JAMA Oncology | Original Investigation

Association of Oral Microbiome With Risk for Incident Head and Neck Squamous Cell Cancer

Richard B. Hayes, DDS, PhD; Jiyoung Ahn, PhD; Xiaozhou Fan, MS; Brandilyn A. Peters, PhD; Yingfei Ma, PhD; Liying Yang, MD; Ilir Agalliu, ScD, MD; Robert D. Burk, MD; Ian Ganly, MD; Mark P. Purdue, PhD; Neal D. Freedman, PhD; Susan M. Gapstur, PhD; Zhiheng Pei, MD, PhD

IMPORTANCE Case-control studies show a possible relationship between oral bacteria and head and neck squamous cell cancer (HNSCC). Prospective studies are needed to examine the temporal relationship between oral microbiome and subsequent risk of HNSCC.

OBJECTIVE To prospectively examine associations between the oral microbiome and incident HNSCC.

DESIGN, SETTING, AND PARTICIPANTS This nested case-control study was carried out in 2 prospective cohort studies: the American Cancer Society Cancer Prevention Study II Nutrition Cohort (CPS-II) and the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO). Among 122 004 participants, 129 incident patient cases of HNSCC were identified during an average 3.9 years of follow-up. Two controls per patient case (n = 254) were selected through incidence density sampling, matched on age, sex, race/ethnicity, and time since mouthwash collection. All participants provided mouthwash samples and were cancer-free at baseline.

EXPOSURES Oral microbiome composition and specific bacterial abundances were determined through bacterial 16S rRNA gene sequencing. Overall oral microbiome composition and specific taxa abundances were compared for the case group and the control group, using PERMANOVA and negative binomial generalized linear models, respectively, controlling for age, sex, race, cohort, smoking, alcohol, and oral human papillomavirus-16 status. Taxa with a 2-sided false discovery rate (FDR)-adjusted *P*-value (*q*-value) <.10 were considered significant.

MAIN OUTCOMES AND MEASURES Incident HNSCC.

RESULTS The study included 58 patient cases from CPS-II (mean [SD] age, 71.0 [6.4] years; 16 [27.6%] women) and 71 patient cases from PLCO (mean [SD] age, 62.7 [4.8] years; 13 [18.3%] women). Two controls per patient case (n = 254) were selected through incidence density sampling, matched on age, sex, race/ethnicity, and time since mouthwash collection. Head and neck squamous cell cancer cases and controls were similar with respect to age, sex, and race. Patients in the case group were more often current tobacco smokers, tended to have greater alcohol consumption (among drinkers), and to be positive for oral carriage of papillomavirus-16. Overall microbiome composition was not associated with risk of HNSCC. Greater abundance of genera *Corynebacterium* (fold change [FC], 0.58; 95% confidence interval [CI], 0.41-0.80; q = .06) and *Kingella* (FC, 0.63; 95% CI, 0.46-0.86; q = .08) were associated with decreased risk of HNSCC, potentially owing to carcinogen metabolism capacity. These findings were consistent for both cohorts and by cohort follow-up time. The observed relationships tended to be stronger for larynx cancer and for individuals with a history of tobacco use.

CONCLUSIONS AND RELEVANCE This study demonstrates that greater oral abundance of commensal *Corynebacterium* and *Kingella* is associated with decreased risk of HNSCC, with potential implications for cancer prevention.

JAMA Oncol. doi:10.1001/jamaoncol.2017.4777
Published online January 11, 2018.

Supplemental content

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Richard B. Hayes, DDS, PhD, Division of Epidemiology, Department of Population Health, NYU Langone Health, New York University School of Medicine, 650 First Ave, New York, NY, 10016 (Richard.B.Hayes@nyumc.org).

More than 550 000 new cases of head and neck cancer (oral cavity, pharynx, and larynx) and 380 000 deaths related to the disease occur worldwide per year.[1] Head and neck squamous cell cancer (HNSCC) comprises approximately 85% of these cases, leading to considerable physical disfigurement, decreased quality of life, and an approximately 40% 5-year mortality rate. There is a critical need to develop effective new approaches for prevention of HNSCC, to complement efforts for smoking, alcohol, and human papillomavirus (HPV) control.

The human mouth hosts a diverse community of bacteria referred to as the oral microbiome.[2] Approximately 700 bacterial species have been identified in the human oral cavity to date. These bacteria are involved in a wide variety of functions, and many are important in maintaining oral health.[3] The role that bacteria play in the etiology and predisposition to cancer is of increasing interest, and several investigations have been carried out relating to the oral microbiome and its association with head and neck cancer.[4-10] Bacterial profiles in cancer cases were identified in these studies, but the investigations were limited to oral cavity cancers, were of a small sample size, and were often limited to a small number of bacterial species analyzed. Furthermore, direct analysis of bacterial communities in patients with tumors cannot distinguish whether observed microbiome profiles reflect secondary overgrowth of certain bacteria with preference to the cancer microenvironment, or whether they are associated with future development of cancer.

We conducted a prospective study nested in 2 large US cohorts to determine if the oral microbiome was associated with subsequent risk of HNSCC. We directly assessed the oral microbiota from high-throughput sequencing of the 16S ribosomal RNA (16S rRNA) gene in prediagnostic oral samples from 129 HNSCC cases and 254 controls in these cohorts and compared patient cases and controls for overall microbiota composition, and abundance of specific bacterial taxa. We also examined whether the associations of oral microbiome with subsequent HNSCC differed according to tobacco use, cancer site (oral cavity, pharynx, and larynx), or by time from microbiome measurement to cancer diagnosis.

