetes diagnosis was comparable with White individuals with a new diabetes diagnosis at all timepoints with a cumulative probability of 0.422% (95%CI 0.367 to

Table 3 Flexible survival spline hazards model parameter estimates

	(95% CI)	SE	HR (95% CI)	P value
γ0	−16.48 (-17.50 to −15.45)	0.52	–	0.00
γ1	1.12 (0.87 to 1.37)	0.13	–	0.00
γ2	0.03 (0.01 to 0.05)	0.01	–	0.00
γ3	−0.07 (-0.10 to −0.04)	0.02	–	0.00
Newly Diagosed Diabetes	4.47 (3.76 to 5.19)	0.36	87.57 (42.94 to 178.59)	0.00
Age	4.05 (2.29 to 5.81)	0.90	57.63 (9.91 to 335.02)	0.00
Males (ref. female)	0.25 (0.2 to 0.31)	0.03	1.29 (1.22 to 1.36)	0.00
Black (ref. White)	0.24 (0.14 to 0.33)	0.05	1.27 (1.15 to 1.39)	0.00
Hispanic (ref. White)	0.02 (-0.1 to 0.14)	0.06	1.02 (0.91 to 1.15)	0.77
Asian (ref. White)	−0.08 (-0.27 to 0.10)	0.09	0.92 (0.76 to 1)	0.38
Unknown race (ref. White)	0.03 (-0.14 to 0.21)	0.09	1.03 (0.87 to 1.23)	0.73
NOD*Age	−0.91 (-1.22 to −0.6)	0.16	0.4 (0.29 to 0.55)	0.00
NOD*Black	−0.26 (-0.43 to −0.08)	0.09	0.77 (0.65 to 0.92)	0.00
NOD*Hispanic	−0.33 (-0.56 to −0.1)	0.12	0.72 (0.57 to 0.9)	0.00
NOD*Asian	−0.28 (-0.65 to 0.09)	0.19	0.76 (0.52 to 1.09)	0.14
NOD*Unknown race	−0.01 (-0.36 to 0.35)	0.18	0.99 (0.7 to 1.41)	0.97
γ1 (NOD)	−0.36 (-0.53 to 0.19)	0.09	0.7 (0.59 to 0.83)	0.00
γ1 (Age)	0.14 (-0.29 to 0.56)	0.22	1.15 (0.75 to 1.76)	0.54
γ2 (NOD)	0.00 (0.00 to 0.01)	0.00	1.00 (1.00 to 1.01)	0.19
γ2 (Age)	0.01 (-0.01 to 0.02)	0.01	1.01 (0.99 to 1.02)	0.28

0.485) at 1 year after diabetes diagnosis in a 75-year-old Black male. In individuals remaining diabetes-free, the cumulative probability of pancreatic cancer was higher in Black individuals compared with White individuals with a cumulative probability in 1 year of 0.112% (0.102 to 0.124) and 0.090% (0.084 to 0.096) in a 75-year-old non-diabetic Black and White male, respectively. Individuals of Hispanic and Asian race who remained diabetes free had

