

Georgia State University
ScholarWorks @ Georgia State University

Mathematics Theses

Department of Mathematics and Statistics

Spring 5-2-2018

Analysis of Prevalence of Oral Human Papillomavirus Infection and Risk Factors among Adults in the United States-National Health and Nutrition Examination Survey 2013-2014

Tengguo Li

Follow this and additional works at: https://scholarworks.gsu.edu/math_theses

Recommended Citation

Li, Tengguo, "Analysis of Prevalence of Oral Human Papillomavirus Infection and Risk Factors among Adults in the United States-National Health and Nutrition Examination Survey 2013-2014." Thesis, Georgia State University, 2018.
https://scholarworks.gsu.edu/math_theses/161

This Thesis is brought to you for free and open access by the Department of Mathematics and Statistics at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Mathematics Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

ANALYSIS OF PREVALENCE OF ORAL HUMAN PAPILLOMAVIRUS INFECTION AND
RISK FACTORS AMONG ADULTS IN THE UNITED STATES
-NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY 2013-2014

by

TENGGUO LI

Under the Direction of Gengsheng Qin, PhD

ABSTRACT

Oral Human Papillomavirus (HPV) infection is a major cause of oropharyngeal cancer that is increasing in incidence over the last decades. However, little is known about the epidemiology of oral HPV infection. Using complex data and HPV testing results from 4724 adults collected through National Health and Nutrition Examination Survey (2013-2014), we estimated the prevalence of oral HPV and evaluated risk factors associated with HPV infection. The oral HPV prevalence was 7.18% for any types and 4.5% for high-risk types in adults aged 18-69 years. The prevalence varied significantly by age, demographic characteristics, and sexual history. Multiple factors including age, gender, and sexual history were identified as risk factors for oral HPV infection. Prevalence of HPV was higher in men (11.1%) than in women (3.3%). Several characteristics, including race, tobacco and alcohol use, and education, were identified as risk modifiers for HPV infection among men and women.

INDEX WORDS: NHANES, Oral HPV, Prevalence, Oropharyngeal cancer, Complex survey,

Risk factor

ANALYSIS OF PREVALENCE OF ORAL HUMAN PAPILLOMAVIRUS INFECTION AND
RISK FACTORS AMONG ADULTS IN THE UNITED STATES
-NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY 2013-2014

by

TENGGUO LI

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2018

Copyright by
Tengguo Li
2018

ANALYSIS OF PREVALENCE OF ORAL HUMAN PAPILLOMAVIRUS INFECTION AND
RISK FACTORS AMONG ADULTS IN THE UNITED STATES
-NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY 2013-2014

by

TENGGUO LI

Committee Chair: Gengsheng Qin

Committee: Xin Qi

Xiaoyi Min

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2018

DEDICATION

This thesis work is dedicated to my wife and my children. They provided me with constant support and continuous encouragement throughout the years of my study and the research for this thesis. This accomplishment would not have been possible without their support and love.

ACKNOWLEDGEMENTS

I would first like to express my gratitude to my thesis advisor, Dr. Gengsheng Qin, for his consistent support, guidance, and timely responses to all my questions. I am grateful to his encouragement and suggestions. This thesis would have not been possible without his support and help.

I would like to thank members of my thesis committee, Dr. Xin Qi and Dr. Xiaoyi Min, for taking time to review the thesis, Thank you for the comments, suggestions, and positive feedbacks on this research work.

Finally, I want to acknowledge the faculty members and classmates at Department of Mathematics and Statistics of Georgia State University for their help and support over the years.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	V
LIST OF TABLES	VIII
LIST OF FIGURES	IX
LIST OF ABBREVIATIONS	XI
1 INTRODUCTION	1
1.1 Background.....	1
1.2 Significance	6
1.3 Data Source.....	6
<i>1.3.1 Survey Design.....</i>	<i>7</i>
<i>1.3.2 Sample Weights.....</i>	<i>7</i>
<i>1.3.3 Variance Estimation</i>	<i>8</i>
1.4 Purpose of Study	8
1.5 Expected Results.....	9
2 METHODS.....	10
2.1 Study Population and Survey Design	10
2.2 Demographic and Behavioral Characteristics.....	11
2.3 Laboratory Methods	11
<i>2.3.1 Oral Specimen Collection.....</i>	<i>11</i>
<i>2.3.2 DNA Isolation</i>	<i>11</i>

2.3.3	<i>HPV Genotyping Test</i>	12
2.4	Statistical Analysis.....	12
3	RESULTS	15
3.1	Overall Oral HPV Prevalence.....	15
3.2	Oral HPV Prevalence by Demographic and Life Style Characteristics.....	16
3.3	Oral HPV Prevalence by Sexual History	22
3.4	Univariate Analysis of Factors Associated with HPV Infection	25
3.5	Multivariable Analysis of Factors Associated with Oral HPV Infection	28
3.6	Factors Associated with Oral HPV Infection by Gender	30
4	DISCUSSION.....	39
4.1	Overall Prevalence of Oral HPV Infection	39
4.2	Prevalence of Oral HPV Infection Varies by Characteristics.....	40
4.3	Factors Associated with Oral HPV Infection	40
4.4	Differences of Oral HPV infection in Males and Females.....	41
4.5	Limitations	43
5	CONCLUSIONS.....	44
	REFERENCES.....	45

LIST OF TABLES

Table 1 Number of HPV-associated and HPV-attributable cancer cases per year	5
Table 2 Frequencies of single and multiple infections of oral HPV in U.S. population aged 18-69 years	16
Table 3 Weighted prevalence of oral HPV among U.S. adults (18-69 years) by demographic characteristics, National Health and Nutrition Examination Survey, 2013-2014.....	17
Table 4 Weighted prevalence of oral HPV among U.S. adults aged 18-69 years by sexual behavior, National Health and Nutrition Examination Survey, 2013-2014.....	24
Table 5 Univariate analysis of characteristics associated with oral HPV infection (any type) in U.S. population aged 18-69 years	26
Table 6 Multivariable analysis of characteristics independently associated with HPV infection among people aged 18-69 years.....	29
Table 7 Adjusted Odds Ratio for HPV infection by gender in U.S. population aged 18-69 years	33