Key Points

Question Is the prediagnostic oral microbiome associated with subsequent risk of head and neck squamous cell cancer (HNSCC)?

Findings In this prospective study in 2 large well-established cohorts, oral commensal bacterial genera *Corynebacterium* and *Kingella* were associated with decreased risk for HNSCC.

Meaning This first prospective study provides evidence that the commensal oral microbiome influences HNSCC risk, with potential implications for cancer prevention.

Methods

Study Population

Parent Cohorts

This study included 129 men and women diagnosed with HNSCC during follow-up (58 from the American Cancer Society Cancer Prevention Study Nutrition Cohort [CPS-II] and 71 from the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial [PLCO]) and 254 matched controls (114 from CPS-II and 140 from PLCO). All participants provided informed consent and all protocols were approved by the New York University School of Medicine institutional review board. The research cohorts comprised large US research populations with stored oral wash samples, comprehensive demographic and risk factor information, and prospective follow-up for cancer incidence. The characteristics of the 2 cohorts were comparable, and oral samples were collected in a similar fashion.

The CPS-II cohort[11] included more than 184 000 participants, aged 50 to 74 years, from 21 US states who completed a mailed baseline diet and lifestyle questionnaire in 1992. Since 1997, follow-up questionnaires have been sent to cohort members every other year to update exposure information and to ascertain incident cancer cases; a response rate of more than 80% has been achieved for each follow-up questionnaire. Incident cancers were verified through medical records, state cancer registries, or death certificates. During 2001 and 2002, oral wash samples were collected by mail from 70 004 cohort members. The present analysis includes HNSCC diagnosed between oral wash collection and June 2009.

The PLCO cohort[12] was a large population-based randomized trial that examined effects of screening on cancer-related mortality and secondary endpoints, in men and women aged 55 to 74 years, recruited between 1993 and 2001, and followed for cancer incidence. Participants were randomized to either a screening or control arm. Oral wash samples were collected in the control arm (n = 52 000). During 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 cases of HNSCC diagnosed between oral wash collection and December 2010.

Nested Case-Control Study Selection

Incident cases of HNSCC (n = 129) were identified during a mean 3.9 years of follow-up. Patients had histologically confirmed incident HNSCC, including cancers of the oral cavity (excluding salivary glands) according to the *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10)* (*ICD-10* codes: C02.0-C06.9), pharynx (excluding nasopharynx) (*ICD-10* codes: C01 [base of tongue], C09.0-C10.9, C12-C13.9), and larynx (*ICD-10* codes: C32.0-32.9). Patients selected for study had valid consent, prediagnostic oral wash samples, and had no prior medical history of cancer (except nonmelanoma skin cancer). Cohort nested controls were selected by incidence-density sampling among cohort members who had no cancer prior to selection, provided a valid consent, and provided an oral wash sample. Controls were matched to patient cases, in a ratio of 2 controls per patient case, by cohort, sex, race/ethnicity (white, black, or other/unknown), age at oral sample collection (within 1 year) and date of oral wash collection (within 30 days for

Table 1. Characteristics of 129 Incident Cases of Squamous Cell Head and Neck Cancer and 254 Participants in the Control Group in the 2 Cohorts[a]

Characteristic	CPS-II		PLCO	
	Cases (N = 58)	Controls (N = 114)	Cases (N = 71)	Controls (N = 140)
Age, mean (SD), y	71.0 (6.4)	71.0 (6.4)	62.7 (4.8)	63.2 (5.0)
Sex				
Male	42 (72.4)	83 (72.8)	58 (81.7)	115 (82.1)
Female	16 (27.6)	31 (27.2)	13 (18.3)	25 (17.9)
Race				
White	58 (100)	114 (100)	66 (93.0)	131 (93.6)
African American or other	0	0	5 (7.0)	9 (6.4)
Smoking status				
Never smoker	10 (17.2)	59 (51.8)	8 (11.3)	70 (50.0)
Former smoker	32 (55.2)	52 (45.6)	38 (53.5)	63 (45.0)
Current smoker[b]	16 (27.6)	3 (2.6)	25 (35.2)	7 (5.0)
Cigarettes per day in ever smokers, mean (SD)[c]	24.8 (12.1)	19.6 (11.9)	26.1 (12.6)	21.3 (13.8)
Alcohol consumption				
None	12 (20.7)	25 (21.9)	8 (11.3)	43 (30.7)
Drinker	36 (62.1)	68 (59.7)	50 (70.4)	89 (63.6)
Missing	10 (17.2)	21 (18.4)	13 (18.3)	8 (5.7)
Grams of ethanol per day in drinkers, mean (SD)[d]	17.1 (14.5)	12.5 (17.3)	38.3 (80.6)	11.3 (16.5)
Time to diagnosis, mean (SD), y[e]	3.0 (2.0)		4.3 (2.5)	
HPV-16 status[c]				
Positive	5 (8.6)	1 (0.9)	7 (9.9)	1 (0.7)
Negative	53 (91.4)	113 (99.1)	64 (90.1)	139 (99.3)
Cancer site				
Oral	19 (32.8)		22 (31.0)	
Pharynx	12 (20.7)		18 (25.4)	
Larynx	27 (46.5)		31 (43.6)	

Abbreviations: CPS-II, American Cancer Society Cancer Prevention Study II Nutrition Cohort; HPV-16, human papillomavirus 16; PLCO, Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial.

[a] Unless otherwise indicated, data are reported as number (percentage) of participants.

[b] $P < .001$ from χ^2 test for comparison between those in the case group and those in the control group in each cohort study.

