Human oral microbiome and prospective risk for pancreatic cancer: a population-based nested case-control study

Xiaozhou Fan,[1] Alexander V Alekseyenko,[2] Jing Wu,[1] Brandilyn A Peters,[1] Eric J Jacobs,[3] Susan M Gapstur,[3] Mark P Purdue,[4] Christian C Abnet,[4] Rachael Stolzenberg-Solomon,[4] George Miller,[5,6,7] Jacques Ravel,[8] Richard B Hayes,[1,7] Jiyoung Ahn[1,7]

Correspondence to
Dr Jiyoung Ahn, Division of Epidemiology, Department of Population Health, New York University School of Medicine, 650 First Ave, New York, NY 10016, USA; Jiyoung.Ahn@nyumc.org

ABSTRACT

Objective A history of periodontal disease and the presence of circulating antibodies to selected oral pathogens have been associated with increased risk of pancreatic cancer; however, direct relationships of oral microbes with pancreatic cancer have not been evaluated in prospective studies. We examine the relationship of oral microbiota with subsequent risk of pancreatic cancer in a large nested case–control study.

Design We selected 361 incident adenocarcinoma of pancreas and 371 matched controls from two prospective cohort studies, the American Cancer Society Cancer Prevention Study II and the National Cancer Institute Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. From pre-diagnostic oral wash samples, we characterised the composition of the oral microbiota using bacterial 16S ribosomal RNA (16S rRNA) gene sequencing. The associations between oral microbiota and risk of pancreatic cancer, controlling for the random effect of cohorts and other covariates, were examined using traditional and L1-penalised least absolute shrinkage and selection operator logistic regression.

Results Carriage of oral pathogens, *Porphyromonas gingivalis* and *Aggregatibacter actinomycetemcomitans*, were associated with higher risk of pancreatic cancer (adjusted OR for presence vs absence=1.60 and 95% CI 1.15 to 2.22; OR=2.20 and 95% CI 1.16 to 4.18, respectively). Phylum *Fusobacteria* and its genus *Leptotrichia* were associated with decreased pancreatic cancer risk (OR per per cent increase of relative abundance=0.94 and 95% CI 0.89 to 0.99; OR=0.87 and 95% CI 0.79 to 0.95, respectively). Risks related to these phylotypes remained after exclusion of cases that developed within 2 years of sample collection, reducing the likelihood of reverse causation in this prospective study.

Conclusions This study provides supportive evidence that oral microbiota may play a role in the aetiology of pancreatic cancer.

Significance of this study

What is already known on this subject?
► Pancreatic cancer is highly lethal; 93% of patients die within 5 years of diagnosis.
► Poor oral health and history of periodontal diseases are associated with higher risk of pancreatic cancer.
► Serum antibody levels to selected periodontal pathogens are associated with increased risk for pancreatic cancer.

What are the new findings?
► In the direct assessment of genomic-based microbiome in oral samples, carriage of the oral pathogens *Porphyromonas gingivalis* and *Aggregatibacter actinomycetemcomitans* was associated with increased risk of pancreatic cancer.
► *Fusobacteria* and its genus *Leptotrichia* were associated with decreased risk of pancreatic cancer.
► Risks related to these phylotypes remained after exclusion of cases that developed within 2 years of sample collection, reducing the likelihood of reverse causation in this prospective study.

How might it impact on clinical practice in the foreseeable future?
► The identification of specific oral bacteria related to pancreatic cancer advances scientific knowledge on the aetiology of pancreatic cancer.
► These oral bacteria may additionally serve as readily accessible, non-invasive biomarkers for subsequent pancreatic cancer risk, which helps to identify people at high risk for this disease.
► Targeted prophylactic therapies may be developed to combat periodontal pathogens and decrease risk for pancreatic cancer.

INTRODUCTION

Pancreatic cancer is highly lethal; 93% of patients die within 5 years of diagnosis, and it is the fourth leading cause of cancer death in the USA.[1] Little is known about opportunities to prevent this cancer. In addition to inherited genetic factors, pancreatitis, smoking and excess body weight are risk factors, yet these only partially explain risk for this disease.[2–5] To reduce the public health burden,

there is a critical need to improve scientific knowledge on the causes of pancreatic cancer and to provide guidance for preventive measures.

More than 700 different bacterial species colonise the human oral cavity, known collectively as the oral microbiome.[6] Emerging evidence shows that oral microbiota play important roles in human health, including in immune response, carcinogen metabolism and nutrient digestion.[7–9] A series of recent epidemiological studies have showed that poor oral health status[10–14] and blood antibodies to selected oral pathogens[15] are associated with increased risk of pancreatic cancer. Because oral health is closely tied to oral microbial status,[16] we hypothesise that the oral health—pancreatic cancer association may have an underlying microbial basis, yet no studies have directly evaluated the relationship between the oral microbiome and subsequent risk for pancreatic cancer.

We conducted a prospective study nested in two large US cohorts, the American Cancer Society Cancer Prevention Study (CPS) II and the National Cancer Institute Prostate, Lung, Colorectal and Ovarian Cancer (PLCO) Screening Trial, to determine if oral microbiome was associated with subsequent risk of pancreatic cancer. We directly assessed the oral microbiota from high-throughput sequencing of the 16S Ribosomal RNA (16S rRNA) gene in pre-diagnostic oral samples from 361 pancreatic cancer cases and 371 controls in these cohorts, and compared these groups for baseline overall microbiota composition, carriage of known oral pathogens, and relative abundance of other specific bacterial taxa.