LIST OF FIGURES

Figure 1 Weighted prevalence of Oral HPV infection by individual HPV types in U.S. population aged 18-69 years	15
Figure 2 Weighted oral HPV prevalence by age (14-69 years)	20
Figure 3 Weighted oral HPV prevalence by age group	21
Figure 4 Weighted prevalence of HPV16 by age	22
Figure 5 Weighted prevalence of oral HPV (any type) among U.S. populations by age and gender.....	32
Figure 6 Weighted prevalence of oral HPV16 infection among U.S. adults by gender	32
Figure 7 Adjusted Odds Ratio for oral HPV (any type) infection among U.S. adults by age and gender.....	34
Figure 8 Adjusted Odds Ratio for oral HRHPV infection among U.S. adults by age and gender	34
Figure 9 Adjusted Odds Ratio for oral HPV (any type) infection among U.S. population aged 18-69 years by education and gender	35
Figure 10 Adjusted Odds Ratio for oral HPV infection among U.S. population aged 18-69 years by tobacco use and gender	35
Figure 11 Adjusted Odds Ratio for HRHPV infection among U.S. population aged 18-69 years by gender and alcohol use.....	37
Figure 12 Adjusted Odds Ratio for HRHPV infection among U.S. population aged 18-69 years by gender and race	37
Figure 13 Association of gender and age of sex debut with oral HPV infection among U.S. population aged 18-69 years	38

Figure 14 Weighted prevalence of oral HPV (any type) infection among U.S. adults by race and
number of lifetime opposite sex partners..... 38

LIST OF ABBREVIATIONS

- ACIP:** Advisory Committee on Immunization Practice
- CDC:** Centers for Disease Control and Prevention
- CI:** Confidence Interval
- FDA:** Food and Drug Administration
- HPV:** Human Papillomavirus
- HRHPV:** High-Risk Human Papillomavirus
- LA:** Linear Array
- LRHPV:** Low-Risk Human Papillomavirus
- NCHS:** National Center for Health Statistics
- NHANES:** National Health and Nutrition Examination Survey
- MEC:** Mobile Examination Center
- OHD:** Oral HPV Detection Laboratory
- OPC:** Oropharyngeal Cancer
- PSU:** Primary Sampling Units
- SEER:** Surveillance, Epidemiology, and End Results
- TSL:** Taylor Series Linearization

1 INTRODUCTION

1.1 Background

Human Papillomaviruses (HPV) are small DNA viruses with double stranded circular genome about ~7000-8000bp, surrounded by icosahedral capsids of 52-55nm in diameter [1, 2]. The genome encodes 7 early proteins (E1, E2, E4, E5A, E5B, E6, and E7) and 2 late capsid proteins (L1 and L2). Based on DNA sequence similarity of the late gene L1, HPVs are classified into α , β , and γ genera comprising more than 200 HPV types [3, 4].

HPV infection in epithelial cells is the most common sexually transmitted disease [5]. HPV infection in cutaneous cells can cause warts, while in the mucosal cells HPV can cause genital warts and various forms of cancers including anogenital, oropharyngeal cancers, and possibly skin cancers [6]. It is estimated that 79 million people are infected by some types of HPVs in the United States [7]. While majority of HPV infection are cleared by the host immune system, persistent infection of HPV in a small portion of people can cause cancer [8]. About 600,000 incident cancers worldwide and 250,000 premature deaths are related to HPV infection [9]. According to the data from National Program of Cancer Registries of Centers for Disease Control and Prevention (CDC), about 39,800 new cases of cancer occur annually in the human body sites where HPVs are frequently found in the United States. Among these cases, 79% (31,500) are caused by HPV infection [10]. As shown in Table 1, HPV causes 91% of cervical cancers and a portion of cancers on various sites including vagina (75%), vulva (69%), penis (63%), anus (91%), rectum (69%), and oropharynx (70%) [10]. HPV types differ in their potential in causing epithelial abnormalities. Most of the 200 known HPV types infect cutaneous epithelial cells and lead to skin wart. More than 90% of HPV attributable cancers are caused by a group of HPVs (~40 types) in α genera [11]. Depending on their oncogenic potential, HPV

types are grouped into high-risk HPV (HRHPV) or low-risk HPV (LRHPV) [4, 12]. HRHPV currently include 18 types, HPV 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68, 69, 73, and 82 [13, 14]. HRHPV act as carcinogen and can result in the development of anogenital cancers. In cervix, HRHPV infection lead to low-grade and high-grade cervical cell abnormalities, which can progress to cervical cancer. HPV16 and HPV18 account for about 70% of all cases of cervical cancers worldwide [15, 16]. Infection with a high-risk HPV type is necessary, but not sufficient for the development of cervical cancer, because the vast majority of women with HPV infection do not develop cancer [12, 17]. LRHPV includes HPV type 6, 11, 40, 42, 54, 55, 61, 62, 64, 67, 69, 70, 71, 72, 81, subtype IS39 of HPV82, 83, 84, and 89 [cp6108] [13, 14]. Infection with low-risk HPV types such as type 6 and 11, generally does not result in cancer, but can cause benign or low-grade cervical cell abnormalities, genital warts, and laryngeal papillomas [18].

Human is the only natural reservoir for HPV. HPV is predominantly transmitted through sexual contact with an infected individual [19]. As a result, HPV infection is common among sexually active populations. Using data for 4150 females collected between 2003 to 2006, the overall prevalence of genital HPV was estimated as 42.5% in females aged 14-59 years [20]. In a cross-sectional study using data from National Health and Nutrition Examination Survey (NHANES), the overall prevalence of genital HPV was estimated as 45.2% among 1868 men aged 18-59 years [21]. The key risks for HPV infection are related to sexual behaviors, including number of sex partners, the age of sex debut, and the likelihood of sex partners being a HPV carrier [22, 23]. People of young age (<25 years) are vulnerable to HPV infections. A prospective study showed that exposure to new partners conferred high risk for incident HPV infections [24]. It was reported that the number of sex partners of men was a key determinant of

cervical cancer risk for their wives. The prevalence of penile HPV infection had a positive trend along with increasing number of sexual partners [25]. Additional risk factor for HPV infection include compromised immune system, long term use of oral contraceptives, inconsistent condom use, as well as smoking and alcohol use. Despite the compelling evidence that HPVs are mainly sexually transmitted, other transmission route, including transmission from mother to child and from skin to skin, were documented, although they appeared to be uncommon [19].