[c] $P < .05$ from t test for comparison between those in the case group and those in the control group in each cohort study.

[d] $P < .05$ from t test for comparison between those in the case group and those in the control group in the PLCO cohort study.

[e] Time from oral sample collection to HNSCC diagnosis.

CPS-II and within 3 months for PLCO); 4 controls were excluded owing to inadequate materials, resulting in 4 patient cases having only 1 matched control.

Measurements

Demographic and Risk Factor Information

At enrollment and follow-up periods, participants in both cohorts completed a structured questionnaire that included questions about age, sex, race, smoking status, and alcohol consumption. For each participant, we used covariates from the questionnaire most closely preceding oral wash sample collection. Oral HPV status was available in the cohorts.[13] The relevant cohort data were harmonized and transferred to NYU School of Medicine.

Oral Wash Samples

Oral mouthwash samples were collected in both cohorts using similar protocols.[12,14] DNA is highly stable in a frozen state, which allowed us to use the frozen oral wash pellets archived in the CPS-II and PLCO cohorts. We have previously characterized HPV-16 status in these samples[13] and shown suitability of this sample type for oral bacterial microbiome measurement.[15]

Oral Microbiota Characterization Using 16S rRNA Gene Sequencing

Bacterial genomic DNA was extracted from thawed mouthwash samples using the PowerSoil DNA Isolation Kit (MO BIO). From extracted DNA, we amplified the 16S rRNA gene V3 to V4 regions using universal primers (347F 5′-GGAGGCAGCAG-TRRGGAAT'-3′ and 803R 5′-CTACCRGGGTATCTAATCC-3′), while incorporating adapters and sample-specific barcode sequences.[16] The 454 FLX Titanium pyrosequencing system (Roche) was used to sequence the resulting amplicons. The sequencing data will be submitted to dbGaP.

Upstream Sequence Analysis of Microbiome Data

Multiplexed sequences from pyrosequencing were demultiplexed based on sample-specific barcodes. Poor-quality sequences were excluded using QIIME.[17,18] Filtered sequence reads were clustered into operational taxonomic units (OTUs), and subsequently assigned to taxa using the Human Oral Microbiome Database predefined taxonomy map of reference sequences with 98% or greater identity.[19] From the 383 prediagnostic oral wash samples, we obtained 3 217 274 (mean [SD], 8400 [2992]) high-quality 16S rRNA gene sequence reads, with similar number of reads (library size) per participant in the case and control groups in both cohorts (eTable 1 in the Supplement).

Quality Control

Laboratory personnel were blinded to case or control status of participants, and matched pairs were processed side by side. Blinded positive quality control (QC) specimens were used across all sequencing batches. We previously reported that mul-

Table 2. Median Counts and Fold Changes (FCs) for the Association[a] Between Selected Taxa[b] and Risk of Head and Neck Squamous Cell Cancer (HNSCC) in a Case-Control Study Nested in 2 Cohorts

Taxon	Transformed Counts,[c] Median		FC (95% CI)	*P* Value	*q*[d]	Maximum Cook D[e]
	Case	Control				
Phylum *Actinobacteria*	585.70	477.91	1.21 (1.05-1.39)	.01	.05	16.1
Actinobacteria; Corynebacteriales	1.45	3.71	0.61 (0.46-0.82)	<.001	.01	0.01
Actinobacteria; Corynebacteriales; Corynebacteriaceae;	1.23	3.63	0.58 (0.42-0.80)	<.001	.02	0.01
Actinobacteria; Corynebacteriales; Corynebacteriaceae; Corynebacterium	1.23	3.62	0.58 (0.41-0.80)	.001	.06	0.01
Actinobacteria; Actinomycetales; Actinomycetaceae; Actinomyces;oral_taxon_170	0	0.94	1.94 (1.30-2.89)	.001	.06	19.6
Phylum *Proteobacteria*	453.84	533.99	0.91 (0.82-1.01)	.07	.15	0.36
Betaproteobacteria	27.22	81.28	0.77 (0.64-0.93)	.007	.08	0.24
Betaproteobacteria; Neisseriales	21.47	63.07	0.70 (0.53-0.91)	.008	.07	0.43
Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria; Sicca	0.86	12.40	0.48 (0.32-0.73)	<.001	.04	21.9
Betaproteobacteria; Neisseriales; Neisseriaceae; Kingella	3.26	5.59	0.63 (0.46-0.86)	.003	.08	1.88
Phylum *Bacteroidetes*	796.23	614.25	1.14 (0.98-1.31)	.09	.15	1.25
Bacteroidia;Bacteroidales;Prevotellaceae;Prevotella; Nanceiensis	3.30	6.39	0.51 (0.35-0.74)	<.001	.04	0.42
Flavobacteriia; Flavobacteriales; Flavobacteriaceae; Capnocytophaga; Leadbetteri	1.82	3.87	0.58 (0.41-0.82)	.002	.07	7.02
Phylum *Firmicutes*	5467.45	4646.84	1.09 (0.97-1.23)	.14	.18	3.59
Negativicutes; Selenomonadales; Veillonellaceae; Selenomonas; Sputigena	0.71	1.19	0.55 (0.38-0.80)	.002	.07	3.75
Phylum *Fusobacteria*	238.39	187.24	1.08 (0.94-1.25)	.29	.29	2.36

[a] The association between taxonomic abundance and HNSCC was detected by differential gene expression analysis based on the negative binomial distribution (DESeq) function, adjusting for cohorts, matching factors (age at oral sample collection [within 1 year), sex, and race [white, other]), smoking status (current, former, never), number of cigarettes per day for ever-smokers, alcohol drinking status (never, ever), grams ethanol per day for ever-drinkers, and human papillomavirus (HPV)-16 status.

