Long term persistence of oral HPV over 7 years of follow-up

Gypsyamber D’Souza PhD1,7, Gwendolyn Clemens MS2, Howard D. Strickler MD3, Dorothy J. Wiley4 PhD, Tanya Troy MA1, Linda Struijk PhD5, Maura Gillison MD PhD,5 Carole Fakhry MD MPH1,7

1 Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore MD
2 Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore MD
3 Department of Epidemiology & Public Health, Albert Einstein College of Medicine, Bronx, NY
4 University of California, Los Angeles School of Nursing, Los Angeles, California
5 DDL Diagnostic Laboratory, Rijswijk, the Netherlands
5 Department of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX
7 Department of Otolaryngology - Head and Neck Surgery, Johns Hopkins Hospital, Baltimore MD

Corresponding Authors:

Carole Fakhry and Gypsyamber D’Souza

Johns Hopkins University

cfakhry@jhmi.edu and Gdsouza2@jhu.edu

410-955-0289

© The Author(s) 2020. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
Abstract

Background: HPV-related oropharyngeal cancer (HPV-OPC) incidence is increasing but the natural history of the precursor, oral HPV, has not been well described.

Methods: This observational cohort study of people living with HIV and at-risk HIV uninfected people evaluated participants semi-annually using 30-second oral rinse/gargle specimens over 7 years. Initially, 447 subjects were followed for four years as part of the POPS Study, and a subset of 128 showing persistent infections at the last POPS visit had an additional visit, as part of MOUTH Study, on average 2.5 years later. Extracted DNA from oral rinse/gargle specimens were amplified using PCR and type-specification of 13 oncogenic HPV types. Risk factors for oncogenic oral HPV clearance were evaluated using Cox models.

Results: The majority of oncogenic oral HPV infections cleared quickly with median time to clearance of 1.4 years [IQR=0.5-3.9]. After 7 years of follow-up, 97% of incident and 71% of prevalent infections had cleared. Lower HPV16 viral load was statistically significantly associated with clearance (Per 10-fold decrease in copy number: adjusted hazard ratio [HR]=2.51, 95% confidence interval [CI]=1.20-5.26, p=.01) Adjusted analyses showed oncogenic oral HPV clearance was lower among prevalent than incident detected infections (aHR=0.44, 95%CI=0.35-0.55), among men than women (aHR=0.74, 95%CI=0.60-0.91), for older participants (aHR per 10 years increasing age=0.81, 95%CI=0.74-0.89), and among people living with HIV (aHR=0.76, 95% CI=0.60-0.95). One participant who had oral HPV16 consistently detected at 10 study visits over 4.5 years was subsequently diagnosed with HPV-OPC.

Conclusions: This prospective study of oncogenic oral HPV infection is the longest and largest quantification of oral HPV16 infections to date.
Persistent oncogenic oral human papillomavirus (HPV) infection is a risk factor for oropharyngeal squamous cell carcinoma, >70% of which are HPV-related (HPV-OPC),\(^1\) yet, risk factors for oncogenic HPV persistence remain poorly understood. Persistence of cervical HPV serves as a strong biomarker for cervical cancer risk.\(^7,8\) As such, oncogenic cervical HPV detection is now incorporated as an option for screening in the United States.\(^9\) This model was implemented after understanding the long-term natural history of cervical HPV, cytologic change, and risk of malignancy. Whether such a paradigm applies to oral HPV and oropharyngeal cancer is unknown. Oral HPV natural history data are limited; three studies report variable follow-up (≤4 years) and limited number of oral HPV16 infections (range: 10-48 infections)\(^2,3,5,10\).

Initial natural history studies suggest that factors associated with acquisition and persistence of oral HPV infection include male sex, cigarette smoking, and immunosuppression.\(^2,4\) For example, people living with HIV (PLWH) have 2-3 times higher oral HPV prevalence than HIV-uninfected individuals and the risk of both incident oral HPV detection and its persistence in PLWH increases with diminishing CD4 cell count after controlling for sexual behavior and other relevant covariates.\(^4\) Short-term natural history studies have shown that most people clear oral HPV infections within 1-2 years\(^2-6\), however, there is a paucity of long-term natural history studies. We therefore aimed to describe for the first time the natural history of persistent oral HPV over 7 years.

**Methods**

**Study Population**

The study population was a subcohort comprised of 1833 participants enrolled from two prospective, multi-center cohort studies of people living with HIV (PLWH) and at-risk HIV-uninfected individuals in the United States: the Multicenter AIDS Cohort Study (MACS) and the Women’s Interagency HIV Study (WIHS).\(^11\) Participants were followed prospectively every six months with oral
rinse sample collection for four years between October 2009 and March 2016 as part of the Persistent Oral Papillomavirus Study (POPS; as previously described). After completion of POPS, the 128 participants who had persistent oncogenic oral HPV infection at the last POPS visit were enrolled in an additional follow-up visit as part of the Men and Women Understanding Throat HPV study (MOUTH; ClinicalTrial.gov study number NCT03644563). As in POPS, at the baseline MOUTH study visit participants provided an oral rinse specimen for HPV testing, completed a questionnaire and had CD4 cell count and HIV viral load tested. There was a median of 2.5 years (IQR 1.9-3.0) between the last POPS visit and the additional MOUTH visit, and the 128 MOUTH participants shared similar characteristics as POPS participants as a whole (Table 1). All participants completed informed consent and the study was IRB approved at each study site.