METHODS
Study population
Parent cohorts
The CPS II and PLCO cohorts comprise large US research populations with stored oral wash samples, comprehensive demographic information, and prospective follow-up for cancer incidence. The characteristics of the two cohorts are comparable, and oral wash samples were collected in a similar fashion in both cohorts. Study design and identification of incident cancers are described in detail in supplementary methods (see online supplementary methods).

Briefly, the CPS II cohort[17] includes more than 184 000 participants, aged 50–74, from 21 US states who completed a mailed baseline diet and lifestyle questionnaire in 1992. Follow-up questionnaires have been sent to cohort members every other year to update exposure information and to ascertain incident cancer cases; a >87% response rate has been achieved for each follow-up questionnaire. Incident cancers were verified through medical records, state cancer registries, or death certificates. During 2000–2002, oral wash samples were collected by mail from 70 004 cohort members.[18] This analysis includes pancreatic cancer cases diagnosed between oral wash collection and December 2008.

The PLCO cohort[19] is a large population-based randomised trial that examined the effects of screening on cancer-related mortality and secondary endpoints, in men and women aged 55–74, recruited between 1993 and 2001, and followed for cancer incidence. Participants were randomised to either a screening or control arm. Oral wash samples were collected in the control arm (n=52 000). During the follow-up, incident cancers were ascertained by an annual mailed questionnaire (>95% follow-up rate), and verified through medical records or death certificates. This analysis includes pancreatic cancer cases diagnosed between oral wash collection and December 2010.

Nested case–control study selection
Cases are subjects from the two cohorts with incident primary pancreatic adenocarcinoma (ICD–O–2 codes C25.0–C25.3, C25.7–25.9), with no prior history of cancer (except non-melanoma skin cancer), a valid consent and pre-diagnostic oral wash samples. Cohort nested controls were selected by incidence-density sampling without replacement among cohort members who had no cancer prior to selection, provided a valid consent and an oral wash sample. Controls were matched to cases by cohort, age (5-year), sex, race (white, other) and calendar year of oral wash collection. As of December 2012, 170 cases and 170 controls from CPS II and 191 cases and 201 controls from PLCO were eligible for this project (see online supplementary figure S1).

Measurements
Assessment of questionnaire-based covariates
At enrolment and follow-up periods, participants in both cohorts completed a structured questionnaire that included questions about age, sex, race, body mass index (BMI), smoking status, alcohol consumption and history of diabetes. We used covariates from the questionnaires most closely in time to oral wash sample collection for each participant. The relevant cohort data were harmonised and transferred to New York University School of Medicine.

Oral wash samples
Oral mouthwash samples were originally collected in both cohorts in lieu of blood samples for the purpose of obtaining buccal cell DNA for human genome research.[17 19] Participants in both cohorts were asked to swish vigorously with 10 mL Scope mouthwash (P&G), and then to expectorate into a specimen tube. These samples were shipped to each cohort's biorepository, pelleted and stored at -80°C until use.[18 19] Our assay uses DNA for oral microbiome sequencing; DNA is highly stable in a frozen state, thus allowing us to use the frozen oral wash pellets archived in the CPS II and PLCO cohorts. We have previously shown suitability of this sample type for oral microbiome measurement,[20] and excellent quality control (QC) agreement for these samples.[21] Because our hypothesis is that oral bacteria influence risk of pancreas cancer through intermittent yet persistent migration of oral bacteria to the pancreas, oral wash samples which contain the broad spectrum of oral bacteria, are suitable to test this hypothesis.[22]

Oral microbiota characterisation using 16S rRNA gene amplification and sequencing
Bacterial genomic DNA was extracted from mouth wash samples using the MoBio PowerSoil DNA Isolation Kit (Carlsbad, California, USA), with the bead-beating method in the MoBio Powerlyzer instrument at Dr Ahn's Laboratory. From extracted DNA, we amplified the 16S rRNA gene V3-V4 regions using universal primers (347F 5′-GGAGGCAGCAG TRRGGAAT′-3′ and 803R 5′-CTACCRGGGTATCTAATCC-3′), while incorporating adapters and a sample-specific barcode sequence.[23] The Roche 454 FLX Titanium pyrosequencing system was used to sequence the resulting amplicons. The sequencing data will be submitted to dbGaP.

Derivation of microbiome data
Multiplexed and barcoded sequences were deconvoluted. Poor-quality sequences were excluded using the default parameters of the QIIME script *split_libraries.py* (minimum average

quality score=25, minimum/maximum sequence length=200/1000 base pairs, no ambiguous base calls, and no mismatches allowed in the primer sequence).[24] Finally, chimeric sequences were removed with ChimeraSlayer.[25] From 732 pre-diagnostic oral wash samples, we obtained 9 354 571 high quality 16S rRNA gene sequence reads (mean 11 782 (SD 2799) sequences per sample), with similar number of reads in both cohorts (see online supplementary table S1). Filtered sequences were clustered into operational taxonomic units (OTUs), with 97% identity, and assigned to taxonomy using the Human Oral Microbiome Database (HOMD).[26]

Quality control

Laboratory personnel were blinded to case/control status, and matched pairs were processed side by side. Blinded positive QC specimens were used across all sequencing batches. Twenty-two repeats from two blinded subjects were inserted for the CPS II samples, and nine repeats from two blinded subjects were inserted for the PLCO samples. We previously reported that QC samples had good agreement in these QC subjects:[21] coefficient of variability ranged from 8.90% to 11.10% for the Shannon-Wiener index, and 0.92% to 4.68% for the Simpson index, both measures of within-subject bacterial community diversity. Negative controls samples (without DNA template) were used to detect possible reagent and environment contamination in all sequencing batches as well.