[b] All taxa with a false discovery rate (FDR)-adjusted *q* <0.10 are included in the table.

[c] Sequence read counts were normalized by dividing raw counts by DESeq size factors.

[d] FDR-adjusted *P* value. The FDR adjustment was conducted at each taxonomic level (ie, class, genus) separately.

[e] Maximum Cook distance for DESeq model.

tiple repeated QC samples had good agreement (total 31 assays from 3 nonparticipant volunteers), with coefficient of variability of 0.45% to 6.22% for the Shannon-Wiener index for the PLCO and CPS-II sequencing batches.[15] In addition, we also assessed temporal stability of microbiome from oral wash samples, with repeat-collected QC samples, each week for 4 to 6 weeks from 5 volunteers.

Statistical Analysis

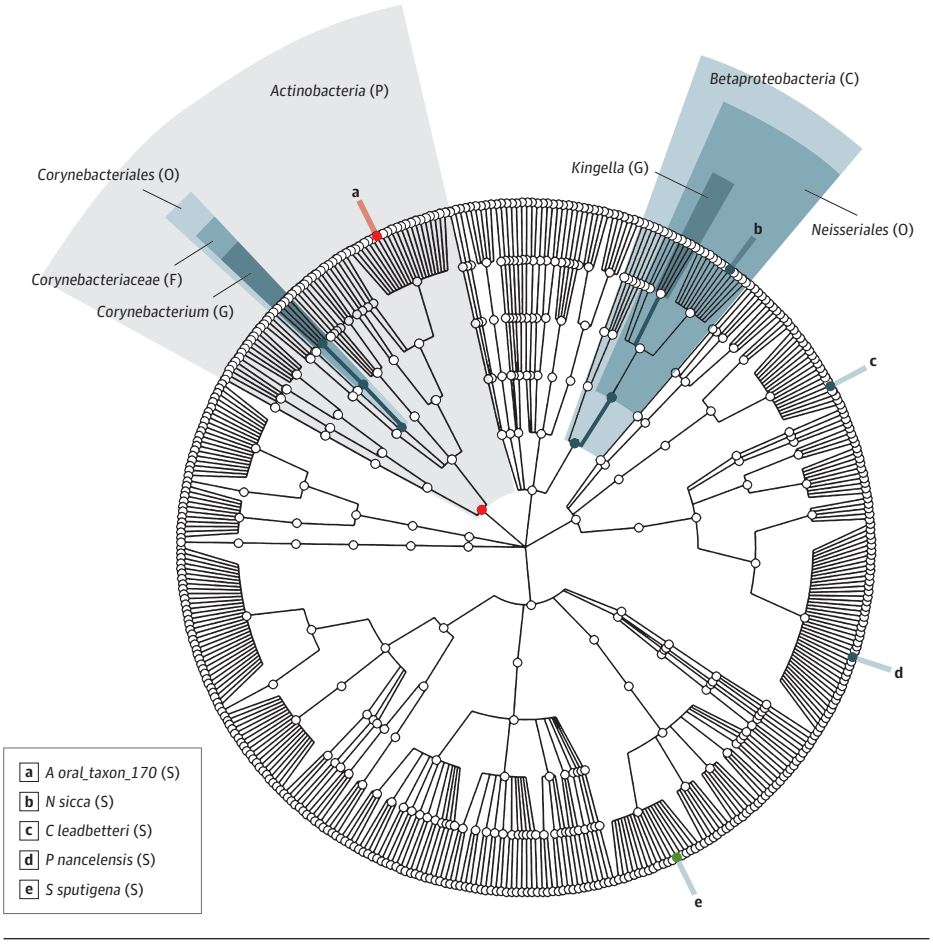
Permutational multivariate analysis of variance (PERMANOVA; adonis function, vegan package; R Foundation) was used to examine statistically whether overall bacterial community composition (β-diversity) differed by case or control status, adjusting for age at oral sample collection, sex, race, smoking status, number of cigarettes per day for ever-smokers, alcohol drinking status, grams ethanol per day for ever-drinkers, and oral HPV-16 status. We assessed within-participant diversity (α-diversity) using numbers of observed OTUs (richness) and the inverse Simpson index (evenness) in linear regression, adjusting for covariates, and based on 500 iterations of rarefied OTU tables with sequencing depth of 3000 (1 participant excluded).

We compared bacterial taxa between patient cases and controls, with correction for over-dispersion and library size, at each taxonomic level, using the differential gene expression analysis based on the negative binomial distribution (DESeq) in the DESeq2 package,[20,21] with variance and mean linked by local multivariable regression. We identified 10 phyla, 22 classes, 36 orders, 62 families, 127 genera, and 439 species before filtering. We analyzed taxa within 5 phyla (99.7% of reads, eTable 2 in the Supplement). In addition, we filtered to include only taxa with at least 2 sequences in at least 40 participants to remove low-count taxa for the phylum through species-level analysis. Our analyses included 5 phyla, 11 classes, 17 orders, 28 families, 50 genera, and 167 species. We used DESeq2 default outlier replacement and filtering of count outliers (maximum Cook distance, D > 25). To account for multiple comparisons at each taxonomic level, we considered a false discovery rate (FDR)-adjusted *P* value (*q* value) <.10, 2-sided, as significant.[22] Further, we conducted stratified analyses according to smoking status, alcohol consumption status, site (oral cavity, pharynx and larynx), study cohort and time from oral wash sample collection to diagnosis. Analyses were carried out using R statistical software (version 3.3.1, R Foundation)