The current analysis includes 447 subjects from POPS who had at least one oncogenic oral HPV infection detected, including 128 subjects who subsequently had an additional visit as part of MOUTH study (between October 2017- September 2018; Figure 1). At each visit participants completed a computer assisted self-interview (CASI) questionnaire which included questions on behaviors, medication use and clinical outcomes, had a physical examination and CD4 cell count and HIV viral load tested.

**Oral rinse sample collection, processing, and oral HPV detection**

Oral rinse samples were collected using 10mL saline or Scope and a 30-second oral rinse and gargle as previously described. Samples were stored at 4°C until processed. Oral rinse samples were tested in the laboratory of Dr. Maura Gillison at The Ohio State University (for the POPS study) and in the DDL Diagnostic Laboratory (https://www.ddl.nl, for the MOUTH study). Oral HPV DNA detection involved PCR with the PGMY09/11 primer system in the Gillison Lab and with the SPF10 primer system in the DDL lab (version 1 system Labo Biomedical Products, Rijswijk, The Netherlands). HPV type specification was conducted using Roche linear array in the Gillison lab and the SPF10 DEIA/LIPA system in the DDL lab. Oncogenic oral HPV types were defined identically for data from both labs, to
include types 16/18/31/33/35/39/45/51/52/56/58/59/66. Similar methods were used in both laboratories; 100 samples were tested in both and yielded 97% concordance. HPV16 viral load was tested for all samples with any HPV detected using either linear array band intensity or TaqMan quantitative PCR (LightCycler® 480 Probes Master kit on Roche LC480 II instrument).

**Statistical analysis**

Oral HPV prevalence and incidence the POPS study have been reported elsewhere. The current analyses included 676 oncogenic oral HPV among 447 individuals, who had at least one follow-up visit after the infection was detected. All analyses were type-specific, i.e. following the same HPV type in the same person until it cleared. Infections detected at the first study visit were defined as prevalent, and infections detected for the first time during follow-up were defined as incident. Incident infections detected only at the participants’ last visit were excluded from analysis of clearance and risk factors for clearance.

Time-to-clearance was calculated using Kaplan-Meier for each type-specific oncogenic infection as the time from first detection until cleared, where the primary definition for clearance was defined as two consecutive negative tests. For infections considered cleared, the visit of clearance was the first of the two consecutive negative visits. As the MOUTH baseline visit was several years after the last sample collection, infections still detected at the MOUTH baseline visit were considered persistent while those with a single negative result at the MOUTH baseline visit were considered cleared for this analysis. Risk factors for clearance were explored using Cox proportional hazard models and clustered by ID to account for multiple infections within the same person. The multivariable model considered all variables with p-values<0.05 in unadjusted Cox proportional hazard models and those known to be relevant based on previous literature (both current smoking and packyears of tobacco were considered in multivariate models based on their importance in priori literature but as both showed no association in adjusted models they were not retained in the final model). All significance tests were two-sided. Variables were removed one at a time to develop the final multivariable model.
HPV16 viral load was tested at the visit of first HPV16 detection. Number of HPV16 copies in each oral rinse sample was standardized to the number of human cells in the sample to calculate the number of HPV16 copies per cell as a measure of HPV16 viral load standardized across samples. Effect of HPV16 viral load on HPV16 persistence was evaluated using Cox proportional hazard models, and adjusted for sex, age, and infection type. The number of HPV16 copies was categorized by 10-fold changes per cell as: at least one HPV16 copy per every human cell, 0.1 to 0.99 copies per cell, 0.01-0.099 copies per cell, 0.001-0.0099 copies per cell, and less than 0.001 copies per cell.

Oral HPV detection and viral load were also considered in a graphic where intensity of the line blot results were plotted by color to visualize the pattern and strength (viral load) of infection. For viewability, this figure was restricted to 99 (of the 110) people with HPV16 infection detected and at least 4 oral rinses collected in the study (results were similar when examined for all oncogenic oral HPV infections). For POPS samples, signal intensity of Roche HPV Linear Array [LA], an established surrogate for HPV DNA viral load (signal strength)\(^{16-18}\) available at every study visit, was analyzed on a semi-quantitative 4 point scale: strong[3,4], medium[1,2], weak[-1], and negative. For MOUTH samples, quantitative HPV16 viral load was measured by qPCR (among those with HPV16 detected by LA, otherwise assumes to be 0) and categorized on the same semi-quantitative 4-point scale (defined by number of HPV16 copies/cell: \(\geq 1\), \(<1\) but \(>0\), target detected but no quantification, 0).

**Results**

**Participant Characteristics**

Median follow-up for all participants in the current analysis was 4.2 years (IQR 2.6-6.0) for POPS and MOUTH combined, including 7 years (IQR 6.3-7.6) median follow-up for those in MOUTH. Participants were 59.1% men, 50.3% Black and 33.1% White, and median age was 50 years (Table 1). Current smoking (45.1%) and alcohol consumption (70.0%) were common and 73.4% of participants were living with HIV (of whom 81.1% were currently using highly active antiretroviral therapy.
HAART). The median CD4 cell count was 522 cells/mm³, and HIV viral load was undetectable in 53.0%.

Time to clearance of oncogenic oral HPV infections

HPV16 was the most common oncogenic infection detected and represented 15.5% of all oncogenic infections (105 of 676). Other common oncogenic types detected included HPV33, 35, 45, 52, 59, 66 (8-9% of infections each). The majority of oncogenic oral HPV infections cleared within 1.4 years [IQR=0.5-3.9]. However, after 7 years of follow-up, 5.5% (n=37) of oncogenic oral HPV infections were still persistently detected (Figure 2); these 37 infections were among 31 individuals, as some individuals had multiple persistent infections. Median time to clearance was more rapid for incident (0.7 years [IQR=0.5-2.5]) than prevalent (2.4 years [IQR=0.7-7.5]) infections (Figure 3A). Most (70%) incident infections cleared by 2 years, with 93% and 97% clearance by 5 and 7 years, suggests that long-term persistence of incident infections was rare. In contrast, only 47% of prevalent infections cleared by 2 years, and clearance was 65% and 71% at 5 and 7 years.