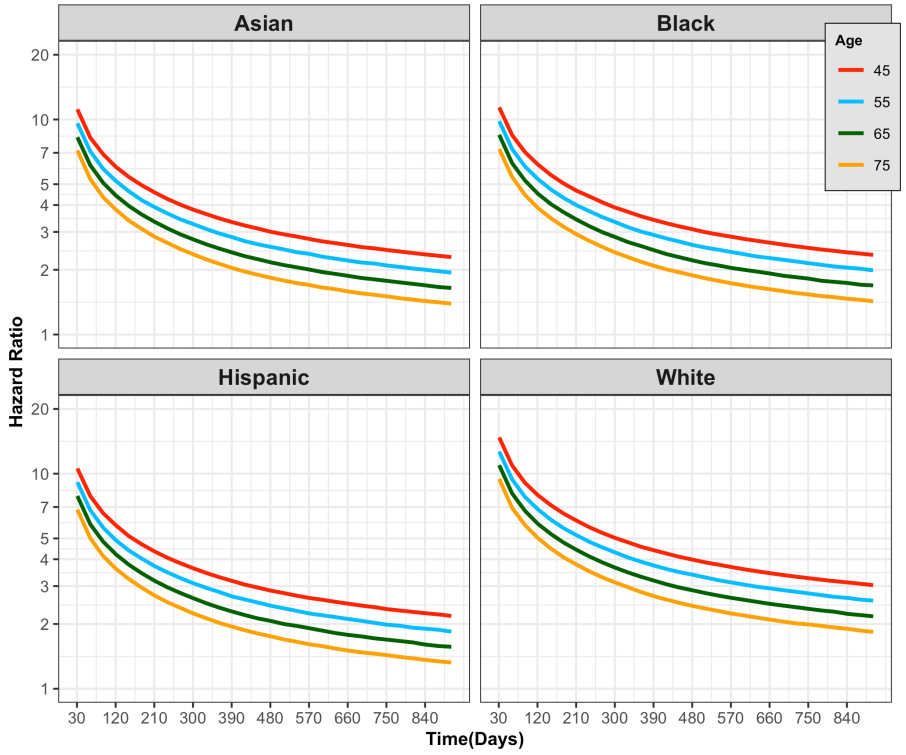

Figure 1 HR of pancreatic cancer in individuals with new diabetes versus Individuals without diabetes by age, race and days since diabetes diagnosis/cohort entry.

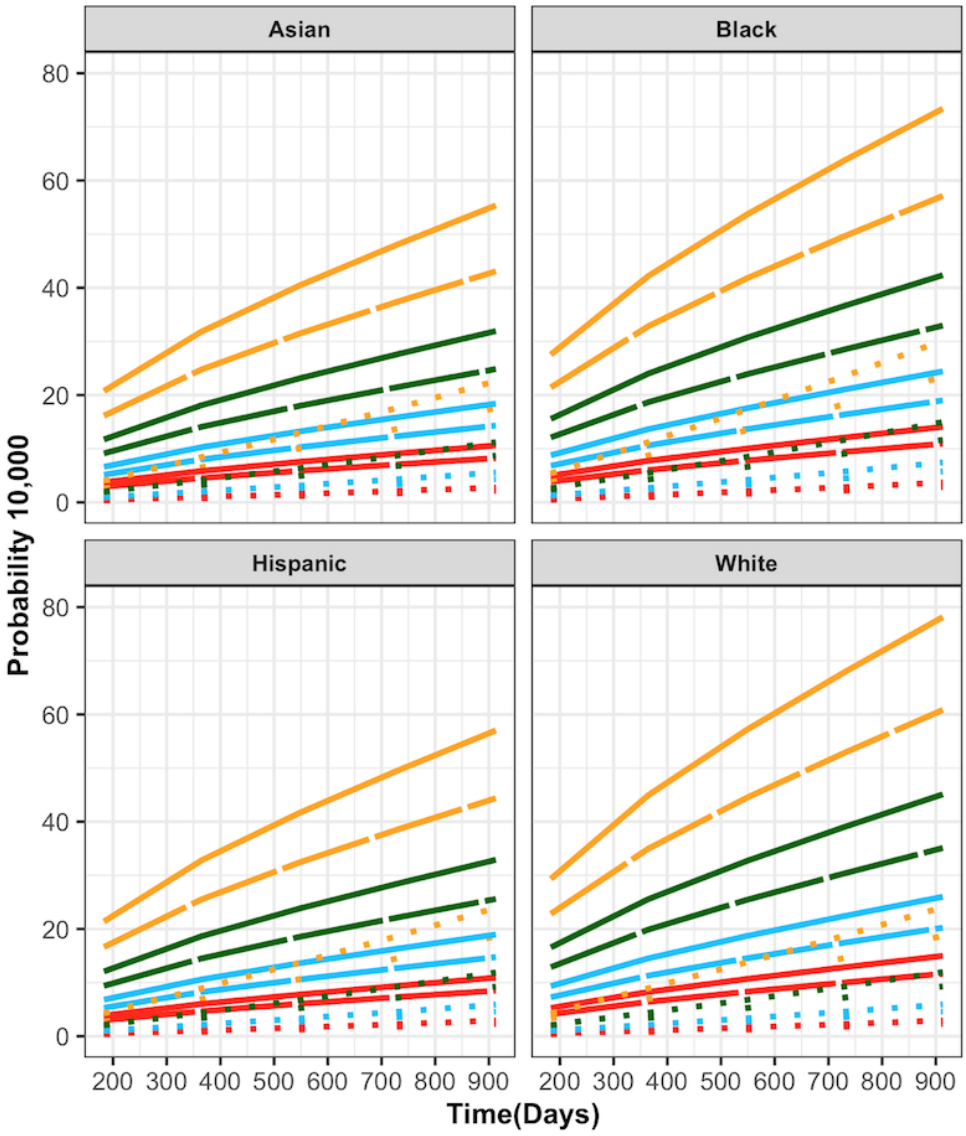

Figure 2 Cumulative probability of a pancreatic cancer diagnosis (per 10 000 individuals) by days since diabetes diagnosis/cohort entry by age, race and diabetes strata for selected ages (1-S_t*100).

pancreatic cancer probabilities comparable with those of White individuals free of diabetes at the same timepoint, cumulative probability at 1 year of 0.089 (95%CI 0.080 to 0.101) and 0.084 (95% CI 0.071 to 0.101)

for a 75-year-old Hispanic and Asian male, respectively. Hispanic and Asian individuals with newly diagnosed diabetes were at a lower probability of pancreatic cancer than Black or White individuals with newly diagnosed diabetes, cumulative probability at 1 years of 0.328 (95%CI 0.270 to 0.392) and 0.318 (95% CI 0.238 to 0.434), respectively.

