Research

Weiqi Liao, Ashley K Clift, Martina Patone, Carol Coupland, Arturo González-Izquierdo, Stephen P Pereira and Julia Hippisley-Cox

Identifying symptoms associated with diagnosis of pancreatic exocrine and neuroendocrine neoplasms:

a nested case-control study of the UK primary care population

Abstract

Background
Pancreatic cancer has the worst survival rate among all cancers. Almost 70% of patients in the UK were diagnosed at Stage IV.

Aim
This study aimed to investigate the symptoms associated with the diagnoses of pancreatic ductal adenocarcinoma (PDAC) and pancreatic neuroendocrine neoplasms (PNEN), and comparatively characterise the symptomatology between the two tumour types to inform earlier diagnosis.

Design and setting
A nested case-control study in primary care was conducted using data from the QResearch® database. Patients aged ≥25 years and diagnosed with PDAC or PNEN during 2000 to 2019 were included as cases. Up to 10 controls from the same general practice were matched with each case by age, sex, and calendar year using incidence density sampling.

Method
Conditional logistic regression was used to investigate the association between the 42 shortlisted symptoms and the diagnoses of PDAC and (or) PNEN in different timeframes relative to the index date, adjusting for patients' sociodemographic characteristics, lifestyle, and relevant comorbidities.

Results
A total of 23 640 patients were identified as diagnosed with PDAC and 596 with PNEN. Of the symptoms identified, 23 were significantly associated with PDAC, and nine symptoms with PNEN. The two alarm symptoms for both tumours were jaundice and gastrointestinal bleeding. The two newly identified symptoms for PDAC were thirst and dark urine. The risk of unintentional weight loss may be longer than 2 years before the diagnosis of PNEN.

Conclusion
PDAC and PNEN have overlapping symptom profiles. The QCancer® (pancreas) risk prediction model could be updated by including the newly identified symptoms and comorbidities, which could help GPs identify high-risk patients for timely investigation in primary care.

Keywords
early diagnosis; pancreatic ductal adenocarcinoma (PDAC); pancreatic neoplasms; pancreatic neuroendocrine neoplasms (PNEN); primary health care; symptom.

INTRODUCTION

Pancreatic cancer is the 10th most common cancer in incidence but represents the fifth most common cause of death owing to cancer in the UK. Pancreatic cancer is very aggressive and has the worst survival rate among all types of cancer.[1] Almost 70% of patients were diagnosed at Stage IV. Tumours arising from the pancreas can be classified as exocrine (approximately 95%) or neuroendocrine (≤5%) neoplasms,[2] known as pancreatic ductal adenocarcinoma (PDAC) and pancreatic neuroendocrine neoplasms (PNEN),[3] respectively. Though there are differences in tumour pathobiology and treatment strategies between PDAC and PNEN, the two tumour types share a proclivity to metastasise.[4–6] More favourable outcomes were observed in diagnosis at earlier stages.[7]

In the absence of a screening programme for pancreatic cancer in the UK, symptomatic presentation in general practice remains a key avenue for earlier diagnosis. However, besides jaundice, current literature reporting symptoms associated with pancreatic cancer are vague and non-specific.[8,9] GPs face the challenges of differentiating a potential malignancy from other benign diseases when patients present with non-specific symptoms, which could be easily 'missed' or delay diagnosis of the tumour. In addition, as a rarer type of cancer, there is still no large-scale study characterising the symptomatology of PNEN, nor have studies comprehensively compared the symptomatology of PDAC with PNEN. Therefore, a better understanding of the symptomatology and the timings at which patients present with symptoms could help GPs better manage patients and make clinical decisions. Furthermore, such symptoms could be used in social campaigns to increase public awareness of pancreatic cancer.

To address this research gap, the authors conducted this study with three aims: to explore the symptoms that patients presented in primary care in different time windows, which may indicate the diagnosis of PDAC/PNEN; to comparatively characterise the symptomatology of PDAC and PNEN; and to inform the update of the QCancer® (pancreas) prediction model.[10,11] The QCancer (pancreas) score quantifies the risk of an incident diagnosis of pancreatic cancer in the next 2 years, based on an individual patient's characteristics. Such a risk score may help GPs make different decisions: a 2-week-wait referral

W Liao (ORCID: 0000-0002-8605-3749), MMed, MRes, data scientist; **M Patone**, PhD, data scientist; **J Hippisley-Cox** (ORCID: 0000-0002-2479-7283), MD, FRCP, FRCGP, professor of clinical epidemiology and general practice, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford. **AK Clift**, MA, MBBS, research fellow, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, and Cancer Research UK Oxford Centre, University of Oxford, Oxford. **C Coupland**, PhD, professor of medical statistics in primary care, Division of Primary Care, School of Medicine, University of Nottingham, Nottingham. **A González-Izquierdo**, PhD, senior research associate, UCL Institute of Health Informatics, Health Data Research UK, London. **SP Pereira**, MD, professor of hepatology and gastroenterology, UCL Institute for Liver and Digestive Health, University College London.

Address for correspondence
Julia Hippisley-Cox, Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK.

Email: julia.hippisley-cox@phc.ox.ac.uk

Submitted: 4 March 2021; **Editor's response:** 13 April 2021; **final acceptance:** 2 July 2021.

How this fits in

This is the largest population-based study of its kind, systematically examining the symptomatology of pancreatic ductal adenocarcinoma (PDAC) and pancreatic neuroendocrine neoplasms (PNEN), and quantifying the association of 42 potential symptoms in different time windows relative to the date of diagnosis. This study confirmed several symptoms as risk factors for PDAC reported in previous UK studies, which had much smaller sample sizes. A deeper understanding of symptoms associated with PNEN is gained. Considering that most symptoms associated with pancreatic cancer are non-specific and do not qualify for urgent referral for investigation (2-week-wait) in the current NICE guideline, GPs should be vigilant of patients presenting with several concurrent non-specific symptoms and make proper safety-netting strategies. GPs should also increase the awareness of the risk of pancreatic cancer among people with comorbidities, and be careful not to attribute potential symptoms of pancreatic cancer to patients with existing health conditions.

for patients at high risk; watchful wait and safety netting for patients with low risk.

METHOD
Study design setting
This was a nested case-control study using the QResearch database (version 44), an extensive, validated, anonymised primary care database comprising records of >35 million patients registered in approximately 1500 GP surgeries, spread throughout the UK, which have been using the EMIS System since 1989. The population in the QResearch database is representative of the UK population. Patients' records have been linked with cancer registries, Office for National Statistics (ONS) mortality records, and Hospital Episode Statistics (HES). The cancer registration data include information on the date of diagnosis, type and location of the tumour, morphology, grade and stage, and treatment.

Study population
The eligible study population was an open cohort of patients aged ≥25 years and registered in the QResearch database between 1 January 2000 and 31 December 2019. Patients with an existing diagnosis of any type of pancreatic cancer before the entry date were excluded. The entry date to the cohort was the latest of the patient's 25th birthday, the date of patients registered with the practice plus 1 year, the date on which the practice computer system was installed plus 1 year, or the beginning of the study period. The right censor date was the earliest date of the following: the date of pancreatic cancer diagnosis, the date of death, the date of leaving the practice, or the study end date. Person years were calculated between the study entry date and the right censor date.

Identification of cases and controls
Cases were patients in the study cohort with an incident diagnosis of PDAC/PNEN, recorded in ≥1 of the four linked sources — GP records, HES, cancer registry, or ONS. The index date for cases was the earliest date the diagnosis was recorded in any four data sources. Cases were matched with up to 10 controls in the same practice, age, sex, and calendar year using incidence density sampling.[12] Each control was allocated an index date, which was the date of diagnosis of their matched case.