Statistical analysis

The relationship between overall oral microbiota composition and pancreatic cancer was assessed by analysis of weighted and unweighted UniFrac distances[27]; these metrics assess the phylogenetic similarity of bacterial community pairs, taking into account OTU relative abundance or presence/absence, respectively. To visualise separation of subjects based on pairwise distances, principal coordinate analysis (PCoA) plots were generated using the first two principal coordinates and labelled according to pancreatic cancer status. Permutational mutivariate analysis of variance using distance matrices (PERMANOVA) ('Adonis' function, vegan package, R)[28] of the UniFrac distance was used to test differences in overall oral microbiome composition according to pancreatic cancer status, controlling for potential confounders (age, sex, race, BMI, smoking status, alcohol consumption status and history of diabetes).

We first assessed pancreatic cancer risk in relation to four pre-defined periodontal pathogens (*Porphyromonas gingivalis*, *Aggregatibacter actinomycetemcomitans*, *Tannerella forsythia* and *Prevotella intermedia*).[29] Because relative abundance of these pathogens was low, we characterised individuals as carriers and non-carriers of the pathogens, with non-carriers having zero sequence reads for the specific pathogen. For case–control comparison of carriage of these pathogens, we used logistic regression models controlling for the random effects of cohorts,[30] with inclusion of other covariates (age, sex, race, BMI, smoking status, alcohol consumption status and history of diabetes). Results were similar, when we used conditional logistic regression modelling (see online supplementary results).

We also broadly compared bacterial taxa (phylum to genus) between cases and controls. We limited our analysis of bacterial phyla to those with mean relative abundance ≥0.01%. For lower level taxa (class to genus), we limited analysis to those with mean relative abundance ≥0.0001%. Our analyses included 5 phyla, 15 classes, 22 orders, 36 families and 62 genera. To explore the association of taxa relative abundance at each level (phylum to genus) with pancreatic cancer risk, we conducted L1

penalised least absolute shrinkage and selection operator (LASSO) logistic regression[31] implemented in R 'glmnet' package,[32] following a previous approach to disentangle the effect of metformin treatment on human gut microbiota.[33] This method penalises the sum of the absolute values of regression coefficients, resulting in a parsimonious model where only the taxa with the strongest associations with the outcome will be selected. Covariates (cohort, age, sex, race, BMI, smoking status, alcohol consumption status, history of diabetes) were controlled in the LASSO taxa selection process. To evaluate the risk associations for the selected taxa, we fit traditional logistic regression models using each selected taxon as a predictor of pancreatic cancer risk individually, as described in the carriage of pathogens analysis. Further, we conducted stratified analyses according to smoking status, alcohol consumption status and time to cancer diagnosis. Analyses were carried out using R V.3.2.0. Further description of each methodology can be found in the see online supplementary material methods.

RESULTS

Demographic characteristics of cases and controls are shown in table 1. The majority of participants in both cohorts were white, while subjects in the CPS II cohort tended to be older than those in the PLCO cohort. Cases and controls in both cohorts were similar with respect to the matching factors of age, sex and race. Cases were more likely to be current smokers in the PLCO cohort, and to consume more alcoholic beverages in the CPS II cohort.

We first examined overall oral microbiota composition in relation to subsequent risk of pancreatic cancer. The first and second components of PCoA based on weighted UniFrac distances were plotted (see online supplementary figure S2). PERMANOVA analysis, controlling for covariates, showed no differences in either cohort, with respect to overall phylogenetic distance of oral microbiome composition, between subjects who went on to develop pancreatic cancer (cases) and those that did not (controls) (CPS II, p=0.85 and PLCO, p=0.49) (see online supplementary figure S2). Similar results were found using the unweighted UniFrac distance (CPS II, p=0.43 and PLCO, p=0.87).

We next examined the relationship of pre-selected oral periodontal pathogens with subsequent risk of pancreatic cancer (table 2). Carriage of *P. gingivalis* was associated with a higher risk of pancreatic cancer (adjusted OR for presence vs absence=1.60 and 95% CI 1.15 to 2.22). This association was observed for low carriers (below median relative abundance, OR 1.47 (0.96 to 2.26)) and high carriers (above median relative abundance, OR 1.73 (1.14 to 2.63)), exhibiting a significant dose–response relationship (p trend=0.0047). Furthermore, these associations were consistent in CPS II (p trend=0.032; see online supplementary table S2a) and PLCO (p trend=0.070; see online supplementary table S2b). Carriage of *A. actinomycetemcomitans* was also associated with increased risk of pancreatic cancer (OR 2.20 (1.16 to 4.18)), but overall carriage prevalence was low and this relationship could not be evaluated for dose–response. Carriage of *T. forsythia* and *P. intermedia* were not associated with the risk of pancreatic cancer.

We carried out a comparison in cases and controls of the relative abundance of oral bacterial taxa on all taxonomic levels (table 3 and see online supplementary figure S3). We found that greater abundance of *Fusobacteria* was associated with decreased risk of pancreatic cancer (adjusted OR per per cent increase of relative abundance=0.94 and 95% CI (0.89 to 0.99)), as well as its genus *Leptotrichia* (OR 0.87 (0.79 to 0.95)). Genus

Table 1 Distribution of demographic factors among pancreatic cancer cases and controls in the CPS II and PLCO cohorts