As pancreatic cancer risk increases dramatically with age, the cumulative probability of pancreatic cancer in older individuals without diabetes is comparable with the probability in younger individuals with a new diabetes diagnosis (figure 2, table 4, online supplemental table S5). Based on these probability estimates, we provide estimates of the number of individuals with newly diagnosed diabetes who would need to undergo screening in order to detect one individual with pancreatic cancer (online supplemental table S3). The estimated number to be screened was slightly higher in Asian individuals and Hispanic individuals compared with Black individuals and White individuals. While the overall number needed to screen was lower in individuals with newly diagnosed diabetes compared with individuals remaining diabetes free, this number varied by age. For example, to detect one individual with pancreatic cancer, 2198 45-year-old Asian females or 1710 45-year-old Asian males would need to be screened at their time of diabetes diagnosis. This is about the same number of 75-year-old Asian individuals without diabetes one would need to screen to detect one individual with pancreatic cancer (1527 and 1188 in females and males, respectively). Similar patterns are seen for other racial groups (online supplemental table S3). These estimates assume a screening test with perfect sensitivity (100%) that would only detect cancers that would present symptomatically within 1 year.

Table 4 Cumulative probability (%) of pancreatic cancer by age, race, sex and time since diagnosis/cohort

		Males				Females			
		3 months	6 months	12 months	24 months	3 months	6 months	12 months	24 months
Age		% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)
Asian									
45	Diabetic	0.021 (0.014 to 0.030)	0.037 (0.027 to 0.054)	0.058 (0.043 to 0.082)	0.091 (0.067 to 0.128)	0.016 (0.011 to 0.023)	0.029 (0.021 to 0.041)	0.045 (0.033 to 0.064)	0.071 (0.052 to 0.100)
	Non-D	0.002 (0.001 to 0.002)	0.005 (0.004 to 0.006)	0.010 (0.00 to 0.012)	0.021 (0.018 to 0.026)	0.001 (0.001 to 0.002)	0.004 (0.003 to 0.004)	0.008 (0.006 to 0.009)	0.017 (0.014 to 0.020)
55	Diabetic	0.037 (0.027 to 0.053)	0.066 (0.048 to 0.092)	0.103 (0.075 to 0.142)	0.159 (0.116 to 0.217)	0.029 (0.021 to 0.041)	0.051 (0.037 to 0.070)	0.080 (0.057 to 0.109)	0.123 (0.089 to 0.167)
	Non-D	0.004 (0.003 to 0.005)	0.010 (0.008 to 0.011)	0.020 (0.017 to 0.024)	0.043 (0.036 to 0.051)	0.003 (0.002 to 0.004)	0.007 (0.006 to 0.009)	0.016 (0.013 to 0.019)	0.033 (0.028 to 0.040)
65	Diabetic	0.068 (0.049 to 0.094)	0.117 (0.085 to 0.159)	0.181 (0.131 to 0.245)	0.276 (0.200 to 0.374)	0.053 (0.038 to 0.072)	0.091 (0.068 to 0.123)	0.141 (0.106 to 0.189)	0.215 (0.163 to 0.288)
	Non-D	0.008 (0.007 to 0.010)	0.020 (0.016 to 0.023)	0.041 (0.034 to 0.049)	0.087 (0.073 to 0.103)	0.006 (0.005 to 0.008)	0.015 (0.013 to 0.018)	0.032 (0.027 to 0.038)	0.068 (0.057 to 0.081)
75	Diabetic	0.123 (0.089 to 0.170)	0.208 (0.153 to 0.283)	0.318 (0.238 to 0.434)	0.481 (0.358 to 0.654)	0.096 (0.069 to 0.131)	0.161 (0.118 to 0.217)	0.248 (0.180 to 0.332)	0.374 (0.277 to 0.501)
	Non-D	0.018 (0.014 to 0.022)	0.040 (0.034 to 0.049)	0.084 (0.071 to 0.101)	0.176 (0.148 to 0.211)	0.014 (0.011 to 0.017)	0.031 (0.026 to 0.038)	0.065 (0.054 to 0.079)	0.137 (0.114 to 0.165)
Black									
45	Diabetic	0.029 (0.024 to 0.037)	0.050 (0.042 to 0.060)	0.078 (0.066 to 0.093)	0.121 (0.104 to 0.143)	0.023 (0.018 to 0.028)	0.039 (0.032 to 0.046)	0.060 (0.051 to 0.071)	0.094 (0.080 to 0.110)
	Non-D	0.003 (0.002 to 0.003)	0.006 (0.005 to 0.007)	0.013 (0.012 to 0.015)	0.028 (0.026 to 0.032)	0.002 (0.002 to 0.002)	0.005 (0.004 to 0.006)	0.010 (0.009 to 0.012)	0.022 (0.020 to 0.025)
55	Diabetic	0.053 (0.045 to 0.062)	0.088 (0.073 to 0.103)	0.137 (0.116 to 0.160)	0.210 (0.181 to 0.244)	0.041 (0.035 to 0.049)	0.068 (0.059 to 0.080)	0.106 (0.092 to 0.123)	0.164 (0.143 to 0.188)
	Non-D	0.005 (0.005 to 0.006)	0.013 (0.011 to 0.014)	0.027 (0.024 to.030)	0.057 (0.052 to 0.063)	0.004 (0.004 to 0.005)	0.010 (0.009 to 0.011)	0.021 (0.019 to 0.023)	0.045 (0.040 to 0.049)
65	Diabetic	0.096 (0.082 to 0.114)	0.156 (0.134 to 0.179)	0.240 (0.208 to 0.275)	0.366 (0.317 to 0.418)	0.074 (0.063 to 0.087)	0.121 (0.105 to 0.139)	0.187 (0.164 to 0.213)	0.285 (0.252 to 0.324)
	Non-D	0.011 (0.010 to 0.013)	0.026 (0.024 to 0.029)	0.055 (0.050 to 0.061)	0.116 (0.107 to 0.128)	0.009 (0.008 to 0.010)	0.020 (0.018 to 0.023)	0.043 (0.039 to 0.047)	0.090 (0.083 to 0.099)
75	Diabetic	0.174 (0.148 to 0.204)	0.275 (0.238 to 0.318)	0.422 (0.367 to 0.485)	0.638 (0.552 to 0.733)	0.135 (0.115 to 0.159)	0.214 (0.184 to 0.247)	0.329 (0.285 to 0.378)	0.496 (0.431 to 0.570)
	Non-D	0.024 (0.021 to 0.027)	0.054 (0.048 to 0.060)	0.112 (0.102 to 0.124)	0.235 (0.215 to 0.257)	0.019 (0.016 to 0.021)	0.042 (0.038 to 0.046)	0.087 (0.080 to 0.096)	0.183 (0.168 to 0.199)