Among oral HPV16 infections detected, clearance was 42%, 51%, 68%, and 76% at 1, 2, 5 and 7 years respectively (Figure 3B). This represents one year oral HPV16 persistence of 58% (95%CI=48%-67%). Among those with a persistent infection at one year, 55% (95%CI=40%-68%) remained persistent at five years. Five year oral HPV16 persistence was 32% (95%CI=22%-42%) overall, with 50% of prevalent infections and 16% of incident infections persisting at 5 years (p<.001).

Risk factors for oral HPV clearance

Univariate factors associated with clearance of oncogenic oral HPV infection were explored (Figure 3 and Table 2). HPV16 infections were statistically significantly less likely to clear than other (non-16) oncogenic HPV types (5-year persistence 32% vs 18%, HR=1.48, 95%CI=1.03-2.14, p=.02,
Figure 3B). Consistent with prior shorter natural history studies, infections were statistically significantly less likely to clear among men than women (5-year persistence 25% vs 15%, HR=0.63, 95%CI=0.51-0.79, p=.002, Figure 3C), and among older than younger individuals (5-year persistence in ≥60 vs <40; 41% vs 11%, HR=0.42, 95%CI=0.30-0.61, p<.001, Figure 3D). Current tobacco, alcohol, marijuana use, HIV, history of tonsillectomy and oral hygiene did not influence oral HPV clearance (p range .60-.72; Table 2).

Time to clearance curves were similar in magnitude when analyses were restricted to HPV16 or prevalent infections, however the effect of sex was not statistically different (Supplementary Figures 1-2). When examining incident infections only, sex, age and infection type (HPV16 vs. non16) were not statistically significant predictors of clearance (Supplementary Figure 3), perhaps because clearance was rapid for most incident infections, among both men and women and all age groups (each median<12 months).

In multivariate analysis (Table 2), oncogenic oral HPV clearance was statistically significantly lower among prevalent than incident detected infections (aHR=0.44, 95%CI=0.35-0.55), among men than women (aHR=0.74, 95%CI=0.60-0.91), for older participants (aHR per 10 years increasing age=0.81, 95%CI=0.74-0.89), and among people living with HIV (aHR=0.76, 95% CI=0.60-0.95). However, type of oncogenic HPV infection (HPV16 vs other oncogenic types) was no longer statistically significant.

Viral load and pattern of oral HPV16 infection

Given the importance of HPV type 16 specifically, we next considered the signal strength (viral load) and pattern of HPV16 infections across visits (Figure 4). Among the 99 infections included in this analysis, nearly half of infections (n=53, 54%) were detected at only 1 or 2 visits. There were 23 oral HPV16 infections that did not clear during the study (i.e. detected at the last study visit; Figure 4A), including 11 that were detected at ≥7 visits (representing at least 3.5 years of persistent infection). Oral HPV16 clearance was 51%, 55%, 64%, 68%, 73%, 76% at 2, 3, 4, 5, 6, and 7 years respectively. The
longest oral HPV16 infection observed was a prevalent infection that was persistently detected at each of 13 visits for over 7.5 years (see row 1 Figure 4).

Infections with a lower HPV16 viral load were statistically significantly more likely to clear than those with higher viral load. Each 10-fold decrease in copy number was associated with a three-fold increase in odds of clearance (HR=3.3, 95% CI= 1.7-6.3). After adjusting for sex, age, and infection type, lower HPV16 copy number remained a statistically significant predictor of clearance (aHR=2.5 95% CI=1.2-5.2) Among 53 oral HPV16 infections detected at high intensity level at any point during the study, only 54.7% were observed to clear (Figure 4). Clearance was statistically significantly lower among the 53 participants with a high signal strength for HPV16 than among the 52 participants with only medium/low strength results (5 year clearance 39.4% vs 95.3%, p<.001).

Incident case of HPV-OPC

Among those with oral HPV16 persistence at 5 years, there was one histologically confirmed incident case of HPV-OPC (4.5%, 95% CI= 0.1%-23% [1/22]) detected as part of clinical care, not by this observational study (Figure 4). HPV tumor status was determined by p16 immunohistochemistry and oncogenic DNA in-situ hybridization (PATHO-GENE HPV screening probe). This participant had oral HPV16 consistently detected at each of the 10 study visits (Figure 4A – denoted by *) spanning over 4.5 years. The strength of oral HPV16 detection during the first two years of detection was intermittently medium/low (4 times), but then persistently high intensity the latter two years (4 times) preceding the diagnosis of tonsillar cancer AJCC 7th edition stage T2 N1 M0.

The participant was male, living with HIV, former smoker reporting 44 pack-years of tobacco use, daily alcohol use, ≥100 lifetime male oral sex partners, a nadir CD4+ T-lymphocyte count of 211 cells/mm³ and CD4+ count at diagnosis of 1146 cells/mm³.
Discussion

This is the longest natural history study of oncogenic oral HPV infection to date, with the largest number of oral HPV16 infections, and the first temporal description of persistent oral HPV16 infection followed until clinical presentation of HPV-OPC. It provides the first estimates of long-term oncogenic oral HPV persistence; 32% of oral HPV16 infections were persistently detected for 5 or more years, and when persistent, were often found at high viral load. These data inform commonly asked questions about the implications of detecting oncogenic oral HPV infection and provide estimates necessary to design future potential screening trials.