	CPS II cohort					PLCO cohort				
	Cases (n=170)		Controls (n=170)			Cases (n=191)		Controls (n=201)		
	No	Per cent	No	Per cent	p Value*	No	Per cent	No	Per cent	p Value*
Age (years)†	73.7	5.7	73.7	5.7	0.95	63.8	5.2	63.8	5.4	0.96
Sex										
Male	90	52.9	90	52.9	1.00	116	60.7	122	60.7	1.00
Female	80	47.1	80	47.1		75	39.3	79	39.3	
Race										
White	163	95.9	162	95.3	1.00	175	91.6	183	91.0	0.98
Non-white	7	4.1	8	4.7		16	8.4	18	9.0	
BMI (kg/m^2)										
<24.99	67	39.4	72	42.3	0.74	65	34.0	61	30.3	0.63
25–29.99	73	42.9	66	38.8		72	37.7	84	41.8	
≥30	27	15.9	29	17.1		52	27.2	51	25.4	
Missing	3	1.8	3	1.8		2	1.1	5	2.5	
Smoking status										
Never smoker	72	42.4	78	45.9	0.61	82	42.9	108	53.7	0.0033
Former smoker	90	52.9	87	51.2		80	41.9	82	40.8	
Current smoker	8	4.7	5	2.9		29	15.2	11	5.5	
Alcohol consumption										
Never drink	40	23.5	32	18.8	0.039	42	22.0	53	26.3	0.63
<1 standard drink per day	62	36.5	85	50.0		91	47.6	90	44.8	
≥1 standard drink per day	41	24.1	29	17.1		35	18.3	38	18.9	
Missing	27	15.9	24	14.1		23	12.1	20	10.0	
Diabetes										
Yes	25	14.7	23	13.5	0.88	18	9.4	17	8.5	0.87
No	145	85.3	147	86.5		171	89.6	182	90.6	
Missing						2	1.0	2	0.9	

*p Values were based on t-test or χ^2 test (two-sided).
†Mean and SD instead of N and percentage were reported for age.
BMI, body mass index; CPS, Cancer Prevention Study; PLCO, prostate, lung, colorectal and ovarian cancer.

Alloprevotella may also be associated with pancreatic cancer risk (OR 1.20 (1.01 to 1.43)). Results were similar after further adjustment for the carriage of *P. gingivalis* and *A. actinomycetemcomitans*. On further analysis, the risks identified for these two genera could not be associated with specific species. The inverse associations of *Fusobacteria* and *Leptotrichia* with pancreatic cancer risk were consistent in the CPS II cohort (see online supplementary table S3a) and PLCO cohort (see online supplementary table S3b), although the effect of *Alloprevotella* was restricted to the CPS II cohort.

As smoking and alcohol consumption have been related to risk for pancreatic cancer and these factors impact oral microbial community structure, we examined the relationship of selected microbes with pancreatic cancer risk in ever versus never-smokers, and in ever versus never-drinkers (table 4). The association of *A. actinomycetemcomitans* with pancreatic cancer risk tended to be greater in ever-drinkers (OR 3.03 (1.31 to 7.03)) than in never-drinkers (OR 0.47 (0.13 to 1.72)), although these differences were not statistically significant (p for interaction=0.086). Additionally, there was no evidence that the associations of other selected phylotypes with pancreatic cancer risk differed by smoking or alcohol consumption status (all p for interaction >0.05).

Because pancreatic cancer could potentially influence oral microbial composition, we also explored risk associations for selected microbes stratified by time between sample collection and diagnosis of pancreatic cancer (table 4). Risks related to

P. gingivalis and *Leptotrichia* remained after exclusion of cases that developed within 2 years of sample collection, reducing the likelihood of reverse causation in this prospective study.

DISCUSSION

In this first prospective evaluation of the oral microbiome and risk for pancreatic cancer, we demonstrated that carriage of the periodontal pathogens *P. gingivalis* and *A. actinomycetemcomitans*, and decreased relative abundance of *Fusobacteria* and its genus *Leptotrichia*, are associated with subsequent risk of pancreatic cancer. We provided evidence that these associations were unlikely due to smoking or other potential confounders. Because of the prospective design, the likelihood that these associations might be due to reverse causation was minimised.[34] Given these considerations, the findings from our study point to a potential aetiological role of oral bacteria in pancreatic cancer.

Although this is the first study to prospectively examine the association between oral bacterial pathogens and pancreatic cancer risk using direct measurement of oral bacteria, there is extensive epidemiological evidence supporting this finding. *P. gingivalis* and *A. actinomycetemcomitans* are keystone pathogens among oral bacteria involved in the initiation of periodontal disease and tooth loss.[35] Previous studies reported that a history of periodontal disease and tooth loss are associated prospectively with increased risk of pancreatic cancer.[10–14] Providing further, although indirect, evidence of an association of pancreatic cancer risk with *P. gingivalis*, a large European

Table 2 Carriage of periodontal pathogens in pancreatic cancer cases and controls in the combined CPS II and PLCO cohort

Periodontal pathogens	Cases (n=361) N (%)	Controls (n=371) N (%)	OR*	95% CI*	p Value*	p Trend†
P. Gingivalis						
Non-carriers	233 (64.5)	275 (74.1)	1.00	–	–	
Any carriers	128 (35.5)	96 (25.9)	1.60	1.15 to 2.22	0.0048	
Non-carriers	233 (64.6)	275 (74.2)	1.00	–	–	0.0047
Low carriers‡	59 (16.3)	48 (12.9)	1.47	0.96 to 2.26	0.078	
High carriers‡	69 (19.1)	48 (12.9)	1.73	1.14 to 2.63	0.011	
A. actinomycetemcomitans§						
Non-carriers	330 (91.4)	356 (96.0)	1.00	–	–	
Any carriers	31 (8.6)	15 (4.0)	2.20	1.16 to 4.18	0.016	
T. forsythia						
Non-carriers	191 (52.9)	210 (56.6)	1.00	–	–	
Any carriers	170 (47.1)	161 (43.4)	1.16	0.86 to 1.55	0.34	
P. intermedia						
Non-carriers	279 (77.3)	301 (81.1)	1.00	–	–	
Any carriers	82 (22.7)	70 (18.9)	1.30	0.89 to 1.88	0.17	