Candidate symptoms and potential risk factors
A broad list of the 42 symptoms potentially associated with PDAC and PNEN is summarised in Box 1. These symptoms were identified through literature review,[2,10,13,14] information from the leading charities such as Cancer Research UK[15] and Pancreatic Cancer UK,[16] National Institute for Health and Care Excellence (NICE) guidelines — NG12[17] and NG85,[18] and patient representatives. All occurrences of symptoms in primary care records were extracted, but the analysis was focused on the most recent 5 years before the index date. In addition, the following variables were of research interest and adjusted in the models: patients' sociodemographic characteristics (ethnicity, socioeconomic deprivation using Townsend quintile), lifestyle factors (smoking and drinking statuses, and body mass index [BMI], using most recently available values before the

---

Box 1. Overview of 42 candidate symptoms examined in study

- Symptoms in the QCancer (pancreas) model: abdominal distension (females only), abdominal pain, appetite loss, change in bowel habit, constipation (males only), dysphagia (males only), haematemesis, indigestion, and weight loss

- Symptoms from the literature: back pain, bloating, diarrhoea, fever/shivering, jaundice, itching/pruritus, nausea, steatorrhoea, and vomiting

- Other potential symptoms to be examined in association with the diagnosis of PDAC/PNEN: abdominal mass, gastrointestinal bleeding, rectal bleeding, heartburn/gastro-oesophageal reflux (GOR), food regurgitation, flatulence, faecal urgency, dark urine, dry mouth, cracked lips, thirst, loss of taste, night sweats, flushes, rash, paleness, oedema, tiredness, appetite increase, weight gain, bad breath, bruising, sore lips, and hypersomnia

Table 1. Clinicopathological characteristics of study cases with pancreatic cancers[a]

Characteristics	PDAC n (%) (N = 23 640)		PNEN n (%) (N = 596)	
Age at diagnosis, years				
25–29	11	(0.05)	<5	
30–39	129	(0.55)	25	(4.19)
40–49	666	(2.82)	84	(14.09)
50–59	2325	(9.84)	118	(19.80)
60–69	5176	(21.89)	177	(29.70)
70–79	7717	(32.64)	135	(22.65)
≥80	7616	(32.22)	54	(9.06)
Route to diagnosis				
Death certificate	85	(0.36)	<5	
Emergency presentation	5542	(23.44)	78	(13.09)
GP referral	2442	(10.33)	147	(24.66)
Inpatient elective	364	(1.54)	23	(3.86)
Other outpatient pathway	1125	(4.76)	112	(18.79)
2-week-wait	1887	(7.98)	57	(9.56)
Not recorded	12 195	(51.59)	178	(29.87)
Stage at diagnosis (TNM)				
Stage I	274	(1.16)	62	(10.40)
Stage II	769	(3.25)	53	(8.89)
Stage III	593	(2.51)	27	(4.53)
Stage IV	4016	(16.99)	161	(27.01)
Not recorded	17 988	(76.09)	293	(49.16)
Grade at diagnosis				
Well differentiated	303	(1.28)	230	(38.59)
Moderately differentiated	1683	(7.12)	58	(9.73)
Poorly differentiated	1702	(7.20)	61	(10.23)
Undifferentiated	49	(0.21)	<5	(NA)
Not recorded	19 903	(84.19)	246	(41.28)
Type of tumour				
Functional NEN	NA		16	(2.68)
Non-functional NEN	NA		271	(45.47)
Neuroendocrine carcinoma	NA		274	(45.97)
Mixed adenocarcinoma/NEN	NA		8	(1.34)
Other/NOS	NA		23	(3.86)
PDAC	23 640	(100)	NA	
Treatment				
Surgery				
No	17 026	(72.02)	220	(36.91)
Yes	6614	(27.98)	376	(63.09)
Chemotherapy				
No	19 402	(82.07)	430	(72.15)
Yes	4238	(17.93)	166	(27.85)
Radiotherapy				
No	23 000	(97.29)	557	(93.46)
Yes	640	(2.71)	39	(6.54)
Hormone therapy				
No	23 574	(99.72)	550	(92.28)
Yes	66	(0.28)	46	(7.72)
Other treatment				
No	19 030	(80.50)	458	(76.85)
Yes	4610	(19.50)	138	(23.15)

[a]Rows with counts <5 have been suppressed. NA = not applicable. NEN = neuroendocrine neoplasm. NOS = not otherwise specified. PDAC = pancreatic ductal adenocarcinoma. PNEN = pancreatic neuroendocrine neoplasm. TNM = TNM Classification of Malignant Tumours.

index date), and relevant comorbidities that could cause the symptoms, or be potential risk factors for, pancreatic cancer.

Statistical analysis

Descriptive statistics were used to summarise the sociodemographic and clinical characteristics of patients diagnosed with PDAC and PNEN, and the matched control group. The key clinical characteristics of PDAC and PNEN cases were compared.

Exploratory analyses were conducted to investigate the association between the most recently recorded symptom relative to the index date in seven different periods and patient groups (case/control) using univariable conditional logistic regression. These seven timeframes were <1 month, 1–3 months, 4–6 months, 7–12 months, 1–2 years, 2–3 years, and 3–5 years before the index date. The purpose of setting different timeframes for the same symptoms was to compare how the odds ratio (OR) would change, and the implication of timeframes for earlier diagnosis of PDAC and PNEN based on symptomatic presentation. Based on the exploratory results and the clinical relevance of timeframes for earlier diagnosis of pancreatic cancer, the seven timeframes were narrowed down to the following four in two sets of analysis:

- within 3 months for alarm symptoms, or 1 year for other symptoms (denoted as 3M/1Y); and
- within 6 months for alarm symptoms, or 2 years for other symptoms (denoted as 6M/2Y).

Alarm symptoms included jaundice, dysphagia, and gastrointestinal (GI) bleeding. These three symptoms were analysed within shorter timeframes (within 3 to 6 months before the index date), as they are widely accepted as 'red flag' symptoms and should be promptly investigated in primary care, or referred to secondary care. The timeframes were longer for the other symptoms (within 1 or 2 years before the index date), as they are non-specific, probably caused by other benign conditions, and not easily ascribed to an underlying tumour. A categorical variable was used to denote whether the patients presented to their GPs for each symptom based on the most recent date of presentation, for example, no record of presenting with jaundice (reference category), presenting with jaundice within 3 months, or >3 months before the index date. Symptoms in other

Table 2. Demographics, lifestyle, and comorbidities in case controls[a]