Oncogenic oral HPV infection is the precursor to HPV-OPC, based on the cervical HPV-cancer paradigm, as well as two lines of evidence for oral HPV and OPC. First, in a nested case-control study with a single oral rinse sample tested, oral HPV16 preceded the diagnosis of incident OPC by an average of 3.9 years.\textsuperscript{1} Second, multiple retrospective and prospective clinical studies of HPV-OPC patients suggest detection of oncogenic oral HPV infection is equivalent to microscopic evidence of malignancy.\textsuperscript{19-22} Indeed, 50% of HPV-OPC cases with oral oncogenic HPV infection detected after treatment go on to recur, suggesting oral HPV DNA heralds its recurrence. This is the first study to demonstrate longitudinal visits with consistent detection of oncogenic oral HPV infection, the surrogate of microscopic disease, over many years.

Not only was oncogenic oral HPV detected over multiple time points, but an increasing HPV16 viral load was observed, prior to progressing to HPV-OPC. Indeed, increasing viral load in this analysis was associated with reduced clearance of infection, consistent with the cervical cancer literature.\textsuperscript{23} Notably, one third of those with oral HPV16 infection remained persistent five years later. This suggests that there is a group of individuals with persistent infection that are at increased risk of HPV-OPC; if effective screening methods are developed this could be a potential group to focus trials for prevention of HPV-OPC.
Cervical HPV persistence is known to be a necessary cause for malignancy, and has emerged as an acceptable screening test for cervical cancer.\(^{28}\) In contrast, screening for HPV-OPC is not presently endorsed\(^{29,30}\) because several important criteria have not yet been met\(^{31}\), including: 1) a screening test with sufficient sensitivity/specificity\(^{32}\), 2) identified at-risk population\(^{33}\), 3) screening leads to diagnosis at earlier stage, 4) screening reduces morbidity and mortality, 4) cost-effective, and 5) how to clinically evaluate biomarker-positive subjects. This analysis identifies an at-risk population using one potential biomarker, but other critical barriers and limitations remain.

The biomarker used in this study, oncogenic oral HPV DNA, has moderate sensitivity, but good specificity\(^{14,32}\) for HPV-OPC, which would limit its utility (as a stand-alone tests) for any screening program.\(^{36}\) Even among at-risk groups, oral HPV16 prevalence is low, \(\leq 4\%\).\(^{14,33}\) While oral HPV persistence is clearly a pre-requisite to development of HPV-OPC, our data suggest a large number needed to screen to detect a single cancer (poor positive predictive value). Other biomarkers may be necessary to improve early identification of adults at increased risk of HPV-OPC.\(^{14}\) For example, HPV16 E6 antibody appears to be a promising biomarker, and while too rare in the general population to justify use, might be considered for more targeted screening\(^{14,37,38}(p6)\) due to high sensitivity.\(^{14,33}\) Prior studies have shown that HPV16E6 antibodies are detectable up to ten years prior to clinical diagnosis of OPC and that after seroconversion titers either remain stable or increase \(^{39,40}\) thus potentially enabling a one-time assay to identify individuals at increased risk. Follow-up of these individuals might then rely on serial HPV testing in oral rinse and/or plasma.

Consistent with prior long term cervical HPV and shorter oral HPV natural history studies, our study suggests that oral HPV persistence is increased among men, older individuals, PLWH, and prevalent compared to incident infections.\(^{23}\) Clearance for oral HPV in this study population was similar to that for cervical HPV among PLWH (2-year cumulative incidence of clearance of 59\% for oral and \(~60\%\) cervical; quick median time to clearance for oral [1.4 years] and cervical [0.8 years]).\(^{24}\) This is the first study to show that higher viral load is a strong predictor of oral HPV persistence, which suggests it
may serve as a marker of microscopic HPV-OPC. Differences in immune response by sex have been observed for other infections, supporting the possibility of immunologic differences to viral infection among men and women.

This study had several strengths and limitations. The study included a diverse population (sex and race/ethnicity), with a median age younger than that of HPV-OPC diagnosis. POPS (which collected most of the reported data) was not looking for cancers and did not perform clinical examinations. The study population was selected to be at increased risk of oral HPV and included PLWH in whom persistent oral HPV infection is expected to be higher than the general population. Misclassification of our outcome (persistence) cannot be ruled out as it could include clearance and re-infection with the same type-specific infection. Semi-annual visits permit the evaluation of multi-year persistence, but does not inform short-term dynamics within a six months window. The HPV testing in the POPS and MOUTH studies were performed in different labs, while similar but not identical testing methods.

This study suggests there is a subset of individuals with long-term persistent oncogenic oral HPV infection that can develop into HPV-OPC. For the first time we have estimates of oral HPV persistence over 7 years, based upon the largest number of oral oncogenic HPV infections to date. These data could be used to inform studies examining whether and how screening might be appropriate in select populations.