*ORs, 95% CIs, and p values were calculated from logistic regression after controlling for the random effect of cohorts as well as other covariates (age, race, sex, BMI, smoking status, alcohol consumption status and history of diabetes).
†p Trend was calculated by assigning values 0–2 to the non-carriers, low carriers and high carriers of *P. gingivalis*, respectively, and treating this as a continuous variable in the logistic regression model, after controlling for the random effect of cohorts as well as other covariates (age, race, sex, BMI, smoking status, alcohol consumption status and history of diabetes).
‡Cut-off point was based on the median relative abundance of *P. gingivalis* carriers in the control group of each cohort (the median relative abundance of *P. gingivalis* among controls is 0.12% in CPS II cohort, and 0.056% in PLCO cohort).
§Dose–response could not be evaluated due to low overall carriage prevalence.
A. actinomycetemcomitans, aggregatibacter actinomycetemcomitans; BMI, body mass index; CPS, Cancer Prevention Study; P. gingivalis, porphyromonas gingivalis; P. intermedia, prevotella intermedia; PLCO, prostate, lung, colorectal and ovarian cancer; T. forsythia, tannerella forsythia.

Table 3 Median relative abundance of selected oral microbial taxa* in pancreatic cancer cases and controls in the combined CPS II and PLCO cohort

| | Median relative abundance | | | | | |
	Cases (n=361)	Controls (n=371)	OR†	95% CI†	p Value†	p Value‡
Phylum						
Fusobacteria	2.37	2.70	0.94	0.89 to 0.99	0.014	0.0079
Bacteroidetes	9.06	8.69	1.01	0.99 to 1.03	0.28	0.39
Class						
Fusobacteria	2.37	2.70	0.94	0.89 to 0.99	0.014	0.0079
SR1[C-1]	0.00	0.00	1.14	0.95 to 1.37	0.15	0.25
Order						
Fusobacteriales	2.37	2.70	0.94	0.89 to 0.99	0.014	0.0079
Family						
Leptotrichiaceae	0.78	1.03	0.87	0.79 to 0.95	0.0029	0.0027
Genus						
Leptotrichia	0.78	1.03	0.87	0.79 to 0.95	0.0029	0.0027
Alloprevotella	0.27	0.25	1.20	1.01 to 1.43	0.034	0.10

*Taxa were selected from L1-penalised LASSO logistic models with the optimal value of lambda from 100 repeated 10-fold cross-validation on each taxonomic level. Bacterial phyla with the mean relative abundance greater than 0.01% and taxa (class-genus) with the mean relative abundance greater than 0.0001% were eligible for the variable selection.
†ORs, 95% CIs and p Values for 1% increase in relative abundance were calculated from logistic regression models, after controlling for the random effects of cohorts and other covariates (age, race, sex, BMI, smoking status, alcohol consumption status and history of diabetes).
‡Additionally controlled for the carriage of *Porphyromonas gingivalis* and *Aggregatibacter actinomycetemcomitans*.
BMI, body mass index; CPS, Cancer Prevention Study; PLCO, prostate, lung, colorectal and ovarian cancer; LASSO, least absolute shrinkage and selection operator.

cohort found that elevated serum antibodies to *P. gingivalis* were associated with a two-fold increased risk of pancreatic cancer.[15]

While a consistent picture is forming that pancreatic cancer is related to poor oral health status and to bacterial drivers of oral disease, particularly *P. gingivalis*, several questions regarding the underlying biological pathways to pancreatic cancer remain unanswered. First, do oral bacteria populate distant disease sites? Swidsinski *et al*[36] reported, using fluorescence in situ hybridisation, that a dense multispecies bacterial biofilm,

including oral bacterial types, was present within the pancreatic duct of patients with calcific pancreatitis; however, whether bacteria of the same clonal origin are found in the pancreas and the oral cavity is unknown. Further evidence of potential dissemination from the oral cavity in humans are observations that these bacterial types are found in atherosclerotic plaques,[37 38] distal oesophageal tissue,[39] the brains of Alzheimer's patients,[40] and in the foeto-placental unit.[41] Second, how would oral bacteria populate these sites? While the route of transmission of oral

Table 4 Oral bacteria taxa* and risk of pancreatic cancer stratified by smoking, alcohol consumption and time interval between oral sample collection and diagnosis date in the combined CPS II and PLCO cohort

	Cases/controls (N)	OR†	95% CI†	p Value†
Ever smokers	207/185			
A. Actinomycetemcomitans		2.19	0.87 to 5.48	0.096
P. Gingivalis		1.38	0.89 to 2.16	0.15
Leptotrichia		0.86	0.76 to 0.97	0.014
Alloprevotella		1.16	0.93 to 1.45	0.19
Never smokers	154/186			
A. Actinomycetemcomitans		1.78	0.69 to 4.62	0.23
P. Gingivalis		1.75	1.06 to 2.90	0.029
Leptotrichia		0.85	0.73 to 1.01	0.059
Alloprevotella		1.18	0.87 to 1.62	0.29
Ever drinkers	279/286			
A. actinomycetemcomitans		3.03	1.31 to 7.03	0.0097
P. gingivalis		1.50	1.03 to 2.20	0.035
Leptotrichia		0.82	0.73 to 0.92	<0.001
Alloprevotella		1.20	0.97 to 1.47	0.086
Never drinkers	82/85			
A. actinomycetemcomitans		0.47	0.13 to 1.72	0.25
P. gingivalis		1.70	0.80 to 3.62	0.17
Leptotrichia		1.02	0.81 to 1.28	0.87
Alloprevotella		1.29	0.81 to 2.04	0.28
Oral sample collected ≤2 years prior to cancer diagnosis	85/97‡			
A. actinomycetemcomitans		2.98	0.71 to 12.43	0.13
P. gingivalis		1.60	0.77 to 3.31	0.21
Leptotrichia		0.92	0.76 to 1.12	0.42
Alloprevotella		1.13	0.81 to 1.57	0.47
Oral sample collected >2 years prior to cancer diagnosis	276/280‡			
A. actinomycetemcomitans		1.60	0.75 to 3.38	0.22
P. gingivalis		1.53	1.05 to 2.24	0.027
Leptotrichia		0.84	0.75 to 0.94	0.0027
Alloprevotella		1.21	0.97 to 1.50	0.090