Results

Cases of HNSCC and controls in the CPS-II and PLCO cohorts were similar with respect to age, sex, and race (**Table 1**). Cases

Figure 1: Oral Microbiota Associated With Head and Neck Squamous Cell Cancer (HNSCC) in 2 Cohorts



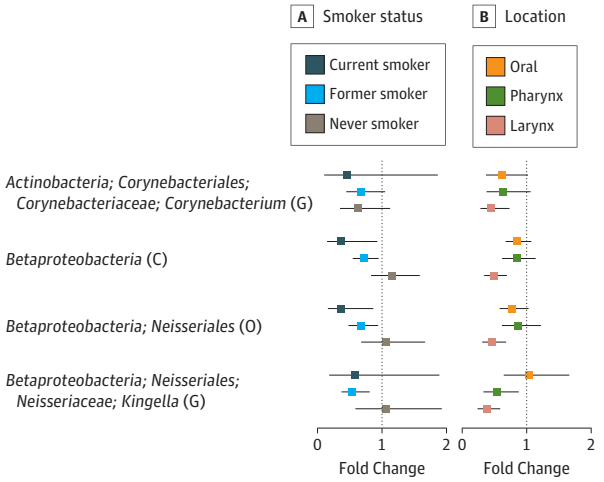
Red indicates taxa enriched in HNSCC case patients, and blue indicates taxa enriched in control participants. The association between taxonomic abundance and HNSCC was detected by the DESeq function, adjusting for cohort, matching factors (age at oral sample collection [within 1 year], gender, and race [white, other]), smoking status, number of cigarettes per day for ever-smokers, alcohol drinking status, grams ethanol per day for ever-drinkers, and human papillomavirus (HPV)-16 status. Figure was constructed using data presented in Table 2. (P), (C), (O), (F), (G), and (S) indicate bacterial taxa at the level of Phylum, Class, Order, Family, Genus, and Species, respectively. False discovery rate $q < 0.10$.

in both cohorts were more likely than controls to be current tobacco smokers, tended to have greater alcohol consumption (among drinkers), and to be positive for oral carriage of HPV-16.[13] Oral microbiome β-diversity (eFigure 1 in the Supplement), species richness, and evenness did not differ significantly for all those in the case group, or for oral cavity, pharynx, and larynx cancer, compared with controls (all $P > .05$).

In analysis of oral microbiome taxa, greater abundance of phylum *Actinobacteria* was associated with increased risk for HNSCC (fold change [FC], 1.21; $P = .01$; $q = .05$), although the maximum Cook D was large (>10), suggesting that this estimate is not robust (**Table 2**) (**Figure 1**). When examining lower-level taxa in phylum *Actinobacteria*, greater abundance of order *Corynebacteriales*, family *Corynebacteriaceae*, and genus *Corynebacterium* were associated with reduced risk of HNSCC (*Corynebacterium:* FC, 0.58; 95% CI, 0.41-0.80; $P = .001$; $q = .06$). In addition, class *Betaproteobacteria*, order *Neisseriales*, family *Neisseriaceae*, and genus *Kingella* in phylum *Proteobacteria* were significantly associated with lower risk of HNSCC (*Kingella:* FC, 0.63; 95% CI, 0.46-0.86; $P = .003$; $q = .08$). When we adjusted for *Corynebacterium* abundance in the *Kingella* multivariable model, and vice versa, both genera remained significant, suggesting independent associations with HNSCC. Three species, *Prevotella nanceiensis, Capnocytophaga leadbetteri* and *Selenomonas sputigena*, were also inversely related to HNSCC ($P < .05$, $q<.10$) (Table 2). None of 4 bacteria known to be associated with periodontal disease (*Porphyromonas gingivalis, Tannerella forsythia* and *Aggregatibacter actinomycetemcomitans*) and dental caries (*Streptococcus mutans*) were associated with risk of HNSCC, with respect to taxonomic abundance (eTable 3A in the Supplement) or overall bacterial carriage (presence/absence) (eTable 3B in the Supplement).

Associations observed for all HNSCC cases and controls (after adjustment for tobacco and alcohol use) tended to be more pronounced in current and former smokers than never smokers, including for *Neisseriales* and *Kingella*, but tests for interaction did not identify any significant differences by smoking status (**Figure 2**A) (eTable 4 in the Supplement). When we excluded current smokers from the analysis (41 patient cases, 10 controls), results were similar to those for all participants for *Corynebacterium* (FC, 0.76; 95% CI, 0.60-0.98) and *Kingella* (FC, 0.69; 95% CI, 0.51-0.92). Relationships of *Corynebacterium* and *Kingella* with HNSCC were largely consistent by age, sex, alcohol use, and by membership in the CPS-II and PLCO cohorts (all tests for heterogeneity, $P > .05$). Risks were also unchanged after exclusion of participants positive for HPV-16 (12 patient cases, 2 controls). The observed as-

Figure 2: Fold Changes (95% CIs) for the Association Between Selected Taxa and Risk of Head and Neck Squamous Cell Cancer (HNSCC) in 2 Cohorts, According to Smoking Status and Site of Cancer (Oral Cavity, Pharynx, and Larynx)



The association between taxonomic abundance and HNSCC was detected by the DESeq function adjusting for covariates. Selected taxa are shown with $q < .10$. Figure was constructed using data presented in eTables 4 and 6 in the Supplement. P, C, O, F, G, and S indicate bacterial taxa on phylum, class, order, family, genus, and species level, respectively.

sociations did not differ by years of follow-up between sample collection and diagnosis (eTable 5 in the Supplement, all tests for heterogeneity $P > .05$), indicating that risks observed in our study are unlikely owing to effects of undiagnosed HNSCC.