HPV vaccination for girls at age of 11 years with catch-up vaccination to age of 26 years was recommended by the Advisory Committee on Immunization Practice (ACIP) of the CDC in 2006 [26]. In 2009, U.S. Food and Drug Administration (FDA) approved the use of quadrivalent HPV vaccine (HPV 6, 11, 16, and 18) in males with guidance from ACIP for genital wart prevention [27]. Routine HPV vaccination was further expanded in 2011 in male individuals aged 11 to 26 years [28]. Recently, a 9-valent HPV vaccine including type 6, 11, 16, 18, 31, 33, 45, 52, and 58 was approved to prevent HPV attributable cancers and decrease HPV transmissions [29].

Oral HPVs are detected at high prevalence in oropharyngeal cancer (OPC) patients [16, 30, 31]. Strong molecular evidence has shown that HPV plays a causal role in the development of OPCs, in which expression of the oncogenes (E6 and E7) and integration of HPV genomes have been shown [32]. HPV infection is considered to accounts for 70% of oropharyngeal squamous cell cancer (OPSC) in the United State (Table 1) [10]. It is estimated that about 3,200 new cases of HPV-associated OPCs in women and about 13,200 in men occur each year in the United States [33]. Notably, the incidence of HPV attributable OPCs (HPV-OPC) increased by 225% from 1988 to 2004, while the incidence of HPV negative OPCs decreased by 50% [34]. This data strongly suggest that HPV infection is the underlying cause for overall increase in OPC

incidence [34-36]. Particularly, OPC incidence among men has doubled over the past 20 years [37]. To date, OPC becomes the most common HPV-related cancers, its incidence among men (7.8 per 100,000) has surpassed incidence rates of cervical cancer among women (7.4 per 100,000) [38]. It was reported that HPV negative OPCs were associated with chronic alcohol and tobacco use. In contrast, HPV positive OPCs were associated with sex behavior, particularly in young individuals [39]. The reason for the predominant increase of HPV positive OPCs in men was not clear [13]. Notably, more than 90% of HPV-positive OPCs are related to HPV16 and oral HPV infection confers higher risks for OPCs [39].

Oral HPV prevalence is low compared to that of genital infections. In one study, the overall prevalence of oral HPV was estimated as 6.9% in people aged 14-69 years using data from NHANES collected from 2009 to 2010. Men had higher prevalence compared to women [13]. It was suggested that oral HPV infection was predominantly transmitted through sexual contact. Sexual behavior and current smoking, including intensity, appeared to be potentially modifiable risk factors for oral HPV infection. Oral HPV prevalence was elevated among people with more than 20 lifetime sexual partners [13]. Recently, another group examined the prevalence of oncogenic oral HPV using data containing participants aged 20-69 enrolled in NHANES from 2009 to 2014, OPC cases from Surveillance, Epidemiology, and End Results (SEER 18), and oropharyngeal cancer mortality from National Center for Health Statistics (NCHS) [40]. They found that oncogenic oral HPV DNA was detected in 3.5% of all adults aged 20–69 years. The highest prevalence (8.1%) was observed in men aged 50–59 years old. Factors conferring high risk for oncogenic oral HPV infection included number of lifetime oral sexual partners, tobacco use, and having more than 5 lifetime oral sexual partners [40].

Table 1 Number of HPV-associated and HPV-attributable cancer cases per year

Cancer site	Average number of cancers per year in sites where HPV is often found	Percentage probably caused by any HPV type ^a	Number probably caused by any HPV type ^a	Percentage probably caused by HPV types 16/18 ^b	Number probably caused by HPV types 16/18 ^b	Percentage probably caused by HPV types 31/33/45/52/58 ^c	Number probably caused by HPV types 31/33/45/52/58 ^c
Cervix	11,693	91%	10,600	66%	7,700	15%	1,700
Vagina	819	75%	600	55%	500	18%	100
Vulva	3,671	69%	2,500	49%	1,800	14%	500
Penis	1,181	63%	700	48%	600	9%	100
Anus	5,229	91%	4,800	79%	4,200	8%	400
Female	3,416	93%	3,200	80%	2,700	11%	400
Male	1,813	89%	1,600	79%	1,400	4%	100
Rectum	772	91%	700	79%	600	8%	100
Female	528	93%	500	80%	400	11%	100
Male	244	89%	200	79%	200	4%	<100
Oropharynx	16,479	70%	11,600	60%	9,900	6%	900
Female	3,203	63%	2,000	51%	1,600	10%	300
Male	13,276	72%	9,600	63%	8,400	4%	600
TOTAL	39,844	79%	31,500	63%	25,300	10%	3,800
Female	23,330	83%	19,400	63%	14,700	13%	3,100
Male	16,514	73%	12,100	64%	10,600	4%	700

-Adopted from the CDC cancer statistics, <https://www.cdc.gov/cancer/hpv/statistics/cases.htm>.

Data are from population-based cancer registries participating in the CDC National Program of Cancer.

Note: Registries and/or the NCI Surveillance, Epidemiology, and End Results Program, meeting criteria for high data quality for all years 2009 to 2013 and cover about 99% of the U.S. population.

^aHPV types detected in genotyping study; most were HRHPV types known to cause cancer.

^bHPV types 16/18 can be prevented by the bivalent, quadrivalent, and 9-valent HPV vaccines.

^cHPV types 31/33/45/52/58 can be prevented by the 9-valent HPV vaccine.

1.2 Significance

With the dramatic increase of incidence of OPCs, there is an increasing interest in identifying healthy individuals at risk for development of oropharyngeal cancer to provide information for screening strategies. Due to the facts that a large portion of OPCs are caused by oral HPV and that HPV-related OPC has better prognosis, monitoring and understanding the trend of prevalence of oral HPV are essential for HPV surveillance, evaluation of the burden of HPV associated cancers and the efficacy of HPV vaccination, as well as cancer control and prevention. Despite the studies outlined above, we still have an incomplete picture regarding to oral HPV infection epidemiology. More data are need to confirm previous findings and provide update data for prevalence of oral HPV. Furthermore, analysis of risk factors for oral HPV prevalence and transmission holds the key for developing prevention programs. However, little is known about the risk factors for oral HPV infection. In this thesis, we conduct a population-based study to investigate the prevalence of oral HPV and identify risk factors for HPV infection, using data collected from 2013 to 2014 through the US National Health and Nutrition Examination Survey (NHANES) studies.