Funding

This work was supported by grant R01DE021395 (NIDCR, NIH; Gypsyamber D’Souza) and R35DE026631 (NIDCR, NIH; Gypsyamber D’Souza). Data in this manuscript were collected by the Multicenter AIDS Cohort Study (MACS) and Womens Interagency HIV Study (WIHS), now the MACS/WIHS Combined Cohort Study (MWCCS). The contents of this publication are solely the responsibility of the authors and do not represent the official views of the National Institutes of Health (NIH). MWCCS (Principal Investigators): Atlanta CRS (Ighovwerha Ofotokun, Anandi Sheth, and Gina...
Wingood), U01-HL146241; Baltimore CRS (Todd Brown and Joseph Margolick), U01-HL146201; 
Bronx CRS (Kathryn Anastos and Anjali Sharma), U01-HL146204; Brooklyn CRS (Deborah Gustafson 
and Tracey Wilson), U01-HL146202; Data Analysis and Coordination Center (Gypsyamber D’Souza, 
Stephen Gange and Elizabeth Golub), U01-HL146193; Chicago-Cook County CRS (Mardge Cohen and 
Audrey French), U01-HL146245; Chicago-Northwestern CRS (Steven Wolinsky), U01-HL146240; 
Connie Wofsy Women’s HIV Study, Northern California CRS (Bradley Aouizerat and Phyllis Tien), 
U01-HL146242; Los Angeles CRS (Roger Detels), U01-HL146333; Metropolitan Washington CRS 
(Seble Kassaye and Daniel Merenstein), U01-HL146205; Miami CRS (Maria Alcaide, Margaret Fischl, 
and Deborah Jones), U01-HL146203; Pittsburgh CRS (Jeremy Martinson and Charles Rinaldo), U01- 
HL146208; UAB-MS CRS (Mirjam-Colette Kempf and Deborah Konkle-Parker), U01-HL146192; UNC 
Crs (Adaora Adimora), U01-HL146194. The MWCCS is funded primarily by the National Heart, Lung, 
and Blood Institute (NHLBI), with additional co-funding from the Eunice Kennedy Shriver National 
Institute Of Child Health & Human Development (NICHD), National Human Genome Research Institute 
(NHGRI), National Institute On Aging (NIA), National Institute Of Dental & Craniofacial Research 
(NIDCR), National Institute Of Allergy And Infectious Diseases (NIAID), National Institute Of 
Neurological Disorders And Stroke (NINDS), National Institute Of Mental Health (NIMH), National 
Institute On Drug Abuse (NIDA), National Institute Of Nursing Research (NINR), National Cancer 
Institute (NCI), National Institute on Alcohol Abuse and Alcoholism (NIAAA), National Institute on 
Deafness and Other Communication Disorders (NIDCD), National Institute of Diabetes and Digestive 
and Kidney Diseases (NIDDK). MWCCS data collection is also supported by UL1-TR000004 (UCSF 
CTSA), P30-AI-050409 (Atlanta CFAR), P30-AI-050410 (UNC CFAR), and P30-AI-027767 (UAB 
CFAR).

Notes
Role of the funders: The funders had no role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.

Disclosures: None of the authors have any conflict of interest to declare

References

1. 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. *JAMA Oncol*. 2016;2(5):599-606. doi:10.1001/jamaoncol.2015.5504

2. Pierce Campbell CM, Kreimer AR, Lin H-Y, et al. Long-term persistence of oral human papillomavirus type 16: the HPV Infection in Men (HIM) study. *Cancer Prev Res (Phila)*. 2015;8(3):190-196. doi:10.1158/1940-6207.CAPR-14-0296

3. Kreimer AR, Pierce Campbell CM, Lin H-Y, et al. Incidence and clearance of oral human papillomavirus infection in men: the HIM cohort study. *Lancet*. 2013;382(9895):877-887. doi:10.1016/S0140-6736(13)60809-0

4. Beachler DC, Sugar EA, Margolick JB, et al. Risk factors for acquisition and clearance of oral human papillomavirus infection among HIV-infected and HIV-uninfected adults. *Am J Epidemiol*. 2015;181(1):40-53. doi:10.1093/aje/kwu247

5. Beachler DC, Lang Kuhs KA, Struijk L, et al. The Natural History of Oral Human Papillomavirus in Young Costa Rican Women. *Sex Transm Dis*. 2017;44(7):442-449. doi:10.1097/OLQ.0000000000000625

6. D’Souza G, Wentz A, Kluz N, et al. Sex Differences in Risk Factors and Natural History of Oral Human Papillomavirus Infection. *J Infect Dis*. 2016;213(12):1893-1896. doi:10.1093/infdis/jiw063

7. Kjær SK, Frederiksen K, Munk C, Iftner T. Long-term absolute risk of cervical intraepithelial neoplasia grade 3 or worse following human papillomavirus infection: role of persistence. *J Natl Cancer Inst*. 2010;102(19):1478-1488. doi:10.1093/jnci/djq356

8. Sand FL, Munk C, Frederiksen K, et al. Risk of CIN3 or worse with persistence of 13 individual oncogenic HPV types. *Int J Cancer*. 2019;144(8):1975-1982. doi:10.1002/ijc.31883

9. US Preventive Services Task Force, Curry SJ, Krist AH, et al. Screening for Cervical Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2018;320(7):674-686. doi:10.1001/jama.2018.10897

10. Kero K, Rautava J, Syrjänen K, Willberg J, Grenman S, Syrjänen S. Smoking increases oral HPV persistence among men: 7-year follow-up study. *Eur J Clin Microbiol Infect Dis*. 2014;33(1):123-133. doi:10.1007/s10096-013-1938-1
11. MACS/WIHS Combined Cohort Study. MACS/WIHS Combined Cohort Study. https://mwccs.org/. Accessed January 28, 2019.

12. Beachler DC, Weber KM, Margolick JB, et al. Risk factors for oral HPV infection among a high prevalence population of HIV-positive and at-risk HIV-negative adults. *Cancer Epidemiol Biomarkers Prev.* 2012;21(1):122-133. doi:10.1158/1055-9965.EPI-11-0734

13. Men and women Offering Understanding of Throat HPV (MOUTH Study). Supported by NIDCR R35 DE026631. http://bit.ly/MOUTHStudy. Accessed September 12, 2018.