Characteristics	Cases with PDAC n (%) (N = 23 640)		Cases with PNEN n (%) (N = 596)		Controls n (%) (N = 230 024)	
Age, years						
Mean (SD)	73 (11.5)		62.3 (13.2)		72.0 (11.4)	
Median (IQR)	74 (66–82)		64 (53–72)		73 (65–81)	
Sex						
Female	11 705	(49.51)	294	(49.33)	114 429	(49.75)
Male	11 935	(50.49)	302	(50.67)	115 595	(50.25)
Ethnicity						
White	12 604	(53.32)	367	(61.58)	154 096	(66.99)
Indian	186	(0.79)	9	(1.51)	2674	(1.16)
Pakistani	100	(0.42)	6	(1.01)	1150	(0.50)
Bangladeshi	47	(0.20)	<5		692	(0.30)
Other Asian	79	(0.33)	10	(1.68)	1223	(0.53)
Caribbean	247	(1.04)	10	(1.68)	2364	(1.03)
Black African	103	(0.44)	<5	(NA)	1369	(0.60)
Chinese	36	(0.15)	<5	(NA)	449	(0.20)
Other	166	(0.70)	12	(2.01)	1853	(0.81)
Not recorded	10 072	(42.61)	173	(29.03)	64 154	(27.89)
Townsend quintile						
1 (most affluent)	7054	(29.84)	190	(31.88)	72 988	(31.73)
2	5873	(24.84)	135	(22.65)	58 901	(25.61)
3	4758	(20.13)	126	(21.14)	44 581	(19.38)
4	3442	(14.56)	91	(15.27)	31 513	(13.70)
5 (most deprived)	2470	(10.45)	54	(9.06)	21 758	(9.46)
Not recorded	43	(0.18)	0	(0.00)	283	(0.12)
Year of index date						
2000–2004	4358	(18.43)	68	(11.41)	42 402	(18.43)
2005–2009	5744	(24.30)	107	(17.95)	55 270	(24.03)
2010–2014	6767	(28.63)	225	(37.75)	66 009	(28.70)
2015–2019	6771	(28.64)	196	(32.89)	66 343	(28.84)
Smoking status						
Non-smoker	10 841	(45.86)	327	(54.87)	118 626	(51.57)
Ex-smoker	7117	(30.11)	170	(28.52)	76 842	(33.41)
Light smoker (<10/day)	3047	(12.89)	45	(7.55)	16 917	(7.35)
Moderate smoker (10–19)	617	(2.61)	8	(1.34)	2908	(1.26)
Heavy smoker (≥20)	394	(1.67)	7	(1.17)	1959	(0.85)
Not recorded	1624	(6.87)	39	(6.54)	12 772	(5.55)
Alcohol consumption						
Non-drinker	14 053	(59.45)	354	(59.40)	139 902	(60.82)
Trivial (<1 units/day)	3224	(13.64)	89	(14.93)	32 028	(13.92)
Light (1–2 units/day)	1480	(6.26)	41	(6.88)	15 194	(6.61)
Moderate (3–6 units/day)	1539	(6.51)	32	(5.37)	14 712	(6.40)
Heavy (7–9 units/day)	142	(0.60)	<5	(NA)	1173	(0.51)
Very heavy (>9 units/day)	67	(0.28)	<5	(NA)	549	(0.24)
Not recorded	3135	(13.26)	75	(12.58)	26 466	(11.51)
Comorbidities/previous medical history						
Type 1 diabetes	534	(2.26)	25	(4.19)	1752	(0.76)
Type 2 diabetes	6581	(27.84)	187	(31.38)	38 666	(16.81)
Venous thrombus embolism	2631	(11.13)	62	(10.40)	12 565	(5.46)
Deep vein thrombosis	1632	(6.90)	35	(5.87)	8246	(3.58)
Pulmonary embolism	1320	(5.58)	34	(5.70)	5640	(2.45)
Acute pancreatitis	857	(3.63)	21	(3.52)	2208	(0.96)
Chronic pancreatitis	477	(2.02)	6	(1.01)	409	(0.18)
Cholangitis	383	(1.62)	8	(1.34)	703	(0.31)
Gallstones	1602	(6.78)	44	(7.38)	12 378	(5.38)
H. pylori infection	996	(4.21)	37	(6.21)	8359	(3.63)
Family history of GI cancer	456	(1.93)	15	(2.52)	4074	(1.77)
Blood cancer	350	(1.48)	21	(3.52)	4123	(1.79)
Colon cancer	371	(1.57)	10	(1.68)	3829	(1.66)

[a]Rows with counts <5 have been suppressed. GI = gastrointestinal. IQR = interquartile range. PDAC = pancreatic ductal adenocarcinoma. PNEN = pancreatic neuroendocrine neoplasm. SD = standard deviation.

timeframes (6 months, 1 or 2 years) were operationalised in the same way.

Because of the large difference in sample sizes, PDAC and PNEN were analysed separately. Five variables contained missing data, including ethnicity, Townsend quintile, BMI, smoking and drinking statuses. Multiple imputation with chained equations was used to impute missing values for these variables under the missing at random assumption. Ten imputations were conducted. Multivariable conditional logistic regression models were used to identify symptoms significantly associated with the diagnoses of PDAC and PNENs, adjusting for patient characteristics and comorbidities, with Rubin's rules used to pool the parameter estimates across the 10 imputed datasets.[19] Possible interactions were considered and tested in the model. Odds ratios (OR) and 95% confidence intervals (CI) for each symptom were calculated and visualised in forest plots. Symptoms with an OR >1.2 at a significance level of $P<0.01$ for PDAC or PNEN were considered clinically and statistically relevant. Sensitivity analyses were conducted in patients (both cases and control) with at least 3 years of electronic health records (EHRs) before the index date. All statistical analyses were conducted in Stata (version 16.1). The reporting of this study followed the recommendations of the STROBE (strengthening the reporting of observational studies in epidemiology) statement.[20]

RESULTS

Population characteristics

The open cohort included 15 194 279 patients aged ≥25 years, with a total of 100 290 294 person years of follow-up. A total of 23 640 PDAC and 596 PNEN cases were identified from the cohort. Case ascertainment from the linked data sources is shown in Supplementary Table S1. The age-standardised incidence rate of PDAC was 23.21 (95% CI = 22.91 to 23.51) per 100 000 person years and 0.95 (95% CI = 0.89 to 1.01) per 100 000 person years for PNEN (data not shown). Table 1 summarises the demographical and clinical characteristics of the cases, organised by tumour type. The comparison of characteristics between cases and controls (N = 230 024) is shown in Table 2.

Symptoms associated with PDAC and PNEN

Figures 1a and 1b present the significant results of multivariable analyses for PDAC and PNEN for symptoms recorded within

Adjusted odds ratio (95% CI) for PDAC after multiple imputation (significant results only)
Symptom period: 3 months/1 year before the index date (*P*<0.01)