1.3 Data Source

NHANES comprises a series of studies conducted by National Center for Health Statistics (NCHS) at the CDC [41]. It is an ongoing health and nutrition complex survey designed to assess the health and nutritional status of adults and children in the United States. The samples for the survey are selected to represent the U.S. population of all ages. NHANES uses a complex, multistage probability design to sample the civilian, noninstitutionalized population in the 50 states and Washington, D.C. The survey contains HPV testing for oral and genital samples [41].

The unique feature that NHANES combines interviews with physical examinations makes it a good choice for population-based HPV prevalence study.

1.3.1 Survey Design

The selection of participants involves 4 consecutive stages. Stage 1 is called primary sampling units (PSUs) selection. All counties are divided into 15 groups. One county is selected from each group, resulting in 15 counties each year. Stage 2 involves selection of segments within PSUs, which forms a block or group of blocks containing a cluster of households. In stage 3, all houses or apartments in the small groups selected in stage 2 are identified, and about 30 households are sampled from each group. In stage 4, interviewers go to each selected house and conduct an interview. People received interviews are invited to a Mobile Examination Center (MEC) for physical examination such as blood pressure, dental examination, and collection of specimens for laboratory testing. Some groups, such as people who are 60 years and older, Hispanics, Asian, and African Americans, are over sampled to produce reliable statistics [42].

1.3.2 Sample Weights

Sample weights are measures of the number of persons represented by the particular sampled participants. Sample data weighting allows estimates that would have been obtained if the entire sampling frame had been surveyed. Weighting considers the different probabilities of selection for sampling domains, nonresponse, and difference between the final sample and total population. In NHANES, samples were first weighted to compensate for unequal selection probabilities among subgroups such as race and sex. Then samples were adjusted for nonresponse to reduce biases. Finally, the sample weights were post-stratified to U.S. Census Bureau estimates of target population totals to compensate inadequacies in the sample frame, as well as to reduce variance. Calculated final interview weights are used for data analysis for

household interview, and the final examination weights are used for examination data or in conjunction with house hold interview data [42].

1.3.3 Variance Estimation

Exact mathematical formulas for variance estimate are not available for complex sample survey. To provide approximately unbiased and design-consistent estimate for variance, variance approximation procedures are required. Because variance computations need to consider survey design, standard statistical software packages assuming random samples are not applicable for survey data. Two variance approximation procedures are used for NHAES survey analysis, replication methods and Taylor Series Linearization (TSL). Both methods consider the complex sample design and compute the design effects. The replication method was used for HNANES III. Now NCHS recommends to use TSL method for variance estimation for all HNANES data. In this method, the linear approximation is generated by taking the first-order Taylor series approximation for the estimator. Then standard variance estimation methods for linear statistics are used to estimate the variance of the linearized estimator. Various software packages including SAS survey procedures, R and SUDDAN are available to estimate variance by this method [42].

1.4 Purpose of Study

In this study, we will select variables from demographic and questionnaire survey and combine with oral HPV testing variables using NHANES data collected during 2013-2014. We will examine the data to achieve the following aims:

Aim 1: To estimate the prevalence of oral HPV in U.S. populations aged 18-69 years

Aim 2: To assess the distribution of HPV infection by risk categories

Aim 3: To identify factors conferring risks for HPV infection

1.5 Expected Results

We expect weighted prevalence for oral HPV infection to be between 5-10% in U.S. adult populations aged 18-69 years. Prevalence of different HPV types are to be estimated. Significant differences in distribution of HPV infection among demographic and behavioral characteristics are expected to be observed. A number of factors may be identified to modify the risk of being infected by HPV. Using multivariate logistic regression, adjusted ORs are expected to be generated for risk evaluation.

2 METHODS

2.1 Study Population and Survey Design

Data were extracted from NHANES website (<https://wwwn.cdc.gov/nchs/nhanes/ContinuousNhanes/Default.aspx?BeginYear=2013>). The consenting participants have household interviews by dietary and health interviewers followed by physical examinations in a Mobile Examination Center (MEC). The interview contains questions covering demographic, socioeconomic, dietary, and health-related topics. The examination includes medical, dental, and physiological measurements, as well as laboratory tests. Survey data were collected in consideration of diseases, health indicators and risk factors, prevalence of chronic conditions, as well as estimates for previously undiagnosed conditions. Like any complex survey, the data has the following features: unequal weights, stratification, clustering and finite population corrections [41].

NHANES 2013-2014 survey data are used for this analysis. The dataset included 14,332 participants from 30 different survey locations. Oversampled subgroups include Hispanic persons, Non-Hispanic Black persons, Non-Hispanic Asian persons, Non-Hispanic White and other persons at or below 130 percent of the poverty level, as well as Non-Hispanic white and other persons aged 80 years and older. Of those participants, 10,175 (71%) completed the interview and 9,813 (68.5%) underwent health examinations in the MEC. We extracted data for 5057 (51.5%) adult participants (18-69 years old) who submitted oral samples for HPV testing. Of these samples, 333 were inadequate for HPV DNA typing and were excluded from the analysis. The final dataset had 4724 participants with HPV test results [43].

2.2 Demographic and Behavioral Characteristics

Demographic data, including age, gender, race, country of birth, education, and marital status, were obtained from the NHANES household interview survey data (2013-2014) [43]. Income-poverty ratio was calculated based on the U.S. Census definition by dividing total family income by the poverty threshold after adjusting for family size at the time of interview.

Data on substance use, such as alcohol and tobacco, and sexual behavior were self-reported by participants through audio computer-assisted self-interview at the MEC. One drink of alcohol was defined as a 12 oz. beer, a 5 oz. glass of wine, or one and a half ounces of liquor. Sexual behavior variables included ever having sex (vaginal, oral, and anal), age of sex debut, life time opposite sex partners and same sex partners, age of first oral sex, having new sex partners in past 12 months, and frequencies of sex (vaginal/anal). Additional variables for status of sexual transmitted diseases, such as ever been told by doctor of having HPV, herpes, genital wart, gonorrhea, and chlamydia, were included as well.