Greater abundance of genera *Corynebacterium* (order *Corynebacteriales*), *Kingella* (order *Neisseriales*), *Neisseria* (order *Neisseriales*), *Abiotrophia* (order *Lactobacillales*), *Capnocytophaga* (order *Flavobacteriales*) and species *Kingella dentificans* and *Streptococcus sanguinis* were associated with reduced risk for larynx cancer, (all $P < .01$, $q < .10$) (Figure 2B) (eTable 6 in the Supplement). No bacterial genera were associated with oral cavity or pharynx cancer ($q > .10$). At the species level, *Actinomyces oris* and *Veillonella denticariosi* were associated with reduced risk of pharynx cancer (eTable 6 in the Supplement). *Parvimonas micra* and *Neisseria sicca* were associated with reduced risk of oral cavity cancer, whereas an unnamed *Actinomyces* (oral-taxon_170) was associated with increased risk at this site (the latter with maximum Cook D>10) (eTable 6 in the Supplement).

Discussion

In this first prospective study of the oral microbiome and HNSCC, we found that greater abundance of the commensal bacterial genera, *Corynebacterium* and *Kingella*, was associated with reduced risk of HNSCC. The findings were consistent in both cohorts and were unrelated to the time period of microbiome measurement, indicating that risk preceded disease occurrence. We found that these associations tended to be strongest in those with a history of tobacco use. Also these genera and a broader spectrum of bacterial commensals were most strongly associated with larynx cancer. These results are consistent with a role of the healthy oral microbiome in HNSCC prevention.

Several studies have reported on the prevalence of oral bacteria in oral cancer. Culture-based studies revealed increases in salivary bacterial counts in oral cancer patients,[4] and bacterial differentials in oral cancer tissue compared with normal tissue of the same patients[5] or with tissue from healthy patients.[6] Studies based on DNA sequencing have also characterized bacterial colonization of oral tumors.[7-10] Although these studies identified certain microbial patterns present in oral cancer, they do not directly address the relationship of the oral microbiome to development of HNSCC. To address this gap, we prospectively examined risks for HNSCC, including oral cavity, pharynx, and larynx cancer, while taking into account risks associated with smoking, alcohol consumption, and oral carriage of HPV.[23-25]

Our findings that reduced risk of HNSCC is associated with greater abundance of *Corynebacterium* and *Kingella* is biologically plausible. In this study population, we recently reported that oral *Corynebacterium* and *Kingella* are functionally related to xenobiotic biodegradation and metabolism pathways, including capacity to metabolize several toxicants found in cigarette smoke.[15] Our findings of stronger inverse association of these oral taxa with HNSCC in current and former smokers, suggest that the preventive effect of those taxa may be more pronounced in the high-carcinogen oral environment in smokers. In HNSCC, tobacco use tends to have the strongest relationship with larynx cancer, with respect to amount smoked[26] and level of continued risk years after smoking cessation.[27] Although our study tends to show the strongest relationships of the oral microbiome for smokers and for larynx cancer, these findings require independent verification.

Others have speculated that oral bacteria could influence risk of HNSCC in consumers of alcohol because several commensal bacteria can metabolize ethanol to carcinogenic acetaldehyde.[28] In a recent analysis in the PLCO Trial,[24] the proportion of head and neck cancer cases attributed to alcohol alone was 14.7%, and to tobacco and alcohol combined was 0.9%, indicating that alcohol-related risks in PLCO (and likely in CPS-II) are moderate, limiting the ability in our study to assess bacterial alcohol metabolism interrelationships with HNSCC.

We recently reported[13] a 22-fold increased risk in these cohorts for oropharyngeal cancer related to oral HPV-16 carriage; here we report that *A oris* and *V denticariosi* are potentially associated with a modestly reduced risk of pharynx cancer and oropharynx cancer, after adjustment for HPV-16 status or exclusion of HPV-positive cases. Human papillomavirus is clearly the dominant microbial factor for cancer at these cancer sites in HPV carriers; however, these oral bacteria may play a protective role in noncarriers.

Because periodontal disease,[29] tooth loss[30] and infrequent tooth brushing[31] are linked with increased risk of HNSCC in epidemiologic studies, it is plausible that oral bacteria associated with these conditions may be risk factors for HNSCC. We found no relationship with HNSCC for selected bacterial taxa associated with periodontal disease and dental car-

ies. It is becoming evident, however, that periodontitis and dental caries are caused by mixed-species communities rather than by individual pathogens working in isolation.[32] In this regard, greater abundance of *Corynebacterium* and *Kingella*, as well as *Neisseria*, *Abiotrophia, Actinomyces, Veillonella*, and *Capnocytophaga*, have been related to good oral health rather than to oral disease.[33,34] Thus, commensal bacteria associated with reduced risk for HNSCC in our study may characterize a healthy oral microbiome that prevents disease development. This hypothesis is also consistent with our finding of no broad differentials in bacterial α- or β-diversity between participants in the case group and those in the control group in our study.

Limitations

There are certain limitations to our study. From this observational study, causality may not be firmly established, although the prospective design provides a rigorous test of the hypothesis that oral bacteria might play a role in development of HNSCC. Also, we were able to access nested case-control oral wash samples from a repository of more than 120 000 volunteers in PLCO and CPS-II to carry out this prospective study; however, it would have been of additional value to also study samples isolated from other oral sites (eg, tongue, other oral mucosa, dental plaque), but no such resources were available. A limitation of our study is measurement of only a single oral sample. Microbiome measures at multiple time points could result in more precise exposure estimates, although the abundance of *Corynebacterium, Kingella*, and other core members of the oral microbiome are fairly stable over time at the genus level.[35,36] Furthermore, most participants in our study were white, potentially limiting generalizability of our research for other races and ethnicities.