14. D'Souza G, Clemens G, Troy T, et al. Evaluating the utility and prevalence of oral rinse and serologic HPV biomarkers for HPV-related oropharyngeal cancer. *Cancer Prev Res (Phila).* August 2019. doi:10.1158/1940-6207.CAPR-19-0185

15. Schiffman M, Clifford G, Buonaguro FM. Classification of weakly carcinogenic human papillomavirus types: addressing the limits of epidemiology at the borderline. *Infect Agents Cancer.* 2009;4:8. doi:10.1186/1750-9378-4-8

16. Wentzensen N, Gravitt PE, Long R, et al. Human papillomavirus load measured by Linear Array correlates with quantitative PCR in cervical cytology specimens. *J Clin Microbiol.* 2012;50(5):1564-1570. doi:10.1128/JCM.06240-11

17. Grabowski MK, Gray RH, Serwadda D, et al. High-risk human papillomavirus viral load and persistence among heterosexual HIV-negative and HIV-positive men. *Sex Transm Infect.* 2014;90(4):337-343. doi:10.1136/sxtrans-2013-051230

18. Wilson LE, Gravitt P, Tobian AAR, et al. Male circumcision reduces penile high-risk human papillomavirus viral load in a randomised clinical trial in Rakai, Uganda. *Sex Transm Infect.* 2013;89(3):262-266. doi:10.1136/sxtrans-2012-050633

19. Ahn SM, Chan JYK, Zhang Z, et al. Saliva and plasma quantitative polymerase chain reaction-based detection and surveillance of human papillomavirus-related head and neck cancer. *JAMA Otolaryngol Head Neck Surg.* 2014;140(9):846-854. doi:10.1001/jamaoto.2014.1338

20. Wang Y, Springer S, Mulvey CL, et al. Detection of somatic mutations and HPV in the saliva and plasma of patients with head and neck squamous cell carcinomas. *Science Translational Medicine.* 2015;7(293):293ra104-293ra104. doi:10.1126/scitranslmed.aaa8507

21. Fakhry C, Blackford AL, Neuner G, et al. Association of Oral Human Papillomavirus DNA Persistence With Cancer Progression After Primary Treatment for Oral Cavity and Oropharyngeal Squamous Cell Carcinoma. *JAMA Oncol.* 2019;5(7):985-992. doi:10.1001/jamaoncol.2019.0439

22. Dahlström KR, Li G, Hussey C, et al. Circulating HPV DNA as a marker for disease extent and recurrence among patients with oropharyngeal cancer. *Cancer.* 2015;121(19):3455-3464. doi:10.1002/cncr.29538

23. de Sanjosé S, Brotons M, Pavón MA. The natural history of human papillomavirus infection. *Best Pract Res Clin Obstet Gynaecol.* 2018;47:2-13. doi:10.1016/j.bpobgyn.2017.08.015
24. Koshiol J, Schroeder J, Jamieson DJ, et al. Smoking and time to clearance of human papillomavirus infection in HIV-seropositive and HIV-seronegative women. *Am J Epidemiol*. 2006;164(2):176-183. doi:10.1093/aje/kwj165

25. van Lunzen J, Altfeld M. Sex differences in infectious diseases-common but neglected. *J Infect Dis*. 2014;209 Suppl 3:S79-80. doi:10.1093/infdis/jiu159

26. Gillison ML, Broutian T, Pickard RKL, et al. Prevalence of oral HPV infection in the United States, 2009-2010. *JAMA*. 2012;307(7):693-703. doi:10.1001/jama.2012.101

27. Chaturvedi AK, Graubard BI, Broutian T, et al. NHANES 2009-2012 Findings: Association of Sexual Behaviors with Higher Prevalence of Oral Oncogenic Human Papillomavirus Infections in U.S. Men. *Cancer Res*. 2015;75(12):2468-2477. doi:10.1158/0008-5472.CAN-14-2843

28. Huh WK, Ault KA, Chelmow D, et al. Use of primary high-risk human papillomavirus testing for cervical cancer screening: Interim clinical guidance. *Gynecologic Oncology*. 2015;136(2):178-182. doi:10.1016/j.ygyno.2014.12.022

29. Lingen MW, Abt E, Agrawal N, et al. Evidence-based clinical practice guideline for the evaluation of potentially malignant disorders in the oral cavity: A report of the American Dental Association. *J Am Dent Assoc*. 2017;148(10):712-727.e10. doi:10.1016/j.adaj.2017.07.032

30. Final Update Summary: Oral Cancer: Screening - US Preventive Services Task Force. https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/oral-cancer-screening1. Accessed November 22, 2019.

31. Kreimer AR, Shiels MS, Fakhry C, et al. Screening for human papillomavirus-driven oropharyngeal cancer: Considerations for feasibility and strategies for research. *Cancer*. 2018;124(9):1859-1866. doi:10.1002/cncr.31256

32. Gipson BJ, Robbins HA, Fakhry C, D'Souza G. Sensitivity and specificity of oral HPV detection for HPV-positive head and neck cancer. *Oral Oncol*. 2018;77:52-56. doi:10.1016/j.oraloncology.2017.12.008

33. D'Souza G, McNeel TS, Fakhry C. Understanding personal risk of oropharyngeal cancer: risk-groups for oncogenic oral HPV infection and oropharyngeal cancer. *Ann Oncol*. 2017;28(12):3065-3069. doi:10.1093/annonc/mdx535

34. U.S. Census Bureau. *Age and Sex Composition: 2010 Census Briefs*. https://www.census.gov/prod/cen2010/briefs/c2010br-03.pdf.