2.3 Laboratory Methods

2.3.1 Oral Specimen Collection

Oral sample were collected by a 30-second oral rinse and gargle with SCOPE™ mouthwash (10ml) at MEC. Subjects alternate a series of three, five second rinses and five second gargles and then expectorated into a sterile collection tube. De-identified specimens were stored at 4°C and shipped to Oral HPV Detection Laboratory (OHD), Polaris Innovation Center, Polaris, Ohio, for HPV detection [44].

2.3.2 DNA Isolation

Cells in samples were collected by centrifuge and lysed in Puregene cell lysis solution followed by DNase-free RNase and proteinase K digestion using the Qiagen Virus/Bacteria Midi

kit and Pathogen Complex 800 program on the Qiasymphony SP instrument (Qiagen). DNA was eluted in 60 μ L of AVE buffer and quantified using NanoDrop 8000 (ThermoFisher Scientific) [44, 45].

2.3.3 HPV Genotyping Test

HPV detection and genotyping were performed using Linear Array (LA) genotyping assay (Roche Diagnostics). The assay detects 37 HPV types using HPV L1 consensus polymerase chain reaction (PCR) with PGMY09/11 primer pools and primers for β -globin as an internal control for multiplex amplification. PCR product were denatured in NaOH and hybridized to an immobilized HPV probe array for detection of β -globin and genotyping of 37 HPV types, including HRHPV types (16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73, 82) and LRHPV types (6, 11, 40, 42, 54, 55, 61, 62, 64, 67, 69, 70, 71, 72, 81, 82 subtype IS39,83, 84, 89 [cp6108]) (Roche Linear Array HPV Genotyping Test; Roche Molecular Systems) [46]. The XR probe on the LA HPV strip had cross reaction with HPV 33, 35, 52 and 58 and required further test. Samples positive for HPV XR, 33, 35, and 58 were tested for HPV 52 using type-specific TaqMan real time PCR to determine HPV 52 status. Samples negative for both β -globin and HPV were considered as insufficient and were excluded from further analysis [47]. All procedures were performed according to the NHANES quality assurance and quality control (QA/QC) protocols that meet the 1988 Clinical Laboratory Improvement Act (CLIA) mandates [44].

2.4 Statistical Analysis

NHANES 2013-2014 MEC respondents aged 18-69 years with valid oral HPV test results (n=4724) were included in the analysis. The NHANES 2013-2014 MEC sample weights were used to account for the complex survey design and to generate unbiased estimate of prevalence.

Data manipulation, variable creation, data merging and analysis were performed using SAS version 9.3 (SAS Institute). SAS survey procedures were used for all statistical analysis to account for complex survey features such as stratification, clustering, and weights. Variance was estimated using Taylor Series Linearization (TSL) method to account for the complex cluster survey design. To adjust for nonresponses and account for the unequal probabilities of selection, all estimates were analyzed using 2-year weighted variables constructed from the MEC examination weights provided with the dataset (WTMEC2YR) [41]. The dataset has 15 strata with 2 clusters per strata (SDMVSTR, SDMVPSU). HPV prevalence with Confidence Intervals (CIs) were estimated using a logit transformation and generated using SAS `surveymeans` or `surveyfreq` procedures.

Overall oral HPV prevalence with any types was estimated by sociodemographic and sexual behavior characteristics. Medical conditions, including status of sexual transmitted diseases and vaccination history of HPV, were included as well. The prevalence of HR- and LR- HPV infections were estimated by the same characteristics. We also evaluated HPV prevalence within characteristics such as gender and race stratified by age group. Weighted estimates of prevalence with 95% Confidence Interval (95% CI) were presented. Rao-Scott χ^2 test was used to examine the independence between HPV prevalence and selected characteristics for complex survey design, in which test are no longer follow the same χ^2 distribution under null hypothesis [44, 48, 49].

Logistics regression models were used to explore the relationship between HPV infection and demographic or behavioral characteristics. Characteristics that was associated with HPV infection in univariate analysis was selected for multivariate logistic analysis. Multiple models were constructed to evaluate the association between HPV infection (overall HPV, HRHPV,

LRHPV, and HPV16) and selected characteristics. Some characteristics with collinearity such as sexual behaviors were analyzed separately. Interactions between characteristics of interest were examined by pairwise logistic regression tests. Odds Ratios (OR) with 95% CI were used to evaluate risks. Wald χ^2 test was used to evaluate associations between the selected characteristics and HPV infection. Two-tailed P values less than 0.05 was considered as statistically significant for all analysis in this study. Logistics regression analysis was conducted using surveylogistic procedure in SAS 9.3.

3 RESULTS

3.1 Overall Oral HPV Prevalence

The weighted prevalence for oral HPV infection was estimated based on HPV DNA positivity using 4724 participants in the dataset. The overall oral HPV prevalence of any types in U.S. population aged 18-69 years was 7.18% (95% CI, 5.9%-8.4%). This represented 14,219,994 people with this condition, based on the weight variable WTMEC2YR that summed to population at that time. The HPV infection prevalence was estimated as 4.5% (95% CI, 3.05%-4.97%) for HRHPV and 3.5% (95% CI, 3.09%-4.58%) for LRHPV, respectively. The prevalence of type-specific oral HPV infections was shown in Figure 1. 36 of the 37 types covered by LA assay, except for HPV type 64, were detected from the NHANES samples collected during 2013-2014. The prevalence of individual HPV types ranged from 0.005% to 0.96%. The most prevalent HPV type was HPV16 (0.97%; 95% CI, 0.7%-1.23%). Of the 359 HPV positive samples among the 4724 subjects, 281 (78.3%) were single HPV infection, and 78 (21.7%) were multiple infections ranging from 2 to 6 types, representing 11 and 3 million people, respectively (Table 2).