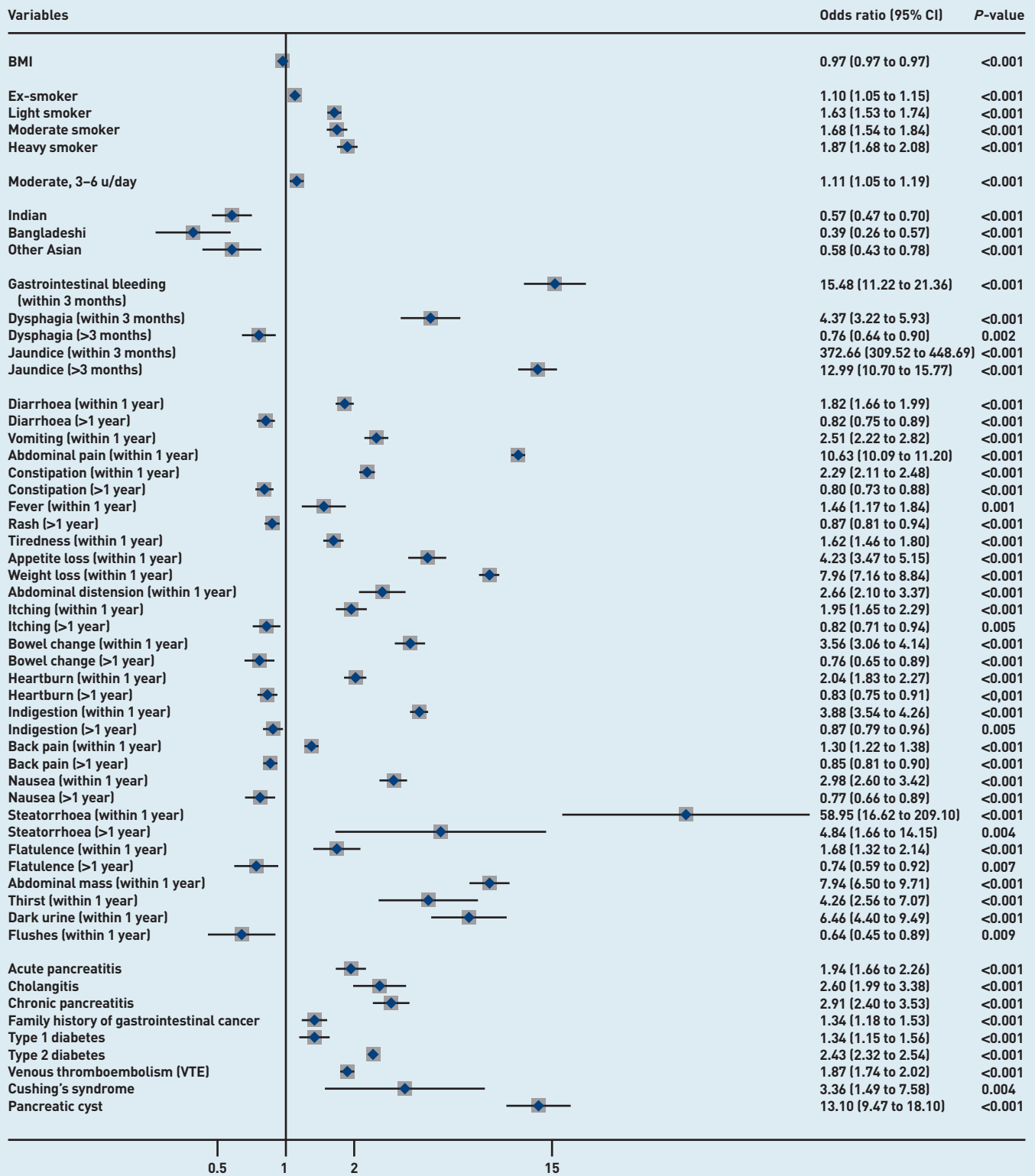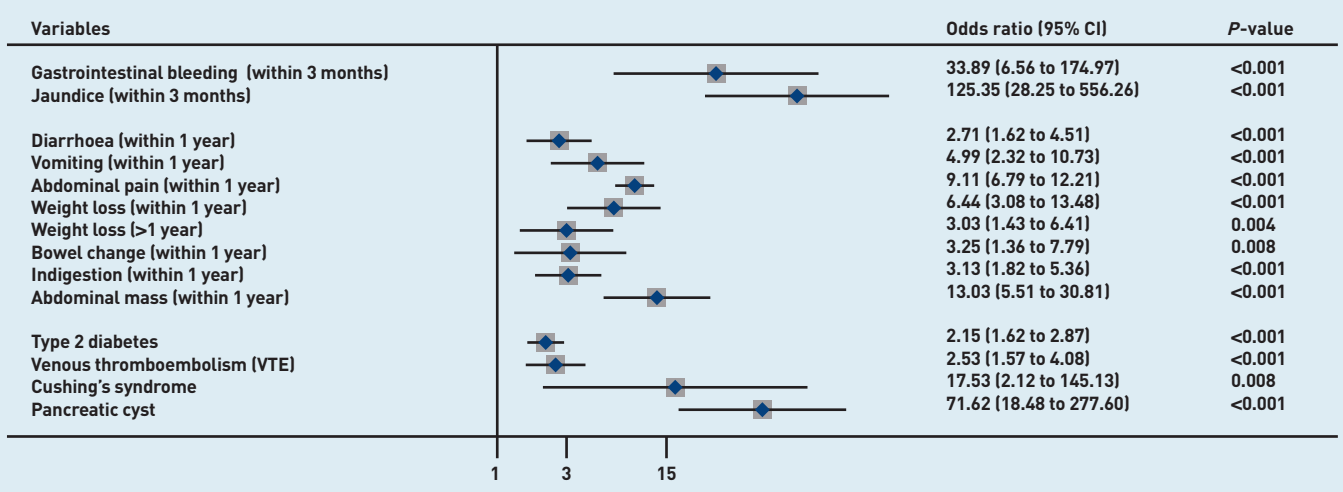| Variables	Odds ratio (95% CI)	*P*-value
BMI	0.97 (0.97 to 0.97)	<0.001
Ex-smoker	1.10 (1.05 to 1.15)	<0.001
Light smoker	1.63 (1.53 to 1.74)	<0.001
Moderate smoker	1.68 (1.54 to 1.84)	<0.001
Heavy smoker	1.87 (1.68 to 2.08)	<0.001
Moderate, 3–6 u/day	1.11 (1.05 to 1.19)	<0.001
Indian	0.57 (0.47 to 0.70)	<0.001
Bangladeshi	0.39 (0.26 to 0.57)	<0.001
Other Asian	0.58 (0.43 to 0.78)	<0.001
Gastrointestinal bleeding (within 3 months)	15.48 (11.22 to 21.36)	<0.001
Dysphagia (within 3 months)	4.37 (3.22 to 5.93)	<0.001
Dysphagia (>3 months)	0.76 (0.64 to 0.90)	0.002
Jaundice (within 3 months)	372.66 (309.52 to 448.69)	<0.001
Jaundice (>3 months)	12.99 (10.70 to 15.77)	<0.001
Diarrhoea (within 1 year)	1.82 (1.66 to 1.99)	<0.001
Diarrhoea (>1 year)	0.82 (0.75 to 0.89)	<0.001
Vomiting (within 1 year)	2.51 (2.22 to 2.82)	<0.001
Abdominal pain (within 1 year)	10.63 (10.09 to 11.20)	<0.001
Constipation (within 1 year)	2.29 (2.11 to 2.48)	<0.001
Constipation (>1 year)	0.80 (0.73 to 0.88)	<0.001
Fever (within 1 year)	1.46 (1.17 to 1.84)	0.001
Rash (>1 year)	0.87 (0.81 to 0.94)	<0.001
Tiredness (within 1 year)	1.62 (1.46 to 1.80)	<0.001
Appetite loss (within 1 year)	4.23 (3.47 to 5.15)	<0.001
Weight loss (within 1 year)	7.96 (7.16 to 8.84)	<0.001
Abdominal distension (within 1 year)	2.66 (2.10 to 3.37)	<0.001
Itching (within 1 year)	1.95 (1.65 to 2.29)	<0.001
Itching (>1 year)	0.82 (0.71 to 0.94)	0.005
Bowel change (within 1 year)	3.56 (3.06 to 4.14)	<0.001
Bowel change (>1 year)	0.76 (0.65 to 0.89)	<0.001
Heartburn (within 1 year)	2.04 (1.83 to 2.27)	<0.001
Heartburn (>1 year)	0.83 (0.75 to 0.91)	<0,001
Indigestion (within 1 year)	3.88 (3.54 to 4.26)	<0.001
Indigestion (>1 year)	0.87 (0.79 to 0.96)	0.005
Back pain (within 1 year)	1.30 (1.22 to 1.38)	<0.001
Back pain (>1 year)	0.85 (0.81 to 0.90)	<0.001
Nausea (within 1 year)	2.98 (2.60 to 3.42)	<0.001
Nausea (>1 year)	0.77 (0.66 to 0.89)	<0.001
Steatorrhoea (within 1 year)	58.95 (16.62 to 209.10)	<0.001
Steatorrhoea (>1 year)	4.84 (1.66 to 14.15)	0.004
Flatulence (within 1 year)	1.68 (1.32 to 2.14)	<0.001
Flatulence (>1 year)	0.74 (0.59 to 0.92)	0.007
Abdominal mass (within 1 year)	7.94 (6.50 to 9.71)	<0.001
Thirst (within 1 year)	4.26 (2.56 to 7.07)	<0.001
Dark urine (within 1 year)	6.46 (4.40 to 9.49)	<0.001
Flushes (within 1 year)	0.64 (0.45 to 0.89)	0.009
Acute pancreatitis	1.94 (1.66 to 2.26)	<0.001
Cholangitis	2.60 (1.99 to 3.38)	<0.001
Chronic pancreatitis	2.91 (2.40 to 3.53)	<0.001
Family history of gastrointestinal cancer	1.34 (1.18 to 1.53)	<0.001
Type 1 diabetes	1.34 (1.15 to 1.56)	<0.001
Type 2 diabetes	2.43 (2.32 to 2.54)	<0.001
Venous thromboembolism (VTE)	1.87 (1.74 to 2.02)	<0.001
Cushing's syndrome	3.36 (1.49 to 7.58)	0.004
Pancreatic cyst	13.10 (9.47 to 18.10)	<0.001