Continued

Table 4 Continued

Age		Males				Females			
		3 months	6 months	12 months	24 months	3 months	6 months	12 months	24 months
		% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)	% (95% CI)
Hispanic									
45	Diabetic	0.022 (0.017 to 0.028)	0.038 (0.031to 0.048)	0.060 (0.049 to 0.075)	0.094 (0.077 to 0.115)	0.017 (0.013 to 0.022)	0.030 (0.024 to 0.037)	0.047 (0.038 to 0.058)	0.073 (0.059 to 0.089)
	Non-D	0.002 (0.002 to 0.003)	0.005 (0.004 to 0.006)	0.010 (0.009 to 0.012)	0.023 (0.020 to 0.026)	0.002 (0.001 to 0.002)	0.004 (0.003 to 0.005)	0.008 (0.007 to 0.009)	0.018 (0.015 to 0.020)
55	Diabetic	0.040 (0.032 to 0.048)	0.068 (0.056 to 0.084)	0.106 (0.088 to 0.130)	0.163 (0.137 to 0.198)	0.031 (0.025 to 0.038)	0.053 (0.043, 0.065)	0.082 (0.067 to 0.100)	0.127 (0.104 to 0.153)
	Non-D	0.004 (0.004 to 0.005)	0.010 (0.009 to 0.012)	0.021 (0.019 to 0.024)	0.046 (0.040 to 0.051)	0.003 (0.003 to 0.004)	0.008 (0.007 to 0.009)	0.017 (0.015 to 0.019)	0.036 (0.031 to 0.040)
65	Diabetic	0.072 (0.058 to 0.089)	0.121 (0.100 to 0.146)	0.186 (0.155 to 0.224)	0.285 (0.237 to 0.340)	0.056 (0.045 to 0.068)	0.094 (0.079 to 0.114)	0.145 (0.122 to 0.175)	0.221 (0.186 to 0.266)
	Non-D	0.009 (0.008 to 0.011)	0.021 (0.018 to 0.024)	0.044 (0.039 to 0.049)	0.092 (0.082 to 0.103)	0.007 (0.006 to 0.008)	0.016 (0.014 to 0.018)	0.034 (0.030 to 0.038)	0.072 (0.064 to 0.081)
75	Diabetic	0.130 (0.106 to 0.159)	0.214 (0.174 to 0.256)	0.328 (0.270 to 0.392)	0.495 (0.410 to 0.590)	0.101 (0.081 to 0.123)	0.166 (0.137 to 0.201)	0.255 (0.212 to 0.307)	0.385 (0.323 to 0.461)
	Non-D	0.019 (0.017 to 0.022)	0.043 (0.038 to 0.049)	0.089 (0.080 to 0.101)	0.187 (0.168 to 0.211)	0.015 (0.013 to 0.018)	0.033 (0.029 to 0.038)	0.069 (0.062 to 0.078)	0.145 (0.129 to 0.163)
White									
45	Diabetic	0.030 (0.025 to 0.036)	0.053 (0.045 to 0.062)	0.083 (0.072 to 0.095)	0.129 (0.114 to 0.147)	0.023 (0.019 to 0.028)	0.041 (0.035 to 0.048)	0.064 (0.056 to 0.075)	0.100 (0.087 to 0.114)
	Non-D	0.002 (0.002 to 0.002)	0.005 (0.004 to 0.006)	0.011 (0.009 to 0.012)	0.023 (0.021 to 0.024)	0.002 (0.001 to 0.002)	0.004 (0.003 to 0.004)	0.008 (0.007 to 0.009)	0.018 (0.016 to 0.019)
55	Diabetic	0.054 (0.047 to 0.062)	0.094 (0.084 to 0.105)	0.145 (0.133 to 0.161)	0.224 (0.206 to 0.245)	0.042 (0.0360 to 0.048)	0.073 (0.064 to 0.081)	0.113 (0.102 to 0.125)	0.174 (0.159 to 0.191)
	Non-D	0.004 (0.004 to 0.005)	0.010 (0.009 to 0.011)	0.021 (0.020 to 0.023)	0.046 (0.043 to 0.048)	0.003 (0.003 to 0.004)	0.008 (0.007 to 0.009)	0.017 (0.016 to 0.018)	0.036 (0.034 to 0.038)