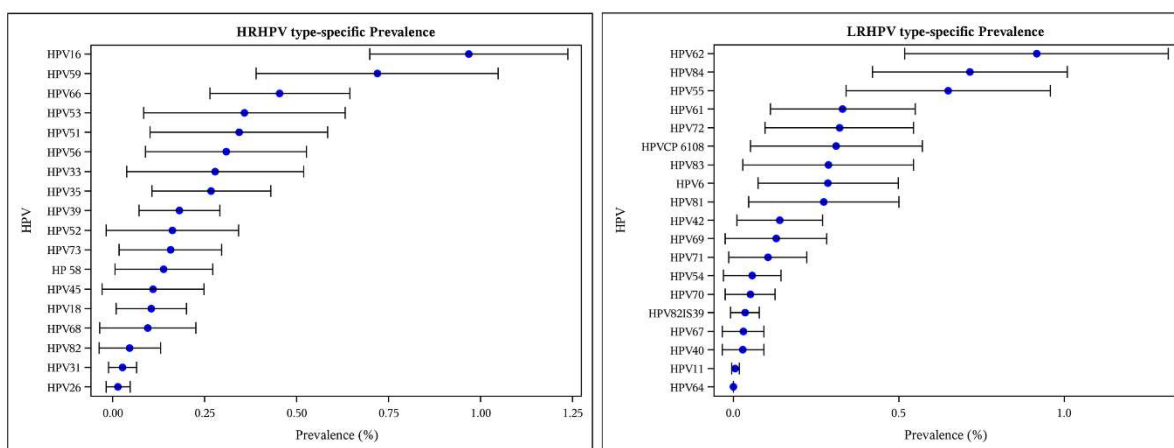


Figure 1 Weighted prevalence of Oral HPV infection by individual HPV types in U.S. population aged 18-69 years

Table 2 Frequencies of single and multiple infections of oral HPV in U.S. population aged 18-69 years

No. of HPV types	Frequency	Weighted Frequency	Percent
1	281	11,157,651	5.63
2	62	2,111,244	1.07
3	12	586,275	0.30
4	3	350,246	0.18
6	1	14,578	0.0074
0	4,365	183,829,791	92.80
Total	4,724	198,049,785	100.0

3.2 Oral HPV Prevalence by Demographic and Life Style Characteristics

The prevalence of oral HPV infection was analyzed by demographic characteristics (Table 3). Bivariate analysis revealed that the distribution of any HPV infection was significantly different across several characteristics including gender, country of birth, race, education, marital status, alcohol use, and tobacco use (Table 3). The prevalence of any oral HPV infection was significantly higher (3.4 folds) in males (11.1%; 95% CI, 8.8%-13.3%) than that in females (3.3%; 95% CI, 4.3%-6.8%). Non-Hispanic Black had the highest overall prevalence (9.5%; 95% CI, 7.8%-11.3%), followed by Non-Hispanic White (7.4; 95% CI, 5.9%-8.9%). Non-Hispanic Asian had the lowest prevalence (2.9; 95% CI, 7.8%-11.3%). Overall HPV prevalence was lower among married people (6.4%; 95% CI, 4.9%-7.9%), compared to those who were not in marriage status (9.6% for widow/divorced/separated; 8.3% for never married). People born in U.S. had higher prevalence (7.5%; 95% CI, 6.1%-9%) than those who were born in other countries (5.1%; 95% CI, 3.3%-7%). Significant difference in distribution across education levels was observed as well (Table 3). Overall prevalence of any HPV infection appeared to be associated with alcohol and tobacco use. HPV prevalence was positively correlated to the number of alcohol consumed per day (Table 3). People who drank 4 alcohols per day had HPV prevalence (13%; 95% CI,

9.2%-16.8%) nearly twice as high as those who didn't drink (6.7%; 95% CI, 5%-8.4%). HPV prevalence was higher in people who smoked at time of interview (12.9%; 95% CI, 9.7%-16.1%) than non-smokers (7.1%; 95% CI, 5.1%-9.2%). Similar to alcohol use, HPV prevalence was positively correlated to the number of cigarette consumed per day (Table 3). The age starting smoking seemed to have little impact on oral HPV prevalence. Oral HPV prevalence seemed to be higher in people with decayed teeth and gum disease. However, these variables had low number participants who responded "No" to the questions and did not had P value calculated. Oral HPV prevalence appeared to be lower in people with HPV vaccination compared to those who did not received vaccination (Table 3). The prevalence of HRHPV infection and LRHPV infection each showed similar distribution patterns across characteristics as any HPV infection. Only two variables, gender and using more than 4 alcohol/day, showed significantly different prevalence distribution patterns for all three categories of HPV ($P < 0.05$) (Table 3).

Table 3 Weighted prevalence of oral HPV among U.S. adults (18-69 years) by demographic characteristics, National Health and Nutrition Examination Survey, 2013-2014

Characteristic	Any HPV			HRHPV (with or without LRHPV)		LRHPV (with or without HRHPV)	
	Sample size	Prevalence (95%CL)	P-value ^a	Prevalence (95%CI)	P-value ^a	Prevalence (95%CI)	P-value ^a
Total	4724	7.1(5.9-8.4)		4(3-4.9)		3.8(3-4.5)	
Age							
18-24	805	4.9(2.7-7.2)	0.5871	2.9(1-4.9)	0.5034	2(1-3)	0.3043
25-29	410	6.4(3.3-9.6)		4.1(1-7.2)		2.9(1.2-4.6)	
30-34	473	6.1(3.5-8.7)		3.6(1.6-5.6)		2.8(0.8-4.7)	
35-39	426	6.9(4.5-9.3)		3.2(1.1-5.4)		4(1.3-6.6)	
40-44	479	7.6(3.5-11.6)		3.7(0.9-6.5)		4.5(1.8-7.3)	
45-49	437	7.2(3.1-11.3)		4.5(1.6-7.3)		3.3(0.9-5.6)	
50-54	433	9.4(5-13.8)		6.4(3.3-9.6)		3.9(1.4-6.4)	
55-59	415	9.1(4.8-13.5)		5(1.9-8.1)		5.1(2.5-7.6)	
60-64	476	8.3(2.7-13.9)		2.8(1-4.5)		6.8(1.9-11.8)	
65-70	370	6.2(3.1-9.4)		3.6(1.1-6)		4(1-6.9)	
Gender			*		*		*
M	2278	11.1(8.8-13.3)	<.0001	6.6(4.9-8.3)	<.0001	5.5(4.3-6.8)	<.0001
F	2446	3.3(2.5-4.1)		1.4(0.9-1.9)		2.1(1.4-2.8)	
Country of birth			*		*		
US	3405	7.5(6.1-9)	0.0163	4.4(3.2-5.5)	0.002	3.9(2.9-4.8)	0.5894
Others	1316	5.1(3.3-7)		2(0.9-3.1)		3.4(2-4.8)	
Missing	3	0(0-0)		0(0-0)		0(0-0)	
Race^c			*				*
Mexican American	708	6.4(3.8-9.1)	0.0044	2.9(1.1-4.8)	0.1055	4(2.1-6)	0.0164