Figure 1a. Forest plots for multivariable conditional logistic regression after multiple imputation (alarm symptoms 3 months before the index date, other symptoms 1 year before the index date) for PDAC. Because of a large number of variables in the model, only significant results (adjusted odds ratio and 95% CI) are presented in the figures. Full results of the models are available in Supplementary Tables S2–S3. The cut-off period for jaundice, gastrointestinal bleeding, and dysphagia was 3 months; 1 year for other symptoms. BMI = body mass index. CI = confidence interval. PDAC = pancreatic ductal adenocarcinoma. u = units.

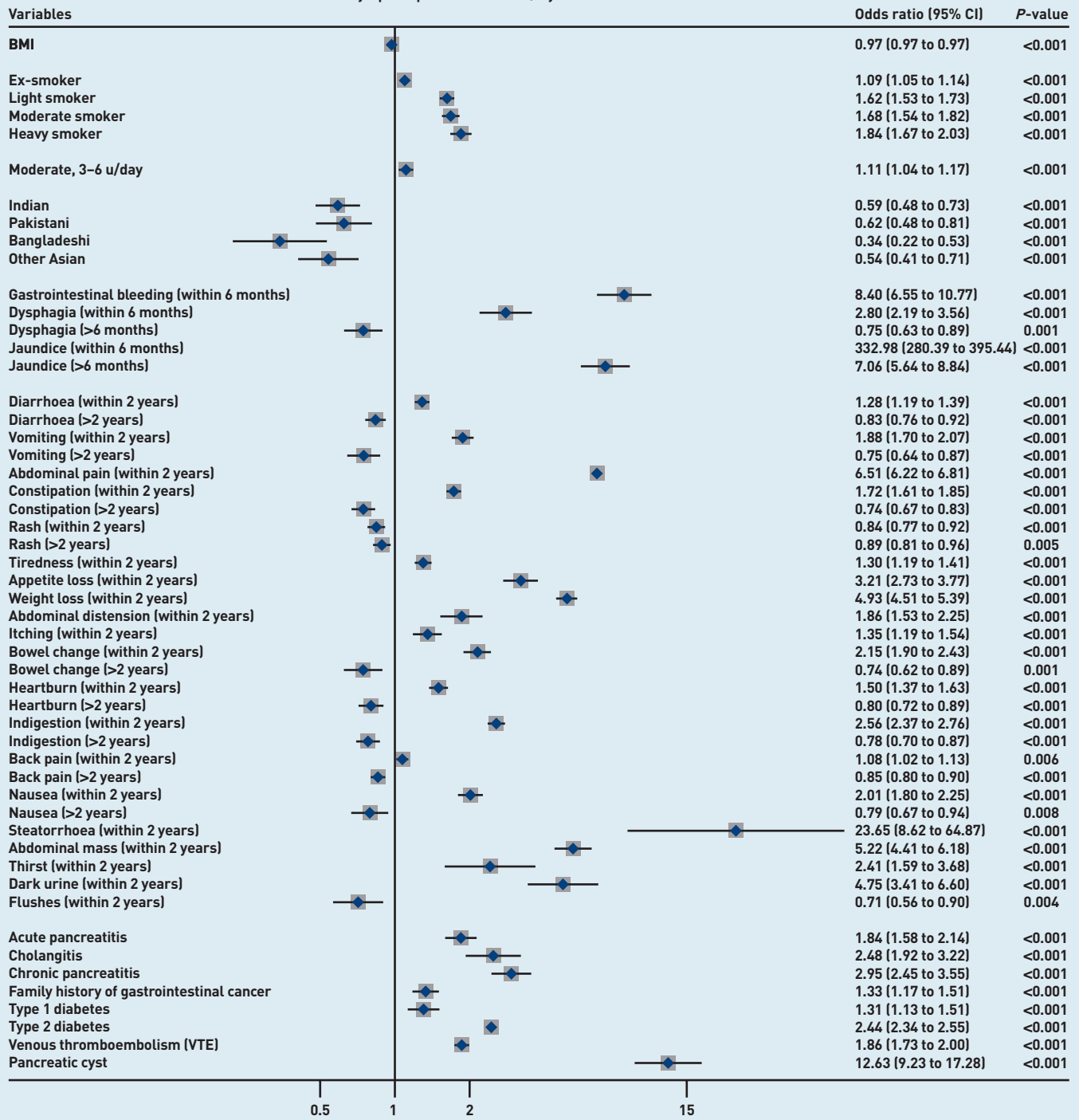| Adjusted odds ratio (95% CI) for PNEN after multiple imputation (significant results only) Symptom period: 3 months/1 year before the index date (*P*<0.01)		
Variables	**Odds ratio (95% CI)**	***P*-value**
Gastrointestinal bleeding (within 3 months)	33.89 (6.56 to 174.97)	<0.001
Jaundice (within 3 months)	125.35 (28.25 to 556.26)	<0.001
Diarrhoea (within 1 year)	2.71 (1.62 to 4.51)	<0.001
Vomiting (within 1 year)	4.99 (2.32 to 10.73)	<0.001
Abdominal pain (within 1 year)	9.11 (6.79 to 12.21)	<0.001
Weight loss (within 1 year)	6.44 (3.08 to 13.48)	<0.001
Weight loss (>1 year)	3.03 (1.43 to 6.41)	0.004
Bowel change (within 1 year)	3.25 (1.36 to 7.79)	0.008
Indigestion (within 1 year)	3.13 (1.82 to 5.36)	<0.001
Abdominal mass (within 1 year)	13.03 (5.51 to 30.81)	<0.001
Type 2 diabetes	2.15 (1.62 to 2.87)	<0.001
Venous thromboembolism (VTE)	2.53 (1.57 to 4.08)	<0.001
Cushing's syndrome	17.53 (2.12 to 145.13)	0.008
Pancreatic cyst	71.62 (18.48 to 277.60)	<0.001

1   3   15

Figure 1b. Forest plots for multivariable conditional logistic regression after multiple imputation (alarm symptoms 3 months before the index date, other symptoms 1 year before the index date) for PNEN. Because of a large number of variables in the model, only significant results (adjusted odds ratio and 95% CI) are presented in the figures. Full results of the models are available in Supplementary Tables S2–S3. The cut-off period for jaundice, gastrointestinal bleeding, and dysphagia was 3 months; 1 year for other symptoms. BMI = body mass index. CI = confidence interval. PNEN = pancreatic neuroendocrine neoplasm.