65	Diabetic	0.098 (0.088 to 0.109)	0.166 (0.152 to 0.180)	0.256 (0.23 to 0.275)	0.390 (0.365 to 0.416)	0.076 (0.068 to 0.084)	0.129 (0.117 to 0.141)	0.199 (0.183 to 0.215)	0.304 (0.283 to 0.326)
	Non-D	0.009 (0.008 to 0.010)	0.021 (0.019 to 0.022)	0.044 (0.041 to 0.046)	0.093 (0.089 to 0.097)	0.007 (0.006 to 0.008)	0.016 (0.015 to 0.017)	0.034 (0.032 to 0.036)	0.072 (0.069 to 0.076)
75	Diabetic	0.177 (0.159 to 0.198)	0.293 (0.268 to 0.323)	0.450 (0.413 to 0.489)	0.679 (0.627 to 0.731)	0.138 (0.123 to 0.154)	0.228 (0.207 to 0.252)	0.350 (0.320 to 0.382)	0.528 (0.490 to 0.572)
	Non-D	0.019 (0.017 to 0.021)	0.043 (0.040 to 0.046)	0.090 (0.084 to 0.096)	0.188 (0.178 to 0.198)	0.015 (0.013 to 0.016)	0.033 (0.031 to 0.036)	0.070 (0.066 to 0.075)	0.146 (0.13 to 0.154)

DISCUSSION

Our goal was to estimate the age, race and sex-specific risks (figure 2, table 4) of pancreatic cancer among individuals with a new diabetes diagnosis and how hazard ratios (figure 1, table 2) decrease with time from diabetes diagnosis. Such estimates inform the window of opportunity and the potential number of pancreatic cancers that may be detected using a new-onset diabetes-based screening strategy. Overall, our estimated 1-year probability of pancreatic cancer in individuals without diabetes was consistent with Surveillance, Epidemiology and End Results (SEER) rates (see online supplemental table S4), suggesting comparability with the US population.

We did observe that the increased HR of pancreatic cancer was higher in younger individuals with newly diagnosed diabetes compared with individuals who developed diabetes at older ages. Furthermore, the HR also varied by race, with higher hazard ratios in White, Asian and Hispanic individuals compared with Black individuals. In the United States, the prevalence of all-cause diabetes also varies by age and race, with prevalence increasing with age and higher prevalence among Black individuals,[28 29] and we observed these patterns in our data. In our analysis, the HR of pancreatic cancer in individuals with newly diagnosed diabetes compared with individuals remaining free of diabetes followed the opposite pattern, such that the HR was lower in groups with a higher underlying prevalence of diabetes. This lower HR in groups at higher risk of diabetes could reflect a higher relative prevalence of diabetes from non-pancreatic cancer-related causes relative to pancreatic cancer-related diabetes.