Conclusions

This survey of oral bacterial microbiota found that increased abundance of *Corynebacterium*, *Kingella*, and potentially other selected genera and species was associated with risk of HNSCC. The study provides the first comprehensive evidence that the oral microbiome may be associated with subsequent risk of HNSCC, with the strongest links for larynx cancer and those with a history of tobacco use. Maintenance of a healthy oral microbiome is essential to oral health[37]; our findings may have implications for HNSCC prevention in conjunction with other control measures.

ARTICLE INFORMATION

Accepted for Publication: October 25, 2017

Published Online: January 11, 2018. doi:10.1001/jamaoncol.2017.4777

Author Affiliations: Department of Population Health, New York University School of Medicine, New York (Hayes, Ahn, Fan, Peters); NYU Perlmutter Cancer Center, New York University School of Medicine, New York (Hayes, Ahn, Pei); Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China (Ma); Department of Medicine, New York University School of Medicine, New York (Yang); Department of Epidemiology and Population Health, Albert Einstein College of Medicine, New York, New York (Agalliu, Burk); Departments of Pediatrics; Microbiology & Immunology; Obstetrics, Gynecology & Women's Health, Albert Einstein College of Medicine, New York, New York (Burk); Department of Surgery, Memorial Sloan Kettering Cancer Center, New York (Ganly); Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland (Purdue, Freedman); Epidemiology Research Program, American Cancer Society, Atlanta, Georgia (Gapstur); Department of Pathology, New York University School of Medicine, New York (Pei); Department of Veterans Affairs New York Harbor Healthcare System, New York (Pei).

Author Contributions: Drs. Hayes and Ahn had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. They both contributed equally to the work.
Study concept and design: Hayes, Ahn, Agalliu, Burk, Pei.
Acquisition, analysis, or interpretation of data: Hayes, Ahn, Fan, Peters, Yang, Burk, Purdue, Freedman, Gapstur, Pei.
Drafting of the manuscript: Hayes, Ahn, Fan, Pei.
Critical revision of the manuscript for important intellectual content: Hayes, Ahn, Peters, Yang, Agalliu, Burk, Purdue, Freedman, Gapstur, Pei.
Statistical analysis: Hayes, Ahn, Fan, Peters.
Obtained funding: Hayes, Ahn, Burk, Pei.
Administrative, technical, or material support: Yang, Burk, Purdue, Freedman, Pei.
Study supervision: Hayes, Burk, Pei.

Conflict of Interest Disclosures: None reported.

Funding/Support: This work was supported primarily by Public Health Service grant R01CA159036 and in part by the NYU Perlmutter Cancer Center Grant (P30CA016087), both of which were funded by the National Cancer Institute (NCI), National Institutes of Health. Additional support was from grants U01CA182370 and R01CA164964 from the NCI to NYU Langone Health and by grants R21CA152785 and P30CA013330 from the NCI to the Albert Einstein College of Medicine. The American Cancer Society supports the follow-up and maintenance of the Cancer Prevention Studies, which contributed data to this project. The Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial was funded through a contract mechanism administered by the NCI's Division of Cancer Prevention (DCP). The major DCP PLCO contracts were for 10 screening centers, a central laboratory, and a coordinating center. In addition, there was a separate DCP contract for a data management and analysis center. Additional contracts were provided by the intramural NCI Division of Cancer Epidemiology and Genetics to house and maintain the PLCO biorepository. ZP is a Staff Physician at the Department of Veterans Affairs New York Harbor Healthcare System.

Role of the Funder/Sponsor: All funders supported the collection, management, analysis, and interpretation of the data but had no role in the design and conduct of the study; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Additional Contributions: We thank the PLCO Cancer Screening Trial investigators and the staff from Information Management Services Inc and Westat Inc for database support, sample selection, and study management. We also acknowledge the contribution to the Cancer Prevention Studies from central cancer registries, which are supported through the Centers for Disease Control and Prevention's National Program of Cancer Registries and cancer registries supported by the National Cancer Institute's Surveillance Epidemiology and End Results Program. Most importantly, we thank the participants of the American Cancer Society Cancer Prevention Study II Nutrition Cohort and the PLCO trial for their contributions that made this study possible.

Acknowledgement: The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the US Department of Veterans Affairs, or the United States Government.

REFERENCES

1. IARC. *Cancer Fact Sheets: Estimated Incidence, Mortality And Prevalence (5 Years)*. Worldwide; Lyon, France. 2012.

2. Dewhirst FE, Chen T, Izard J, et al. The human oral microbiome. *J Bacteriol*. 2010;192(19):5002-5017.

3. Belda-Ferre P, Alcaraz LD, Cabrera-Rubio R, et al. The oral metagenome in health and disease. *ISME J*. 2012;6(1):46-56.

4. Mager DL, Haffajee AD, Devlin PM, Norris CM, Posner MR, Goodson JM. The salivary microbiota as a diagnostic indicator of oral cancer: a descriptive, non-randomized study of cancer-free and oral squamous cell carcinoma subjects. *J Transl Med*. 2005;3:27.

5. Nagy KN, Sonkodi I, Szöke I, Nagy E, Newman HN. The microflora associated with human oral carcinomas. *Oral Oncol*. 1998;34(4):304-308.