35. Senkomago V, Henley SJ, Thomas CC, Mix JM, Markowitz LE, Saraiya M. Human Papillomavirus-Attributable Cancers - United States, 2012-2016. *MMWR Morb Mortal Wkly Rep*. 2019;68(33):724-728. doi:10.15585/mmwr.mm6833a3

36. Levman J. Longitudinal disease detection rates for the evaluation of disease detection technologies with application in high-risk breast cancer screening. *J Clin Diagn Res*. 2013;7(12):2932-2935. doi:10.7860/JCDR/2013/5693.3794
37. Holzinger D, Wichmann G, Baboci L, et al. Sensitivity and specificity of antibodies against HPV16 E6 and other early proteins for the detection of HPV16-driven oropharyngeal squamous cell carcinoma. *Int J Cancer*. 2017;140(12):2748-2757. doi:10.1002/ijc.30697

38. Beachler DC, Waterboer T, Pierce Campbell CM, et al. HPV16 E6 seropositivity among cancer-free men with oral, anal or genital HPV16 infection. *Papillomavirus Res*. 2016;2:141-144. doi:10.1016/j.pvr.2016.07.003

39. Kreimer AR, Johansson M, Yanik EL, et al. Kinetics of the Human Papillomavirus Type 16 E6 Antibody Response Prior to Oropharyngeal Cancer. *J Natl Cancer Inst*. 2017;109(8). doi:10.1093/jnci/djx005

40. Zhang Y, Waterboer T, Haddad RI, et al. Human papillomavirus (HPV) 16 antibodies at diagnosis of HPV-related oropharyngeal cancer and antibody trajectories after treatment. *Oral Oncology*. 2017;67:77-82. doi:10.1016/j.oraloncology.2017.02.004
Table 1. Descriptive characteristics of 264 men from the Multicenter AIDS Cohort Study (MACS) and 183 women from the Women’s Interagency HIV Study (WIHS) with oncogenic oral HPV detected at any point during the POPS study, compared to only those 128 MACS and WIHS participants enrolled in MOUTH.

| Participant characteristics at time of first oral HPV detection | N=447 | N=447 | N=264 | N=183 | N=128 |
|---------------------------------------------------------------|-------|-------|-------|-------|-------|
| Sex                                                     |       |       |       |       |       |
| Men (MACS)                                       | 264   | 59.1% | 100%  | 0%    | 47.7% |
| Women (WIHS)                                     | 183   | 40.9% | 0%    | 100%  | 52.3% |
| Race                                                 |       |       |       |       |       |
| White non-Hispanic                                | 148   | 33.1% | 51.5% | 6.6%  | 35.2% |
| Black non-Hispanic                                | 225   | 50.3% | 34.1% | 73.8% | 57.8% |
| Other (Hispanic, Other)                           | 74    | 16.6% | 14.4% | 19.7% | 7.0%  |
| Education level                                   |       |       |       |       |       |
| <HS                                               | 95    | 21.9% | 4.4%  | 46.2% | 21.8% |
| HS/GED                                           | 210   | 48.5% | 47.8% | 49.5% | 49.2% |
| College                                          | 52    | 12.0% | 17.5% | 4.4%  | 13.7% |
| Adv/Degree                                       | 76    | 17.6% | 30.3% | 0.0%  | 15.3% |
| Smoking status                                    |       |       |       |       |       |
| Never                                             | 82    | 18.6% | 24.8% | 9.5%  | 15.5% |
| Former                                           | 160   | 36.3% | 42.8% | 26.8% | 30.9% |
| Current                                          | 199   | 45.1% | 32.4% | 63.7% | 53.7% |
| Current alcohol use                               |       |       |       |       |       |
| No                                                | 131   | 30.1% | 16.7% | 49.2% | 30.6% |
| Yes                                              | 305   | 70.0% | 83.3% | 50.8% | 69.4% |
| Ever performed oral sex                           |       |       |       |       |       |
| No                                                | 24    | 5.8%  | 2.1%  | 10.9% | 7.1%  |
| Yes                                              | 392   | 94.2% | 97.9% | 89.1% | 92.9% |
| HIV-status                                        |       |       |       |       |       |
| HIV-uninfected                                    | 119   | 26.6% | 28.4% | 24.0% | 25.0% |
| Person living with HIV (PLWH)                      | 328   | 73.4% | 71.6% | 76.0% | 75.0% |
| Currently on HAART (among HIV+)                    |       |       |       |       |       |
| No                                                | 61    | 18.9% | 21.9% | 14.8% | 12.1% |
| Yes                                              | 261   | 81.1% | 78.1% | 85.2% | 87.9% |
| Infection Characteristics                         |       |       |       |       |       |
| Oral HPV16 prevalence                             | 447   | 19.9% | 17.1% | 24.0% | 31.3% |
| Oral non-16 oncogenic HPV                          | 447   | 83.9% | 86.7% | 79.8% | 72.7% |
| Age (in years): Median (IQR)                       | 441   | 50 (43, 56) | 51 (44, 58) | 47 (42, 54) | 50 (44, 57) |
| Current CD4 cell count (cells/ml), among PLWH: Median (IQR) | 322 | 522 (326, 715) | 555 (385, 746) | 434 (228, 640) | 439 (264, 657) |
| Current HIV RNA, among PLWH: Median (IQR)         | 321   | und (und, 1170) | und (und, 263) | 48 (und, 2887) | 48 (und, 600) |
Table 2. Unadjusted and adjusted associations between selected characteristics and oncogenic HPV clearance, MACS and WIHS, 2009-2018