All p for interactions were >0.05 (4 taxa×alcohol, smoking, or time interval).
*Taxa in the models were mutually adjusted.
†Values of ORs, 95% CIs and p values from logistic regression models after controlling for the random effect of cohorts as well as other covariates (age, race, sex, BMI, smoking status, alcohol consumption status and history of diabetes).
‡Controls were included in the stratum of their matched case; six controls were used more than once due to matched cases having different time interval.
A. actinomycetemcomitans, Aggregatibacter actinomycetemcomitans; BMI, body mass index; CPS, Cancer Prevention Study; P. gingivalis, Porphyromonas gingivalis; PLCO, Prostate, Lung, Colorectal and Ovarian Cancer.

bacteria to the pancreas is uncertain, an intermittent yet persistent migration of oral bacteria—by swallowing or via the circulatory system after mastication and personal oral hygiene (tooth brushing and flossing)[42 43]—may reach the pancreas and other sites, as we previously reviewed.[22] Third, can oral pathogens participate in pancreatic carcinogenesis? *P. gingivalis* has been shown to have potential to evade the host immune system by invasion of host cells and disruption of signalling pathways by cytokine and receptor degradation.[44 45] In addition, *P. gingivalis* and *A. actinomycetemcomitans* can initiate the Toll-like receptor (TLR) signalling pathways, and TLR activation is a critical driver of pancreatic carcinogenesis in animal models, as we (Dr Miller) recently demonstrated.[46]

Recognising that evidence of a direct relationship between *P. gingivalis*, related oral pathogens, and pancreatic cancer is preliminary, there is also the alternative hypothesis that pancreatic and oral conditions could be due to parallel, independent systemic inflammatory processes.[47] Immune response is strongly implicated in pancreas carcinogenesis, and several established pancreatic cancer risk factors, including chronic pancreatitis, tobacco smoking and type 2 diabetes, may have an inflammatory component.[48] Good oral hygiene and normal host immune response are sufficient to minimise periodontal pathogens, while failures in immune defence could lead to higher bacterial load and the development of gingival inflammation.[49] While our study points to *P. gingivalis* and related oral pathogens as novel risk factors for pancreatic cancer, it will require further experimental research to establish the mechanistic basis of these relationships.

We also found that a higher relative abundance of the phylum *Fusobacteria* and its genus *Leptotrichia* was associated with decreased risk of pancreatic cancer. This finding is consistent with the results from a nested case–control study from a European cohort, in which higher overall blood antibody levels to commensal oral bacteria were associated with reduced pancreatic cancer risk (including antibodies from *Fusobacteria* species, though not *Leptotrichia* specifically),[15] but inconsistent with a recent cross-sectional study of 8 patients, which found higher abundance of *Leptotrichia* in saliva of pancreatic cancer patients compared to controls.[50] *Fusobacteria* are anaerobic,

non-spore-forming, Gram-negative bacilli and two major families, *Leptotrichiaceae* and *Fusobacteriaceae* (see online supplementary figure S3; phylogenetic tree), are frequently found in the human oral cavity.[6] Within *Leptotrichia* there is wide genetic diversity, with six species belonging to this genus identified thus far;[51] however, the characteristics and the role of *Leptotrichia* in human health remain unclear. *Leptotrichia* is regarded as an opportunistic pathogen, which tends to cause disease in the presence of local or general pre-disposing factors and has been reported to be involved in various human infections,[51–53] including periodontitis, lung abscess, pneumonia, osteomyelitis, endocarditis and bacteraemia, the latter primarily in immune-depressed patients. Moreover, *Leptotrichia* elicits a human immune response, and the serum antibody to it is common.[51] It is possible that the immune response elicited by *Leptotrichia*, rather than *Leptotrichia* itself, may provide protection against pancreatic carcinogenesis.[54] Further human and experimental research is needed to confirm this relationship, and disentangle the complex role of immune response in pancreatic carcinogenesis.

A limitation of our study is the one-time collection of mouthwash samples, reducing our ability to capture a potentially dynamic relationship between the oral microbiota and pancreatic cancer development. We additionally lacked information on the periodontal disease status of the study participants, which would allow us to determine whether periodontal pathogens are implicated in pancreatic cancer independently of periodontal disease. The population studied was predominantly white and was limited to study volunteers, potentially limiting the generalisability of these findings to the general population and other higher-risk groups. We recognise that some controls may have undiagnosed or unreported pancreatic cancer, however, this is unlikely to be common given the low survival rate of pancreatic cancer and the ascertainment of all cancer deaths through death registries in both cohorts.[17 19] Lastly, we did not have information on whether any of the study participants were related to each other; this phenomenon, if present, could have introduced confounding due to the clustering of microbiome composition and cancer within families.