Like prior observational studies, we have shown a significantly higher rate of pancreatic cancer among individuals with newly diagnosed diabetes. However, our large sample size allowed us to closely examine how the hazard ratios of a pancreatic cancer diagnosis rapidly declined with time since the first diabetes diagnosis as compared with those remaining free of diabetes. For effective clinical screening, an understanding of the window of opportunity between the diagnosis of diabetes and when a diagnosis of pancreatic cancer based on clinical symptoms would occur is critical to maximise benefit. In all groups, the hazard ratios of pancreatic cancer were greatest immediately following the initial diabetes claim and decreased rapidly over the following months, as shown in figure 1. Thus, the maximum impact of screening immediately follows diabetes diagnosis. As the HR decreased with time, so too does the yield of screening, as most of the excess incidence occurs in the first few months. For example, the cumulative probability of pancreatic cancer 2 years after a diabetes diagnosis in Black men age 75 is 0.63% versus 0.23% in Black men without diabetes, a difference of about 40 cancers per 10 000 individuals at 2 years; approximately 50% (n=21) of these 'excess' cases occur in the first 6 months following the diabetes diagnosis and 75% (n=30) by the end of year one (table 2). The true benefit of screening is further complicated by how quickly an individual with newly diagnosed diabetes could get screening for a potential pancreatic cancer and to what extent detecting these cancers earlier would improve survival. Our estimates of risk and hazard ratios include different populations, whereas most prior reported studies included predominantly White cohorts.[13 16 17 30] Setiawan *et al* conducted a large population-based cohort study of Black and Hispanic individuals and found that a new diabetes diagnosis was associated with a threefold to fourfold increased risk of pancreatic cancer.[30] Li *et al* examined diabetes and pancreatic cancer risk and included race and ethnicity, but their study was underpowered with respect to estimating interactions between age, sex and race.[17] A large population-based cohort study from Kaiser Permanente Southern California with approximately 7.5 million person-years of follow-up found risk ratios for Hispanic and Asian individuals with a new diabetes diagnosis trended higher compared with Black and White individuals.[16] In our analysis, we were well-powered to detect interactions by race and diabetes status. We found that individuals without diabetes have about a 25% higher risk of pancreatic cancer compared with White individuals without diabetes, and there were slight differences in risk between Asian, Hispanic and White individuals. However, among individuals with a new diabetes diagnosis, pancreatic cancer HR was about 30% higher in White individuals compared with Black, Hispanic or Asian individuals (table 2). It is important to note that the OLDW captures claims data from across the US among individuals with commercial insurance and Medicare Advantage. However, there is great regional diversity within racial groups in the US that may result in differences between our study and studies focused on a specific region. In addition, the results do not capture individuals without insurance, Medicaid or Medicare (without Medicare Advantage).

Age-specific rates of pancreatic cancer have also been on the rise.[31] Based on our study, the younger individuals with a new diabetes diagnosis had a higher HR compared with older individuals. However, the absolute risk of pancreatic cancer in younger individuals with newly diagnosed diabetes is very low because of their overall lower risk of pancreatic cancer. In contrast, in older individuals, the HR of pancreatic cancer diagnosis following to a new diabetes diagnosis is lower, but because they have a much higher underlying risk of pancreatic cancer, their absolute risk is much higher. Absolute risk is a critical metric to consider with planning interventional screening studies.

Recent studies within the Nurse's Health Study and Health Professionals Follow-up Study have demonstrated that pancreatic cancer risk is highest in newly-diagnosed diabetic individuals who experience weight loss.[32] A similar trend was observed in an analysis of Kaiser Permanente data which found more rapid increases in glucose levels, HbA1c levels and recent weight loss in patients with a new diagnosis of diabetes who developed pancreatic cancer compared with those who did not.[16] While EHR data is available within the OLDW, it was available only for a fraction of the study

population, limiting our ability to examine interactions by race and age.

Strengths of our study include the ability to examine interactions between race and diabetes status, as well as fully model the probability of pancreatic cancer as a function of time following diabetes diagnosis and how this probability varies by age. The OLDW database includes broad representation of the US population, including a mixture of representation of ages. In addition, the use of OLDW allows us to evaluate enrollees receiving care in different settings in the US.