Other Hispanic	434	5(2.8-7.1)		3.3(1.6-5.1)		2.3(0.7-3.9)	
NHW	1865	7.4(5.9-8.9)		4.3(2.9-5.6)		3.7(2.9-4.6)	
NHB	1005	9.5(7.8-11.3)		4.7(3.8-5.6)		5.7(4-7.4)	
NH Asian	547	2.9(1.7-4)		1.9(0.8-3)		1.2(0.5-1.8)	
Other	165	6.5(2.3-10.7)		2.9(0.2-5.5)		4.6(1.2-8)	
Education			*				*
Less than high school	295	6.7(3.4-9.9)	0.0003	2.9(0.3-5.5)	0.0633	3.8(1.8-5.7)	0.0051*
High school or equivalent	1558	9.8(7.5-12.2)		4.9(3.8-6.1)		5.9(4-7.9)	
College or above	2541	6.1(5-7.3)		3.7(2.6-4.8)		2.9(2.1-3.7)	
Missing	330	2.7(0.9-4.4)		1.2(0-2.5)		1.4(0.1-2.6)	
Income poverty ratio							
<1.0	1076	8.5(6.6-10.3)	0.5978	4.3(3.3-5.3)	0.814	5(3.6-6.4)	0.1143
≥1.0 to <2.0	1067	7.2(4.4-10)		3.7(2-5.3)		4.2(2.1-6.2)	
≥2.0 to <3.0	576	6.5(3-10.1)		4.7(2.1-7.2)		2(0.7-3.4)	
≥3.0	1651	6.5(4.4-8.7)		4(2.5-5.5)		3.2(2-4.5)	
Missing	354	9.2(3.5-14.9)		2.3(1-3.5)		7.3(1.5-13.1)	
Marital status			*				*
Married/living with partners	2628	6.4(4.9-7.9)	0.0294	3.7(2.5-4.9)	0.3151	3.2(2.4-4)	0.0148*
Widow/divorced/separated	792	9.6(6.9-12.4)		5(2.9-7.1)		5.8(3.7-8)	
Never married	974	8.3(6.5-10.2)		4.5(3-6)		4.6(3.1-6.1)	
Missing	330	2.6(0.9-4.4)		1.2(0-2.5)		1.4(0.1-2.6)	
No. Alcohol/day			*				*
No alcohol	1003	5.8(4.2-7.4)	0.0441	2.3(1.6-3)	0.055	3.7(2-5.3)	0.048
≥2 to <4/day	1632	7.3(5.1-9.5)		4.5(2.9-6.1)		3.6(2.6-4.5)	
≥4 to <8/day	367	11.8(6.2-17.3)		6.8(2.6-11)		6.6(3.3-9.9)	
>8/day	102	14.9(5.8-24)		6.2(0.1-12.4)		9.4(4.3-14.4)	
Missing	1620	6(4.1-7.9)		3.6(2.3-5)		2.9(2-3.9)	
More than 4 alcohol/day			*		*		*
Yes	610	13(9.2-16.8)	0.004	7.3(4.8-9.8)	0.0106	6.8(3.5-10)	0.0327
No	3097	6.7(5-8.4)		3.9(2.5-5.2)		3.5(2.8-4.2)	
Missing	1017	4.3(2.6-6)		1.7(0.7-2.8)		2.7(1.6-3.8)	
Currently smoking			*		*		
Yes	1052	12.9(9.7-16.1)	0.0004	8.2(5.7-10.7)	<.0001	6(4.4-7.6)	0.3001
No	908	7.1(5.1-9.2)		2.8(1.7-3.9)		4.5(2.4-6.7)	
Missing	2764	5(4-6.1)		2.8(2.3-3.3)		2.7(1.8-3.6)	
Age starting smoking regularly							
Never smoke	94	9.1(1.9-16.4)	0.4011	3.9(-0.7-8.6)	0.3968	5.1(-0.8-11.2)	0.432
≤18	1289	9.1(6.3-11.9)		5.1(3-7.3)		4.7(3.2-6.1)	
19-30	528	12(8.5-15.5)		6.9(4.3-9.5)		6.3(3.7-8.8)	
>30	43	18.2(-1.4-37.9)		5.6(0-11.3)		14.1(-4.9-33.1)	
Missing	2770	5(4-6.1)		2.8(2.3-3.3)		2.7(1.8-3.6)	
No. Cigarette/day							
≤1	84	8.3(1-15.6)	0.0855	3.9(-0.6-8.6)	0.2856	4.3(-0.7-9.3)	0.1344
≥2 to <10/day	405	9.9(5-14.8)		6.3(1.6-11)		4.1(2.1-6.2)	
≥10 to <20/day	315	15.3(12.1-18.6)		10.5(6.4-14.6)		6.4(3.5-9.4)	
≥20/day	236	15.6(9.2-21.9)		9.7(3.8-15.5)		8.4(5.1-11.8)	
Missing	3684	5.6(4.6-6.6)		2.8(2.2-3.4)		3.2(2.5-3.9)	
Decayed teeth							
Yes	1590	9.7(7.2-12.2)	0.0961	4.6(2.8-6.3)	NA ^b	5.8(4-7.5)	0.3822
No	10	2.9(-3.5-9.4)		0(0-0)		2.9(-3.5-9.4)	
Missing	3124	6.1(4.9-7.2)		3.7(2.8-4.6)		2.9(2.3-3.6)	
Gum disease							
Yes	1532	9.9(6.9-12.8)	0.9917	4.6(2.9-6.4)	0.6309	6.2(4.2-8.2)	NA ^b

No	7	9.8(-13-32.6)		9.8(-13-32.6)		0(0-0)	
Missing	3185	6.1(4.9-7.3)		3.7(2.7-4.7)		2.9(2.3-3.5)	
HPV vaccination							
Yes	390	5.2(3.7-6.6)	0.0573	2.8(1.2-4.3)	0.0601	2.4(0.8-4.1)	0.2218
No	3279	7.3(5.4-9.3)		4.3(3-5.6)		3.6(2.6-4.7)	
Missing	1055	7.1(3.9-10.4)		3.2(1.7-4.7)		4.9(1.8-8)	

Note: CI: Confidence Interval.