Our study has several strengths. First, the study was conducted prospectively, using oral samples collected up to 10 years before cancer diagnosis, providing the opportunity to determine the temporal relationship between oral microbiota and subsequent development of pancreatic cancer. Additionally, the prospective design of our nested case–control study with incidence density sampling of controls avoids the selection bias that may occur in traditional case–control designs. Our study used comprehensive 16S rRNA screening for the oral microbiome, substantially broadening our understanding of overall bacterial structure and abundance of bacterial types in relation to pancreatic cancer. The relatively large samples size, including 361 pancreatic cancer cases and 371 matched controls, provided sufficient statistical power to detect relevant associations between hypothesised explanatory factors and pancreatic cancer risk. Furthermore, we adjusted for established and possible risk factors for pancreatic cancer throughout the analysis, including age, race, sex, smoking status, alcohol consumption, BMI and history of diabetes.

In conclusion, this prospective nested case–control study demonstrated that carriage of *P. gingivalis* and *A. actinomycetemcomitans*, and decreased relative abundance of phylum *Fusobacteria* and its genus *Leptotrichia*, are related to subsequent increased risk of pancreatic cancer. The study provides evidence that the oral microbiota may play a role in the aetiology of pancreatic cancer.

Author affiliations
¹Department of Population Health, New York University School of Medicine, New York, New York, USA
²Departments of Public Health Sciences and Oral Health Sciences, Biomedical Informatics Center, Program for Human Microbiome Research, Medical University of South Carolina, Charleston, South Carolina, USA
³Epidemiology Research Program, American Cancer Society, Atlanta, Georgia, USA
⁴Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA
⁵Department of Surgery, New York University School of Medicine, New York, New York, USA
⁶Department of Cell Biology, New York University School of Medicine, New York, New York, USA
⁷NYU Perlmutter Cancer Center, New York, New York, USA
⁸Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland, USA

REFERENCES

1. Howlader N, Noone AM, Krapcho M, *et al*. SEER Cancer Statistics Review, 1975–2010. National Cancer Institute. Bethesda, MD, based on November 2 SEER data submission, posted to the SEER web site, 3 April 2016. http://seer.cancer.gov/csr/1975_2010/
2. Yadav D, Lowenfels AB. The epidemiology of pancreatitis and pancreatic cancer. *Gastroenterology* 2013;144:1252–61.
3. Iodice S, Gandini S, Maisonneuve P, *et al*. Tobacco and the risk of pancreatic cancer: a review and meta-analysis. *Langenbecks Arch Surg* 2008;393:535–45.
4. Chari ST, Leibson CL, Rabe KG, *et al*. Pancreatic cancer-associated diabetes mellitus: prevalence and temporal association with diagnosis of cancer. *Gastroenterology* 2008;134:95–101.
5. Berrington de Gonzalez A, Sweetland S, Spencer E. A meta-analysis of obesity and the risk of pancreatic cancer. *Br J Cancer* 2003;89:519–23.
6. Dewhirst FE, Chen T, Izard J, *et al*. The human oral microbiome. *J Bacteriol* 2010;192:5002–17.
7. Slocum C, Kramer C, Genco CA. Immune dysregulation mediated by the oral microbiome: potential link to chronic inflammation and atherosclerosis. *J Intern Med* 2016;280:114–28.
8. Homann N, Tillonen J, Meurman JH, *et al*. Increased salivary acetaldehyde levels in heavy drinkers and smokers: a microbiological approach to oral cavity cancer. *Carcinogenesis* 2000;21:663–8.
9. Moye ZD, Zeng L, Burne RA. Fueling the caries process: carbohydrate metabolism and gene regulation by Streptococcus mutans. *J Oral Microbiol* 2014;6:24878–92.
10. Ahn J, Segers S, Hayes RB. Periodontal disease, *Porphyromonas gingivalis* serum antibody levels and orodigestive cancer mortality. *Carcinogenesis* 2012;33:1055–8.
11. Michaud DS, Joshipura K, Giovannucci E, *et al*. A prospective study of periodontal disease and pancreatic cancer in US Male health professionals. *J Natl Cancer Inst* 2007;99:171–5.
12. Hiraki A, Matsuo K, Suzuki T, *et al*. Teeth loss and risk of cancer at 14 common sites in Japanese. *Cancer Epidemiol Biomarkers Prev* 2008;17:1222–7.
13. Stolzenberg-Solomon RZ, Dodd KW, Blaser MJ, *et al*. Tooth loss, pancreatic cancer, and *Helicobacter pylori*. *Am J Clin Nutr* 2003;78:176–81.
14. Hujoel PP, Drangsholt M, Spiekerman C, *et al*. An exploration of the periodontitis-cancer association. *Ann Epidemiol* 2003;13:312–16.

15 Michaud DS, Izard J, Wilhelm-Benartzi CS, *et al*. Plasma antibodies to oral bacteria and risk of pancreatic cancer in a large European prospective cohort study. *Gut* 2013;62:1764–70.

16 Darveau RP. Periodontitis: a polymicrobial disruption of host homeostasis. *Nat Rev Microbiol* 2010;8:481–90.

17 Calle EE, Rodriguez C, Jacobs EJ, *et al*. The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. *Cancer* 2002;94:2490–501.

18 Feigelson HS, Rodriguez C, Welch R, *et al*. Successful genome-wide scan in paired blood and buccal samples. *Cancer Epidemiol Biomarkers Prev* 2007;16:1023–5.

19 Hayes RB, Reding D, Kopp W, *et al*. Etiologic and early marker studies in the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial. *Control Clin Trials* 2000;21:349S–55S.