There are several limitations to our study. First, our database relies on diagnosis based on ICD9, ICD10 and CPT-4 claims codes, and as such, we are not able to examine how factors that may underlie diabetes, such as obesity, impact our probability estimates. Second, race/ethnicity was imputed using a proprietary algorithm. Prior validations of this imputation method have been previously published and demonstrated moderate sensitivity (48%), high specificity (97%) for Black race and high sensitivity (97%) and moderate specificity (47%) for White population.[32] If we applied these sensitivity and specificity estimates to our study population, the positive predictive value for race in the sample set is predicted to be 85% and 87% for Black and White population, respectively. Under the assumption that misclassification of race is non-differential by diabetes status and pancreatic cancer, it is likely that this misclassification would attenuate our observed effects. Our model-based probabilities in individuals without diabetes are comparable with (SEER) rates (online supplemental table S4),[33] supporting the imputation algorithm. Our population included only those with commercial insurance and older individuals with Medicare Advantage; thus, our results must be carefully considered when applying to individuals without insurance, Medicaid or Medicare. Also, some individuals contributed to multiple enrolment periods, but our time-event model accounted for this and sensitivity analysis (randomly selecting one enrolment period per individual) yielded similar results (results not shown). Future studies examining these findings in Medicare and Medicaid populations may be warranted.

In conclusion, in individuals with newly diagnosed diabetes, the probability of pancreatic cancer varied by age, race, sex and time since diabetes diagnosis. While individuals with a new diabetes diagnosis do have a markedly increased risk of pancreatic cancer, the window of opportunity for earlier detection is short as the hazard rapidly decreases with time following diabetes diagnosis, highlighting both the potential and challenges of early detection in this population.

Patient and public involvement

This study used de-identified data as such patients were not involved in the research.

Author affiliations
[1]Division of Gastroenterology, Department of Medicine, The Johns Hopkins Medical Institutions, Baltimore, MD, USA
[2]Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland, USA
[3]Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins Medicine, Baltimore, Maryland, USA
[4]OptumLabs, Eden Prairie, Minnesota, USA
[5]Department of Pathology, Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins University, Baltimore, Maryland, USA

Contributors EA (writing and interpretation of data), BL (writing, editing and methodology supervision), LKB (data curation, data analysis, editing of manuscript), MG (study design, editing of manuscript, funding) and APK (study design, study supervision, supervision of analysis, writing, editing, funding). APK is the guarantor.

Funding This work was supported by the Stand Up To Cancer-Lustgarten Foundation Pancreatic Cancer Interception Translational Cancer Research Grant (Grant Number: SU2C-AACR-DT25–17). We would also like to acknowledge Nancy Porter for her assistance in this work.

Competing interests Material support for the study described in this publication was provided by Optum Labs via the Stand up to cancer relationship. Dr Klein served as a paid consultant to Optum Labs in 2018. This arrangement was reviewed and approved by Johns Hopkins University in accordance with its conflicts of interest policies. All other authors have no competing interest to declare. Roles: Elham Afghani, MD (writing and interpretation of data) Bryan Lau, PhD ScM MHS (writing, editing and methodology supervision) Laura Becker, MS (data curation, data analysis, editing of manuscript) Michael G Goggins, MD (study design, editing of manuscript, funding) Alison Klein, PhD MHS (study design, study supervision, supervision of analysis, writing, editing, funding). Dr Klein is the guarantor.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. The data underlying the results of this study are third party data owned by Optum Labs and contain sensitive patient information; therefore the data is only available upon request. Interested researchers engaged in HIPAA compliant research may contact connected@optum.com for data access requests.

ORCID iDs
Elham Afghani http://orcid.org/0000-0002-9014-9735
Michael Goggins http://orcid.org/0000-0002-4286-2296
Alison P Klein http://orcid.org/0000-0003-2737-8399

REFERENCES

1 Siegel RL, Miller KD, Fuchs HE, *et al*. Cancer Statistics, 2021. *CA A Cancer J Clinicians* 2021;71:7–33.
2 Everhart J, Wright D. Diabetes mellitus as a risk factor for pancreatic cancer. A meta-analysis. *JAMA* 1995;273:1605–9.
3 Pancreatic Cancer Collaborators. The global, regional, and national burden of pancreatic cancer and its attributable risk factors in 195 countries and territories, 1990-2017: a systematic analysis for the

Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* 2017;4:934–47.
4 Ragozzino M, Melton LJ 3rd, Chu CP, *et al*. Subsequent cancer risk in the incidence cohort of Rochester, Minnesota, residents with diabetes mellitus. *J Chronic Dis* 1982;35:13–9.
5 Friedman GD, van den Eeden SK. Risk factors for pancreatic cancer: an exploratory study. *Int J Epidemiol* 1993;22:30–7.
6 Gullo L, Pezzilli R, Morselli-Labate AM, *et al*. Diabetes and the risk of pancreatic cancer. *N Engl J Med* 1994;331:81–4.
7 Li D, Tang H, Hassan MM, *et al*. Diabetes and risk of pancreatic cancer: a pooled analysis of three large case–control studies. *Cancer Causes Control* 2011;22:189–97.
8 Cuzick J, Babiker AG. Pancreatic cancer, alcohol, diabetes mellitus and gall-bladder disease. *Int J Cancer* 1989;43:415–21.
9 Pannala R, Leirness JB, Bamlet WR, *et al*. Prevalence and clinical profile of pancreatic cancer-associated diabetes mellitus. *Gastroenterology* 2008;134:981–7.
10 Munigala S, Singh A, Gelrud A, *et al*. Predictors for pancreatic cancer diagnosis following new-onset diabetes mellitus. *Clin Transl Gastroenterol* 2015;6:e118.
11 Gupta S, Vittinghoff E, Bertenthal D, *et al*. New-onset diabetes and pancreatic cancer. *Clin Gastroenterol Hepatol* 2006;4:1366–72; .
12 Wang F, Gupta S, Holly EA. Diabetes mellitus and pancreatic cancer in a population-based case-control study in the San Francisco Bay Area, California. *Cancer Epidemiol Biomarkers Prev* 2006;15:1458–63.
13 Chari ST, Leibson CL, Rabe KG, *et al*. Probability of pancreatic cancer following diabetes: a population-based study. *Gastroenterology* 2005;129:504–11.
14 Ben Q, Cai Q, Li Z, *et al*. The relationship between new-onset diabetes mellitus and pancreatic cancer risk: A case–control study. *Eur J Cancer* 2011;47:248–54.
15 Huang BZ, Pandol SJ, Jeon CY, *et al*. New-onset diabetes, longitudinal trends in metabolic markers, and risk of pancreatic cancer in a heterogeneous population. *Clin Gastroenterol Hepatol* 2020;18:1812–21.
16 Huxley R, Ansary-Moghaddam A, Berrington de González A, *et al*. Type-II diabetes and pancreatic cancer: a meta-analysis of 36 studies. *Br J Cancer* 2005;92:2076–83.
17 Liao K-F, Lai S-W, Li C-I, *et al*. Diabetes mellitus correlates with increased risk of pancreatic cancer: a population-based cohort study in Taiwan. *J Gastroenterol Hepatol* 2012;27:709–13.
18 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.
19 Serrano J, Andersen DK, Forsmark CE, *et al*. Consortium for the study of chronic pancreatitis, diabetes, and pancreatic cancer: from concept to reality. *Pancreas* 2018;47:1208–12.
20 Maitra A, Sharma A, Brand RE, *et al*. A prospective study to establish a new-onset diabetes cohort: from the consortium for the study of chronic pancreatitis, diabetes, and pancreatic cancer. *Pancreas* 2018;47:1244–8.
21 Optum Labs. Optum Labs and Optum Labs Data Warehouse (OLDW) Descriptions and Citation. Eden Prairie, MN. PDF. Reproduced with permission from Optum Labs. 2023.
22 Royston P, Parmar MKB. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat Med* 2002;21:2175–97.
23 RStudio Team. RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA. 2020. Available: http://www.rstudio.com
24 Therneau TM, Grambsch PM. *Modeling survival data: extending the cox model*. New York: Springer, 2000.
25 Jackson CH. flexsurv: a platform for parametric survival modeling in R. *J Stat Softw* 2016;70:1–33.
26 Wickham H. *Ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag, 2016.
27 Centers for Disease Control and Prevention. National diabetes statistics report, 2020, 2020. Available: www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf
28 Menke A, Casagrande S, Geiss L, *et al*. Prevalence of and trends in diabetes among adults in the United States, 1988-2012. *JAMA* 2015;314:1021–9.
29 Setiawan VW, Stram DO, Porcel J, *et al*. Pancreatic cancer following incident diabetes in African Americans and latinos: the multiethnic cohort. *J Natl Cancer Inst* 2019;111:27–33.
30 Gordon-Dseagu VL, Devesa SS, Goggins M, *et al*. Pancreatic cancer incidence trends: evidence from the Surveillance, Epidemiology and End Results (SEER) population-based data. *Int J Epidemiol* 2018;47:427–39.
31 Yuan C, Babic A, Khalaf N, *et al*. Diabetes, weight change, and pancreatic cancer risk. *JAMA Oncol* 2020;6:e202948.
32 DeFrank JT, Bowling JM, Rimer BK, *et al*. Triangulating differential nonresponse by race in a telephone survey. *Prev Chronic Dis* 2007;4:A60.
33 American Cancer Society. *Cancer facts & figures for African Americans 2019-2021*. Atlanta: American Cancer Society, 2019.