3 months (alarm symptoms) or 1 year (non-specific) before the index date, respectively, adjusting for patient characteristics and comorbidities. The significant results for symptoms in the 6M/2Y timeframe for PDAC and PNEN are presented in Figures 2a and 2b (only significant results have been plotted, but full results are available in Supplementary Tables S2–S5). Most results were congruent in the two models, though symptoms in a shorter timeframe (3M/1Y) generally had higher ORs than those in a longer period (6M/2Y) and had wider confidence intervals. In addition, symptoms within the cut-off periods (for example, 3 to 6 months or 1 to 2 years) were statistically significant. For symptoms longer than the cut-off periods, there were two main patterns: either the symptoms became non-significant, or the direction of OR in the symptoms reversed (from >1 to <1, significantly higher odds in control), which meant the controls were more likely to consult with those (non-specific) symptoms after the cut-off periods.

Noticeably, the effect of unintentional weight loss may be longer than 2 years before the diagnosis of PNEN. Interaction terms were tested; however, they were not in the final model because including interaction terms increased the number of parameters that need to be estimated. The model was unable to converge, especially for PNEN with a small sample size.

Jaundice had the highest adjusted OR in both PDAC and PNEN. Symptoms associated with PNEN were a subset of symptoms associated with PDAC, but the strength of ORs among the significant symptoms may not be the same in PDAC and PNEN. Nine symptoms were significantly associated with the diagnosis of PNEN in the timeframe

of 3M/1Y, including jaundice, GI bleeding, diarrhoea, bowel change, vomiting, indigestion, abdominal mass, abdominal pain, and weight loss. The additional significant symptoms associated with PDAC included constipation, steatorrhea, abdominal distension, nausea, flatulence, heartburn, fever, tiredness, appetite loss, itching, back pain, thirst, and dark urine (Figure 1a).

Other risk factors
Compared with people of white ethnicity, people of Indian, Bangladeshi, and other Asian (not including Chinese) ethnicity were less likely to develop PDAC (OR<1). Smoking and drinking were risk factors for PDAC. When considering comorbidities, type 2 diabetes mellitus (T2DM), venous thromboembolism, Cushing's syndrome, and presence of pancreatic cysts significantly increased the risks of PDAC and PNEN. Acute pancreatitis, cholangitis, a family history of GI cancer, and type 1 diabetes were significant risk factors for PDAC, but not for PNEN (Figures 1 and 2, comparing a and b).

Sensitivity analysis
About 26% (25.9%) of patients (65 884 out of 254 260, including cases and controls) were excluded from the sensitivity analysis. By comparing the results from the main analysis, the conclusion of symptoms in the sensitivity analysis did not change, though there were some changes in OR. Complete results of sensitivity analyses are available in Supplementary Tables S6–S9.

DISCUSSION
Summary
A total of 24 236 patients diagnosed with pancreatic cancer during 2000 to 2019 were

Adjusted odds ratio (95% CI) for PDAC after multiple imputation (significant results only)
Symptom period: 6 months/2 years before the index date

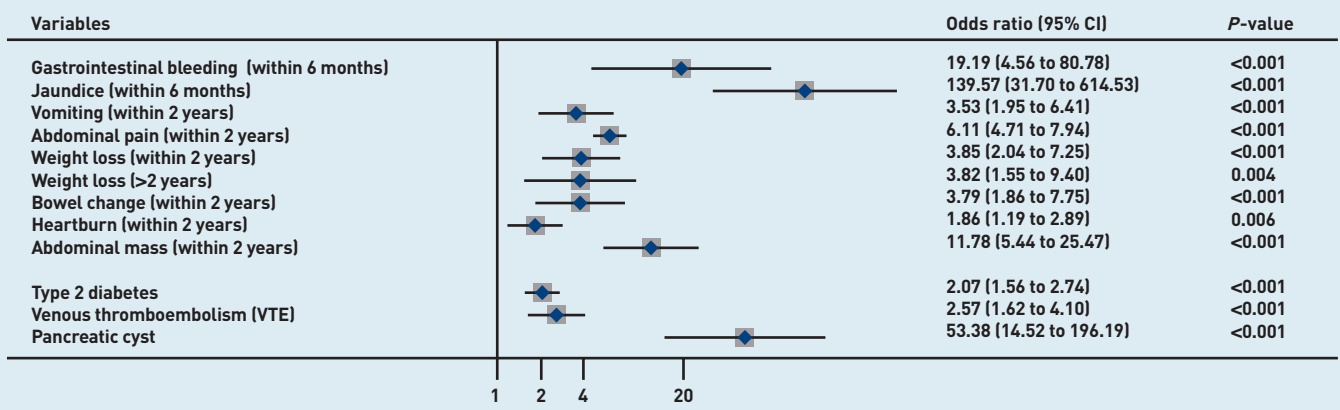| Variables	Odds ratio (95% CI)	P-value
BMI	0.97 (0.97 to 0.97)	<0.001
Ex-smoker	1.09 (1.05 to 1.14)	<0.001
Light smoker	1.62 (1.53 to 1.73)	<0.001
Moderate smoker	1.68 (1.54 to 1.82)	<0.001
Heavy smoker	1.84 (1.67 to 2.03)	<0.001
Moderate, 3–6 u/day	1.11 (1.04 to 1.17)	<0.001
Indian	0.59 (0.48 to 0.73)	<0.001
Pakistani	0.62 (0.48 to 0.81)	<0.001
Bangladeshi	0.34 (0.22 to 0.53)	<0.001
Other Asian	0.54 (0.41 to 0.71)	<0.001
Gastrointestinal bleeding (within 6 months)	8.40 (6.55 to 10.77)	<0.001
Dysphagia (within 6 months)	2.80 (2.19 to 3.56)	<0.001
Dysphagia (>6 months)	0.75 (0.63 to 0.89)	0.001
Jaundice (within 6 months)	332.98 (280.39 to 395.44)	<0.001
Jaundice (>6 months)	7.06 (5.64 to 8.84)	<0.001
Diarrhoea (within 2 years)	1.28 (1.19 to 1.39)	<0.001
Diarrhoea (>2 years)	0.83 (0.76 to 0.92)	<0.001
Vomiting (within 2 years)	1.88 (1.70 to 2.07)	<0.001
Vomiting (>2 years)	0.75 (0.64 to 0.87)	<0.001
Abdominal pain (within 2 years)	6.51 (6.22 to 6.81)	<0.001
Constipation (within 2 years)	1.72 (1.61 to 1.85)	<0.001
Constipation (>2 years)	0.74 (0.67 to 0.83)	<0.001
Rash (within 2 years)	0.84 (0.77 to 0.92)	<0.001
Rash (>2 years)	0.89 (0.81 to 0.96)	0.005
Tiredness (within 2 years)	1.30 (1.19 to 1.41)	<0.001
Appetite loss (within 2 years)	3.21 (2.73 to 3.77)	<0.001
Weight loss (within 2 years)	4.93 (4.51 to 5.39)	<0.001
Abdominal distension (within 2 years)	1.86 (1.53 to 2.25)	<0.001
Itching (within 2 years)	1.35 (1.19 to 1.54)	<0.001
Bowel change (within 2 years)	2.15 (1.90 to 2.43)	<0.001
Bowel change (>2 years)	0.74 (0.62 to 0.89)	0.001
Heartburn (within 2 years)	1.50 (1.37 to 1.63)	<0.001
Heartburn (>2 years)	0.80 (0.72 to 0.89)	<0.001
Indigestion (within 2 years)	2.56 (2.37 to 2.76)	<0.001
Indigestion (>2 years)	0.78 (0.70 to 0.87)	<0.001
Back pain (within 2 years)	1.08 (1.02 to 1.13)	0.006
Back pain (>2 years)	0.85 (0.80 to 0.90)	<0.001
Nausea (within 2 years)	2.01 (1.80 to 2.25)	<0.001
Nausea (>2 years)	0.79 (0.67 to 0.94)	0.008
Steatorrhoea (within 2 years)	23.65 (8.62 to 64.87)	<0.001
Abdominal mass (within 2 years)	5.22 (4.41 to 6.18)	<0.001
Thirst (within 2 years)	2.41 (1.59 to 3.68)	<0.001
Dark urine (within 2 years)	4.75 (3.41 to 6.60)	<0.001
Flushes (within 2 years)	0.71 (0.56 to 0.90)	0.004
Acute pancreatitis	1.84 (1.58 to 2.14)	<0.001
Cholangitis	2.48 (1.92 to 3.22)	<0.001
Chronic pancreatitis	2.95 (2.45 to 3.55)	<0.001
Family history of gastrointestinal cancer	1.33 (1.17 to 1.51)	<0.001
Type 1 diabetes	1.31 (1.13 to 1.51)	<0.001
Type 2 diabetes	2.44 (2.34 to 2.55)	<0.001
Venous thromboembolism (VTE)	1.86 (1.73 to 2.00)	<0.001
Pancreatic cyst	12.63 (9.23 to 17.28)	<0.001