20 Wu J, Lin I, Hayes RB, *et al*. Comparison of DNA extraction methods for human oral microbiome research. *Br J Med Med Res* 2014;4:1980–91.

21 Wu J, Peters BA, Dominianni C, *et al*. Cigarette smoking and the oral microbiome in a large study of American adults. *ISME J* 2016;10:2435–46.

22 Ahn J, Chen CY, Hayes RB. Oral microbiome and oral and gastrointestinal cancer risk. *Cancer Causes Control* 2012;23:399–404.

23 Nossa CW, Oberdorf WE, Yang LY, *et al*. Design of 16S rRNA gene primers for 454 pyrosequencing of the human foregut microbiome. *World J Gastroentero* 2010;16:4135–44.

24 Caporaso JG, Kuczynski J, Stombaugh J, *et al*. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;7:335–6.

25 Haas BJ, Gevers D, Earl AM, *et al*. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* 2011;21:494–504.

26 Chen T, Yu WH, Izard J, *et al*. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database (Oxford)* 2010;2010:baq013.

27 Lozupone C, Knight R. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Appl Environ Microbiol* 2005;71:8228–35.

28 Dixon P. VEGAN, a package of R functions for community ecology. *J Veg Sci* 2003;14:927–30.

29 Newman MG. *Carranza's clinical periodontology*. Philadelphia, PA; London: Saunders, 2012.

30 Mclean RA, Sanders WL, Stroup WW. A unified approach to mixed linear-models. *Am Statist* 1991;45:54–64.

31 Tibshirani R. Regression Shrinkage and Selection via the Lasso. *J R Stat Soc B (Methodol)* 1996;58:267–88.

32 Friedman JH, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *2010* 2010;33:22.

33 Forslund K, Hildebrand F, Nielsen T, *et al*. Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* 2015;528:262–6.

34 Farrell JJ, Zhang L, Zhou H, *et al*. Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer. *Gut* 2012;61:582–8.

35 Costalonga M, Herzberg MC. The oral microbiome and the immunobiology of periodontal disease and caries. *Immunol Lett* 2014;162:22–38.

36 Swidsinski A, Schlien P, Pernthaler A, *et al*. Bacterial biofilm within diseased pancreatic and biliary tracts. *Gut* 2005;54:388–95.

37 Koren O, Spor A, Felin J, *et al*. Human oral, gut, and plaque microbiota in patients with atherosclerosis. *Proc Natl Acad Sci USA* 2011;108(Suppl 1):4592–8.

38 Gaetti-Jardim E Jr, Marcelino SL, Feitosa AC, *et al*. Quantitative detection of periodontopathic bacteria in atherosclerotic plaques from coronary arteries. *J Med Microbiol* 2009;58:1568–75.

39 Pei Z, Bini EJ, Yang L, *et al*. Bacterial biota in the human distal esophagus. *Proc Natl Acad Sci USA* 2004;101:4250–5.

40 Riviere GR, Riviere KH, Smith KS. Molecular and immunological evidence of oral Treponema in the human brain and their association with Alzheimer's disease. *Oral Microbiol Immunol* 2002;17:113–18.

41 Aagaard K, Ma J, Antony KM, *et al*. The placenta harbors a unique microbiome. *Sci Transl Med* 2014;6:237–65.

42 Crasta K, Daly CG, Mitchell D, *et al*. Bacteraemia due to dental flossing. *J Clin Periodontol* 2009;36:323–32.

43 Lockhart PB, Brennan MT, Sasser HC, *et al*. Bacteremia associated with toothbrushing and dental extraction. *Circulation* 2008;117:3118–25.

44 Stathopoulou PG, Benakanakere MR, Galicia JC, *et al*. The host cytokine response to *Porphyromonas gingivalis* is modified by gingipains. *Oral Microbiol Immunol* 2009;24:11–17.

45 Duncan L, Yoshioka M, Chandad F, *et al*. Loss of lipopolysaccharide receptor CD14 from the surface of human macrophage-like cells mediated by *Porphyromonas gingivalis* outer membrane vesicles. *Microb Pathog* 2004;36:319–25.

46 Zambirinis CP, Levie E, Nguy S, *et al*. TLR9 ligation in pancreatic stellate cells promotes tumorigenesis. *J Exp Med* 2015;212:2077–94.

47 Otomo-Corgel J, Pucher JJ, Rethman MP, *et al*. State of the science: chronic periodontitis and systemic health. *J Evid Based Dent Pract* 2012;12:20–8.

48 Greer JB, Whitcomb DC. Inflammation and pancreatic cancer: an evidence-based review. *Curr Opin Pharmacol* 2009;9:411–18.

49 Hajishengallis G, Liang S, Payne MA, *et al*. Low-abundance biofilm species orchestrates inflammatory periodontal disease through the commensal microbiota and complement. *Cell Host Microbe* 2011;10:497–506.

50 Torres PJ, Fletcher EM, Gibbons SM, *et al*. Characterization of the salivary microbiome in patients with pancreatic cancer. *PeerJ* 2015;3:e1373.

51 Eribe ER, Olsen I. Leptotrichia species in human infections. *Anaerobe* 2008;14:131–7.

52 Couturier MR, Slechta ES, Goulston C, *et al*. Leptotrichia bacteremia in patients receiving high-dose chemotherapy. *J Clin Microbiol* 2012;50:1228–32.

53 Cooreman S, Schuermans C, Van Schaeren J, *et al*. Bacteraemia caused by Leptotrichia trevisanii in a neutropenic patient. *Anaerobe* 2011;17:1–3.

54 Inman KS, Francis AA, Murray NR. Complex role for the immune system in initiation and progression of pancreatic cancer. *World J Gastroenterol* 2014;20:11160–81.