6. Bolz J, Dosá E, Schubert J, Eckert AW. Bacterial colonization of microbial biofilms in oral squamous cell carcinoma. *Clin Oral Investig*. 2014;18(2):409-414.

7. Pushalkar S, Ji X, Li Y, et al. Comparison of oral microbiota in tumor and non-tumor tissues of patients with oral squamous cell carcinoma. *BMC Microbiol*. 2012;12:144.

8. Hooper SJ, Crean SJ, Fardy MJ, et al. A molecular analysis of the bacteria present within oral squamous cell carcinoma. *J Med Microbiol*. 2007;56(Pt 12):1651-1659.

9. Schmidt BL, Kuczynski J, Bhattacharya A, et al. Changes in abundance of oral microbiota associated with oral cancer. *PLoS One*. 2014;9(6):e98741.

10. Guerrero-Preston R, Godoy-Vitorino F, Jedlicka A, et al. 16S rRNA amplicon sequencing identifies microbiota associated with oral cancer, human papilloma virus infection and surgical treatment. *Oncotarget*. 2016;7(32):51320-51334.

11. 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(2):500-511.

12. Hayes RB, Reding D, Kopp W, et al; Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Project Team. Etiologic and early marker studies in the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial. *Control Clin Trials*. 2000;21(6)(suppl):349S-355S.

13. Agalliu I, Gapstur S, Chen Z, et al. Associations of oral α-, β-, and γ-human papillomavirus types with risk of incident head and neck cancer [published online January 21, 2017]. *JAMA Oncol*. doi:10.1001/jamaoncol.2015.5504

14. Feigelson HS, Rodriguez C, Robertson AS, et al. Determinants of DNA yield and quality from buccal cell samples collected with mouthwash. *Cancer Epidemiol Biomarkers Prev*. 2001;10(9):1005-1008.

15. 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(10):2435-2446. doi:10.1038/ismej.2016.37

16. Nossa CW, Oberdorf WE, Yang L, et al. Design of 16S rRNA gene primers for 454 pyrosequencing of the human foregut microbiome. *World J Gastroenterol*. 2010;16(33):4135-4144.

17. Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010;7(5):335-336.

18. Haas BJ, Gevers D, Earl AM, et al; Human Microbiome Consortium. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res*. 2011;21(3):494-504.

19. Chen T, Yu WH, Izard J, Baranova OV, Lakshmanan A, Dewhirst FE. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database*. 2010;2010:baq013.

20. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.

21. McMurdie PJ, Holmes S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol*. 2014;10(4):e1003531.

22. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc B Met*. 1995;57(1):289-300.

23. Hashibe M, Brennan P, Chuang SC, et al. Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *Cancer Epidemiol Biomarkers Prev*. 2009;18(2):541-550.

24. Hashibe M, Hunt J, Wei M, Buys S, Gren L, Lee YC. Tobacco, alcohol, body mass index, physical activity, and the risk of head and neck cancer in the prostate, lung, colorectal, and ovarian (PLCO) cohort. *Head Neck*. 2013;35(7):914-922.

25. Forte T, Niu J, Lockwood GA, Bryant HE. Incidence trends in head and neck cancers and human papillomavirus (HPV)-associated oropharyngeal cancer in Canada, 1992-2009. *Cancer Causes Control*. 2012;23(8):1343-1348.

26. Lubin JH, Gaudet MM, Olshan AF, et al. Body mass index, cigarette smoking, and alcohol consumption and cancers of the oral cavity, pharynx, and larynx: modeling odds ratios in pooled case-control data. *Am J Epidemiol*. 2010;171(12):1250-1261.

27. Marron M, Boffetta P, Zhang ZF, et al. Cessation of alcohol drinking, tobacco smoking and the reversal of head and neck cancer risk. *Int J Epidemiol*. 2010;39(1):182-196.

28. Moritani K, Takeshita T, Shibata Y, et al. Acetaldehyde production by major oral microbes. *Oral Dis*. 2015;21(6):748-754.

29. Zeng XT, Deng AP, Li C, Xia LY, Niu YM, Leng WD. Periodontal disease and risk of head and neck cancer: a meta-analysis of observational studies. *PLoS One*. 2013;8(10):e79017.

30. Zeng XT, Luo W, Huang W, Wang Q, Guo Y, Leng WD. Tooth loss and head and neck cancer: a meta-analysis of observational studies. *PLoS One*. 2013;8(11):e79074.

31. Zeng XT, Leng WD, Zhang C, Liu J, Cao SY, Huang W. Meta-analysis on the association between toothbrushing and head and neck cancer. *Oral Oncol*. 2015;51(5):446-451.

32. Jakubovics NS, Yassin SA, Rickard AH. Community interactions of oral streptococci. *Adv Appl Microbiol*. 2014;87:43-110.

33. Kumar PS, Griffen AL, Moeschberger ML, Leys EJ. Identification of candidate periodontal pathogens and beneficial species by quantitative 16S clonal analysis. *J Clin Microbiol*. 2005;43(8):3944-3955.

34. Meuric V, Le Gall-David S, Boyer E, et al. Signature of microbial dysbiosis in periodontitis. *Appl Environ Microbiol*. 2017;83(14):e00462-17.

35. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial community variation in human body habitats across space and time. *Science*. 2009;326(5960):1694-1697.

36. Utter DR, Mark Welch JL, Borisy GG. Individuality, stability, and variability of the plaque microbiome. *Front Microbiol*. 2016;7:564.

37. Kilian M, Chapple IL, Hannig M, et al. The oral microbiome - an update for oral healthcare professionals. *Br Dent J*. 2016;221(10):657-666.