Figure 2a. Forest plots for multivariable conditional logistic regression after multiple imputation (alarm symptoms 6 months before the index date, other symptoms 2 years before the index date) for PDAC. Because of a large number of variables in the model, only significant results (adjusted odds ratio and 95% CI) were plotted in the figures. Full results of the models are available in Supplementary Tables S4–S5. The cut-off period for jaundice, gastrointestinal bleeding, and dysphagia was 6 months; 2 years for other symptoms. BMI = body mass index. CI = confidence interval. PDAC = pancreatic ductal adenocarcinoma. u = units.

identified from the QResearch database. Nine symptoms were significantly associated with the diagnosis of PNEN, which is a subset of 23 significant symptoms for PDAC. A shorter timeframe (3M/1Y) was considered better than a longer one (6M/2Y) for earlier cancer diagnosis, as cases had higher odds of presenting symptoms in the 3M/1Y timeframe. Jaundice had the highest adjusted OR in both PDAC and PNEN. Thirst and dark urine were the two newly identified symptoms associated with PDAC, not previously reported in other studies. Thirst could be a symptom explained by T2DM, which is associated with pancreatic cancer. Dark urine could be caused by progressing liver dysfunction, or the manifestation of biliary duct obstruction.

Adjusted odds ratio (95% CI) for PNEN after multiple imputation (significant results only)
Symptom period: 6 months/2 years before the index date (*P*<0.01)

Variables	Odds ratio (95% CI)	*P*-value
Gastrointestinal bleeding (within 6 months)	19.19 (4.56 to 80.78)	<0.001
Jaundice (within 6 months)	139.57 (31.70 to 614.53)	<0.001
Vomiting (within 2 years)	3.53 (1.95 to 6.41)	<0.001
Abdominal pain (within 2 years)	6.11 (4.71 to 7.94)	<0.001
Weight loss (within 2 years)	3.85 (2.04 to 7.25)	<0.001
Weight loss (>2 years)	3.82 (1.55 to 9.40)	0.004
Bowel change (within 2 years)	3.79 (1.86 to 7.75)	<0.001
Heartburn (within 2 years)	1.86 (1.19 to 2.89)	0.006
Abdominal mass (within 2 years)	11.78 (5.44 to 25.47)	<0.001
Type 2 diabetes	2.07 (1.56 to 2.74)	<0.001
Venous thromboembolism (VTE)	2.57 (1.62 to 4.10)	<0.001
Pancreatic cyst	53.38 (14.52 to 196.19)	<0.001

Figure 2b. Forest plot for multivariable conditional logistic regression after multiple imputation (alarm symptoms 6 months before the index date, other symptoms 2 years before the index date) for PNEN. Because of a large number of variables in the model, only significant results (adjusted odds ratio and 95% CI) were plotted in the figures. Full results of the models are available in Supplementary Tables S4–S5. The cut-off period for jaundice, gastrointestinal bleeding, and dysphagia was 6 months; 2 years for other symptoms. BMI = body mass index. CI = confidence interval. PDAC = pancreatic ductal adenocarcinoma. PNEN = pancreatic neuroendocrine neoplasm.

The complete findings of this study are summarised in Box 2.

Early diagnosis of pancreatic cancer from primary care is still challenging, owing to non-specific symptoms. This study identified 23 symptoms associated with the diagnosis of PDAC and nine symptoms for PNEN. Risk prediction models incorporating comprehensive symptomatology would help identify patients with a high risk of developing pancreatic tumours from primary care. Patients could benefit from an earlier cancer diagnosis and better survival outcomes, which can also save costs for the NHS.

Strengths and limitations
The QResearch database provided rich data for this study, which is by far the largest study of its kind. The representative patient population makes the study findings more generalisable to a broader UK population. The use of EHRs avoided selection, recall, and responder biases from the survey, and also provided benefits from the accuracy of coding and data completeness in the UK general practice. The authors explored the effects of symptoms in seven timeframes first, and then narrowed down to 3M/1Y and 6M/2Y before the index date. Symptoms recorded longer than the timeframes (>3M/1Y, >6M/2Y) were not mixed with no symptom recorded, which provided new information about the symptoms beyond the cut-off periods in cases and controls. Though symptoms were the focus of this study, the background risk factors and relevant comorbidities were taken into account, included, and adjusted for in the model, which is another strength. The authors conducted the study as transparently and thoroughly as possible. The research protocol has been published on the QResearch website. The reporting of this article complies with the STROBE statement.

Information bias in EHRs is the first limitation. The authors could not evaluate how accurately the information was recorded across practices. The recording habit may have considerably differed among GPs. The heterogeneity of recording habits was mitigated by using all possible Read codes for each variable. Because of the small sample size in PNEN cases, it is possible that the full burden of symptoms in

Box 2. Summary of patient and clinical characteristics significantly associated with the diagnosis of PDAC and PNEN

	PDAC	PNEN
Demographics	Reduced risks in Indian, Bangladeshi, other Asian (not including Chinese) compared with white ethnicity (OR<1)	
Lifestyle	BMI (OR<1)	
	All smoking categories associated with increased risks compared with non-smokers	
	Moderate drinking (3–6 units/day)	
Symptoms (within 3 months)	GI bleeding, jaundice, dysphagia	GI bleeding, jaundice
Symptoms (within 1 year)	Diarrhoea, bowel change, vomiting, indigestion, abdominal mass, abdominal pain, weight loss[7]	
	Constipation, steatorrhea, abdominal distension, nausea, flatulence, heartburn, fever, tiredness, appetite loss, itching, back pain, thirst, dark urine	Diarrhoea, bowel change, vomiting, indigestion, abdominal mass, abdominal pain, weight loss (longer than 2 years)
Comorbidities	Type 2 diabetes, venous thromboembolism (VTE), Cushing's syndrome, pancreatic cyst, acute pancreatitis, cholangitis, family history of GI cancer, type 1 diabetes	Type 2 diabetes, VTE, Cushing's syndrome, pancreatic cyst

BMI = body mass index. GI = gastrointestinal. OR = odds ratio. PDAC = pancreatic ductal adenocarcinoma. PNEN = pancreatic neuroendocrine neoplasm.

Discuss this article

Contribute and read comments about this article: **bjgp.org/letters**

PNEN could not be captured. Therefore, the authors could not discern whether a lower number of significant symptoms for PNEN is a lack of statistical power, or PNEN has truly less prominent symptomatology, or both. The researchers planned to explore whether there was any symptom associated with early/late stages at diagnosis in PDAC and PNEN. Unfortunately, the large amount of missing data in cancer staging (Supplementary Table S10) did not allow them to conduct such an analysis.