| Characteristic                          | Number of visits | HR (95% CI) | aHR (95%CI) |
|-----------------------------------------|------------------|-------------|-------------|
| **Sex**                                 |                  |             |             |
| Female                                  | 942              | 1.00 (Ref)  | 1.00 (Ref)  |
| Male                                    | 1,558            | 0.63 (0.51-0.79) | 0.74 (0.60-0.91) |
| **Person Living With HIV (PLWH)**       |                  |             |             |
| HIV-negative                            | 551              | 1.00 (Ref)  | 1.00 (Ref)  |
| PLWH, current CD4 count ≥500            | 933              | 0.91 (0.70-1.12) | 0.74 (0.58-0.95) |
| PLWH, current CD4<500                  | 879              | 0.88 (0.62-1.25) | 0.69 (0.53-0.92) |
| **Oral HPV infection type**             |                  |             |             |
| Incident                                | 928              | 1.00 (Ref)  | 1.00 (Ref)  |
| Prevalent                               | 1,572            | 0.43 (0.35-0.53) | 0.44 (0.35-0.55) |
| **Age (by increasing decade – ref 20-29 year olds)** | 2,363            | 0.81 (0.74-0.89) |             |
| **Age categories (years)**              |                  |             |             |
| 22-40                                   | 292              | 1.00 (Ref)  | 1.00 (Ref)  |
| 40-49                                   | 675              | 0.78 (0.58-1.05) | 0.73 (0.54-0.99) |
| 50-59                                   | 934              | 0.76 (0.57-1.02) | 0.76 (0.57-1.01) |
| 60-79                                   | 462              | 0.42 (0.30-0.61) | 0.40 (0.27-0.59) |
| **Oncogenic oral HPV type**             |                  |             |             |
| Non16 HR HPV*                           | 2,051            | 1.00 (Ref)  | 1.00 (Ref)  |
| HPV16                                   | 449              | 1.48 (1.03-2.14) | 1.19 (0.89-1.61) |
| **History of tonsillectomy**            |                  |             |             |
| No                                      | 1,717            | 1.00 (Ref)  |             |
| Yes                                     | 603              | 0.78 (0.58-1.04) |             |
| **Frequency of toothbrushing**          |                  |             |             |
| ≥2 times per day                        | 2,083            | 1.00 (Ref)  |             |
| <2 times per day                        | 337              | 0.92 (0.65-1.29) |             |
| **Current smoker**                      |                  |             |             |
| No                                      | 1,231            | 1.00 (Ref)  |             |
| Yes                                     | 1,130            | 1.05 (0.86-1.29) |             |
| **Current alcohol use**                 |                  |             |             |
| No                                      | 746              | 1.00 (Ref)  |             |
| Yes                                     | 1,589            | 1.08 (0.86-1.35) |             |
| **Current marijuana use**               |                  |             |             |
| No                                      | 1,620            | 1.00 (Ref)  |             |
| Yes                                     | 702              | 1.00 (0.80-1.26) |             |

* Non-16 oncogenic oral HPV types included all oncogenic types except HPV16: HPV 18/31/33/35/39/45/51/52/56/58/59/66.
Figure Titles and Legends

Figure 1. 447 MACS and WIHS cohort study participants with 676 oncogenic oral HPV infection. 447 MACS and WIHS cohort study participants with 676 oncogenic oral HPV infection. The POPS study was from 2009 to 2015 and MOUTH study visit included data collected from 2017 to 2018. The blue dots represent study visits which occurred at 6 months intervals during the POPS study. There was a lapse between POPS and MOUTH symbolized by the absence of dots. During the MOUTH study there was a one-time visit which ranged between October 2017- September 2018.

Figure 2: Time to clearance of any oncogenic HPV among 676 infections in the 447 MACS/WIHS subjects enrolled in POPS and/or MOUTH study overall. Color band indicated the 95% confidence interval.

Figure 3: Time to clearance of any oncogenic HPV among 676 infections in the 447 MACS/WIHS subjects enrolled in POPS and/or MOUTH study by risk factor. By infection characteristics (prevalent vs incident: panel A); HPV16 vs non16 oncogenic HPV type (panel B), demographic characteristics (men vs women: panel C); and age (panel D).

Figure 4. Oral HPV16 viral load across study visits, among participants with persistent and cleared infections. Each row is a distinct subject in the study who had oral HPV16 DNA detected. Semi-quantitative signal intensity depicted by color as follows: strong (dark red), medium (red), weak (pink) and negative (white with -). Grey represents visits with no oral rinse sample collection. The asterisk (*) denotes the subject diagnosed with HPV-OPC. Figure is restricted to participants with at least 4 oral rinse samples tested. Participants joined study at variable times. Oral rinse results are shown by calendar visit, labeled “V1” for the first bi-annual visit. Gray represents visits where oral rinse sample was not
collected/tested (this includes visits before substudy enrollement for those who entered later, as well as some missed visits, and visits after study completion). MOUTH baseline study visit labeled “M”.
Figure 1.

POPS Study
Oct 2009 – Sep 2015

MOUTH Study
Oct 2017-present

Oral oncHPV persistent
Figure 2.
Figure 3

(A) Prevalent and incident HPV16 infections over 7 years with 95% CI. The p-value is less than 0.001.

(B) Prevalent and incident non-16 onc HPV infections over 7 years with 95% CI. The p-value is 0.02.

(C) Prevalent and incident HPV16 infections in different age groups (40-59, <40, 60+) over 7 years with 95% CI. The p-value is less than 0.001.

(D) Prevalent and incident non-16 onc HPV infections in different age groups (40-59, <40, 60+) over 7 years with 95% CI. The p-value is less than 0.001.
Figure 4.

Persistent infections

Cleared infections