Comparison with existing literature

The current study has some improvements from the authors' previous QCancer (pancreas) prediction model,[10] including a longer study period, a larger sample size of incident cases, more symptoms examined, and the exploration of PNEN, which resulted in an additional 13 significant symptoms identified. Some UK studies examined the symptomatology of PDAC in primary care settings, with a similar study design (matched case-control study) and statistical method (conditional logistic regression), using the GPRD[21] and the Health Improvement Network (THIN)[22] databases. The findings in this study are generally consistent with these two publications. No studies have systematically and robustly evaluated the symptomatology of PNEN in primary care. Patients with PNEN ($n = 64$) reported their symptoms in a voluntary, internet-based survey.

Given the small sample size and potential recall bias in that study,[23] the authors believe that the present population-based approach offers a more robust and generalisable insight of PNEN symptomatology. Older age, smoking, excess alcohol intake, chronic pancreatitis, and T2DM are common risk factors for pancreatic cancer.[24–28] The findings are the same in the present study.

Implications for research and practice

Most symptoms identified in this study do not qualify for a rapid referral in the current NICE guideline for suspected (pancreatic) cancer pathway referral (NG12).[17] GPs should be vigilant to patients presenting with alarm symptoms and non-specific but concerning symptoms, especially when patients have existing comorbidities. Public and patient engagement events could raise public awareness of the symptoms of pancreatic cancer, which may help patients see their GPs more promptly when noticing bodily changes.

Based on the study findings, the authors can update the QCancer (pancreas) prediction model[10] and develop a new model for PNEN. It is also possible to quantify the risk of patients presenting with several concurrent non-specific symptoms, and the predictive values of such symptom combinations. It would be interesting to further understand how GPs managed and investigated patients presenting with different symptom combinations, and the association with the route to diagnosis and cancer stage.

In addition, there is an ongoing project in the authors' team, investigating diabetes as a risk pathway towards pancreatic cancer: *Diabetes as a risk pathway towards early diagnosis and prognostication of pancreatic cancer* (www.qresearch.org/research/approved-research-programs-and-projects/diabetes-as-a-risk-pathway-towards-early-diagnosis-and-prognostication-of-pancreatic-cancer/).

REFERENCES

1. Cancer Research UK. Pancreatic cancer statistics. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/pancreatic-cancer (accessed 17 Aug 2021).

2. Kamisawa T, Wood LD, Itoi T, Takaori K. Pancreatic cancer. *Lancet* 2016; **388(10039):** 73–85.

3. Clift AK, Kidd M, Bodei L, *et al.* Neuroendocrine neoplasms of the small bowel and pancreas. *Neuroendocrinology* 2020; **110(6):** 444–476.

4. Kleeff J, Korc M, Apte M, *et al.* Pancreatic cancer. *Nat Rev Dis Primers* 2016; **2:** 16022.

5. Pavel M, O'Toole D, Costa F, *et al.* ENETS Consensus guidelines update for the management of distant metastatic disease of intestinal, pancreatic, bronchial neuroendocrine neoplasms (NEN) and NEN of unknown primary site. *Neuroendocrinology* 2016; **103(2):** 172–185.

6. Frilling A, Clift AK. Therapeutic strategies for neuroendocrine liver metastases. *Cancer* 2015; **121(8):** 1172–1186.

7. Rawla P, Sunkara T, Gaduputi V. Epidemiology of pancreatic cancer: global trends, etiology and risk factors. *World J Oncol* 2019; **10(1):** 10–27.

8. Gedge K. Pancreatic cancer: a symptomless killer. *J Perioper Pract* 2017; **27(7–8):** 158–161.

9. Apollos JR, Sami S, Prasanth MN, *et al.* Pre-diagnostic delays caused by gastrointestinal investigations do not affect outcomes in pancreatic cancer. *Ann Med Surg* (Lond) 2018; **34:** 66–70.

10. Hippisley-Cox J, Coupland C. Identifying patients with suspected pancreatic cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* 2012; DOI: https://doi.org/10.3399/bjgp12X616355.

11. Collins GS, Altman DG. Identifying patients with undetected pancreatic cancer in primary care: an independent and external validation of QCancer (Pancreas). *Br J Gen Pract* 2013; DOI: https://doi.org/10.3399/bjgp13X671623.

12. Richardson DB. An incidence density sampling program for nested case-control analyses. *Occup Environ Med* 2004; **61(12):** e59.

13. Hippisley-Cox J, Coupland C. Symptoms and risk factors to identify men with suspected cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* 2013; DOI: https://doi.org/10.3399/bjgp20X714137.

14. Hippisley-Cox J, Coupland C. Symptoms and risk factors to identify women with suspected cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* 2013; DOI: https://doi.org/10.3399/bjgp13X660733.

15. Cancer Research UK. Pancreatic cancer symptoms. https://www.cancerresearchuk.org/about-cancer/pancreatic-cancer/symptoms (accessed 11 Jul 2021).

16. Pancreatic Cancer UK. Signs and symptoms of pancreatic cancer. https://www.pancreaticcancer.org.uk/information/signs-and-symptoms-of-pancreatic-cancer/ (accessed 11 Jul 2021).

17. National Institute for Health and Care Excellence. *Suspected cancer: recognition and referral. NG12.* London: NICE, 2015. https://www.nice.org.uk/guidance/ng12 (accessed 11 Jul 2021).

18. National Institute for Health and Care Excellence. *Pancreatic cancer in adults: diagnosis and management. NG85.* London: NICE, 2018. https://www.nice.org.uk/guidance/ng85 (accessed 11 Jul 2021).

19. Rubin DB. *Multiple imputation for non-response in surveys.* New York, NY: John Wiley & Sons, 1987.

20. von Elm E, Altman DG, Egger M, *et al.* The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007; **370(9596):** 1453–1457.

21. Stapley S, Peters TJ, Neal RD, *et al.* The risk of pancreatic cancer in symptomatic patients in primary care: a large case-control study using electronic records. *Br J Cancer* 2012; **106(12):** 1940–1944.

22. Keane MG, Horsfall L, Rait G, Pereira SP. A case-control study comparing the incidence of early symptoms in pancreatic and biliary tract cancer. *BMJ Open* 2014; **4(11):** e005720.

23. Basuroy R, Bouvier C, Ramage JK, *et al.* Delays and routes to diagnosis of neuroendocrine tumours. *BMC Cancer* 2018; **18(1):** 1122.

24. Maisonneuve P, Lowenfels AB. Risk factors for pancreatic cancer: a summary review of meta-analytical studies. *Int J Epidemiol* 2015; **44(1):** 186–198.

25. Bosetti C, Lucenteforte E, Silverman DT, *et al.* Cigarette smoking and pancreatic cancer: an analysis from the International Pancreatic Cancer Case-Control Consortium (Panc4). *Ann Oncol* 2012; **23(7):** 1880–1888.

26. Lucenteforte E, La Vecchia C, Silverman D, *et al.* Alcohol consumption and pancreatic cancer: a pooled analysis in the International Pancreatic Cancer Case-Control Consortium (PanC4). *Ann Oncol* 2012; **23(2):** 374–382.

27. Duell EJ, Lucenteforte E, Olson SH, *et al.* Pancreatitis and pancreatic cancer risk: a pooled analysis in the International Pancreatic Cancer Case-Control Consortium (PanC4). *Ann Oncol* 2012; **23(11):** 2964–2970.

28. Batabyal P, Vander Hoorn S, Christophi C, Nikfarjam M. Association of diabetes mellitus and pancreatic adenocarcinoma: a meta-analysis of 88 studies. *Ann Surg Oncol* 2014; **21(7):** 2453–2462.