^aRao-Scott χ^2 test P values for distribution of prevalence among characteristics, missing values were excluded from the analysis.

^bP values were not determined due to lack of samples in one level of the characteristics. * P<0.05.

^cNHW, Non-Hispanic White, NHB, Non-Hispanic Black, NH Asian, Non-Hispanic Asian.

HPV prevalence over age was shown in Figure 2 and Figure 3. The prevalence for any HPV infection showed a trend to increase by age and reached to a peak (9.4%; 95%CI, 5%-13.8%) in people aged 50-54 and decreased gradually in later ages (Figure 3A and Table 3). Notably, the prevalence of HRHPV and LRHPV by age appeared to follow bimodal patterns. The first peak of infection for HRHPV were observed among people aged 25-29 (4.1%; 95%CI, 1%-7.2%), and the second peak appeared in people aged 50-54 (6.4%; 95%CI, 3.3%-9.6%) (Figure 3B). Unlike HRHPV, peaks for LRHPV prevalence appeared in later ages with first peak in people aged 40-44 (4.5%; 95%CI, 1.8%-7.3%) and second peak in people aged 60-64 (6.8%; 95%CI, 1.9%-11.8%) (Figure 3C). The prevalence of HPV16, the most oncogenic type, was shown in Figure 4. People aged 45-49 years, 55-59 years, and 65-69 years showed higher prevalence compared to people in other age groups.

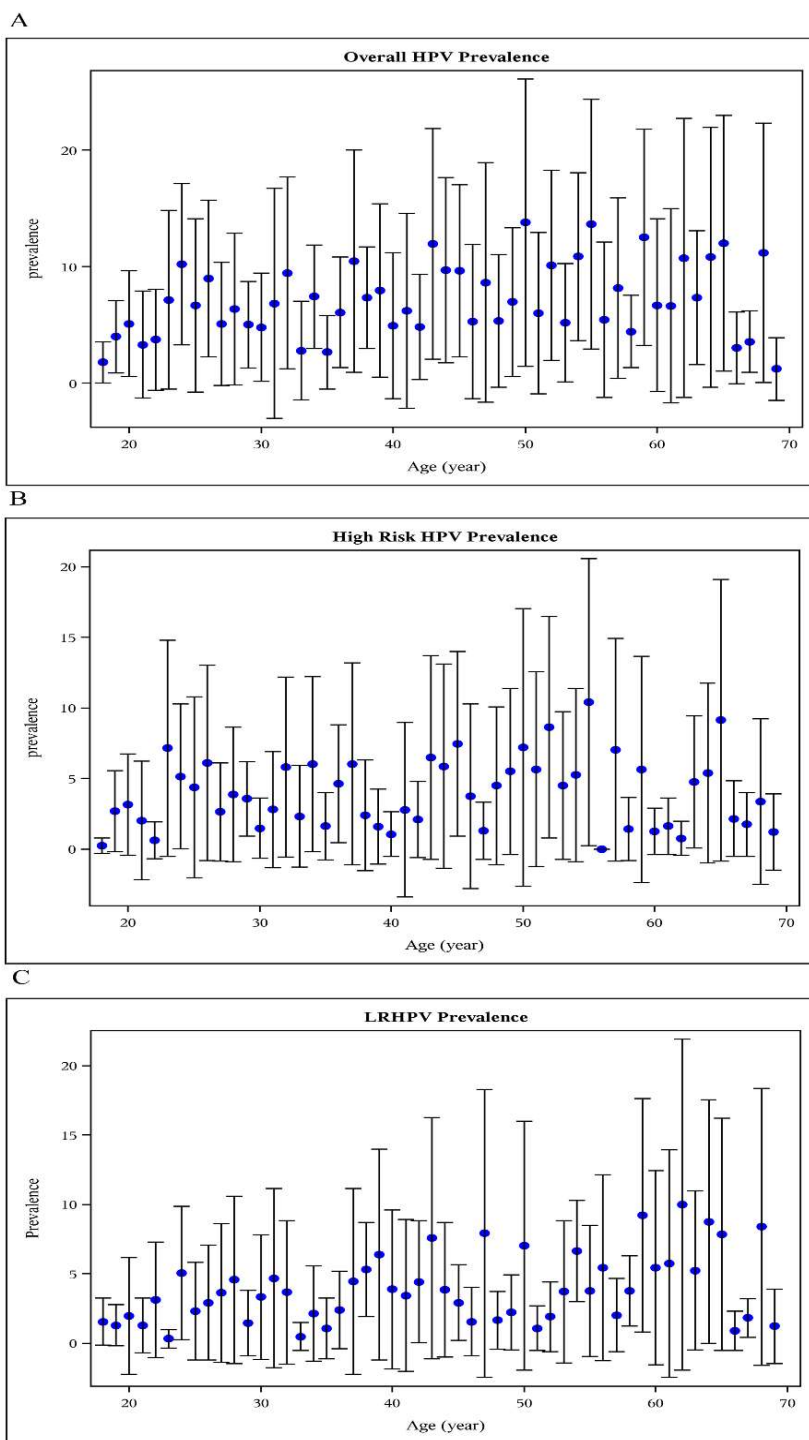


Figure 2 Weighted oral HPV prevalence by age (14-69 years)
 (A) Overall Oral HPV prevalence in percentage of any types; (B) Oral HPV prevalence in percentage of high-risk types; (C) Oral HPV prevalence in percentage of low-risk types. Error bars indicate 95% CI.

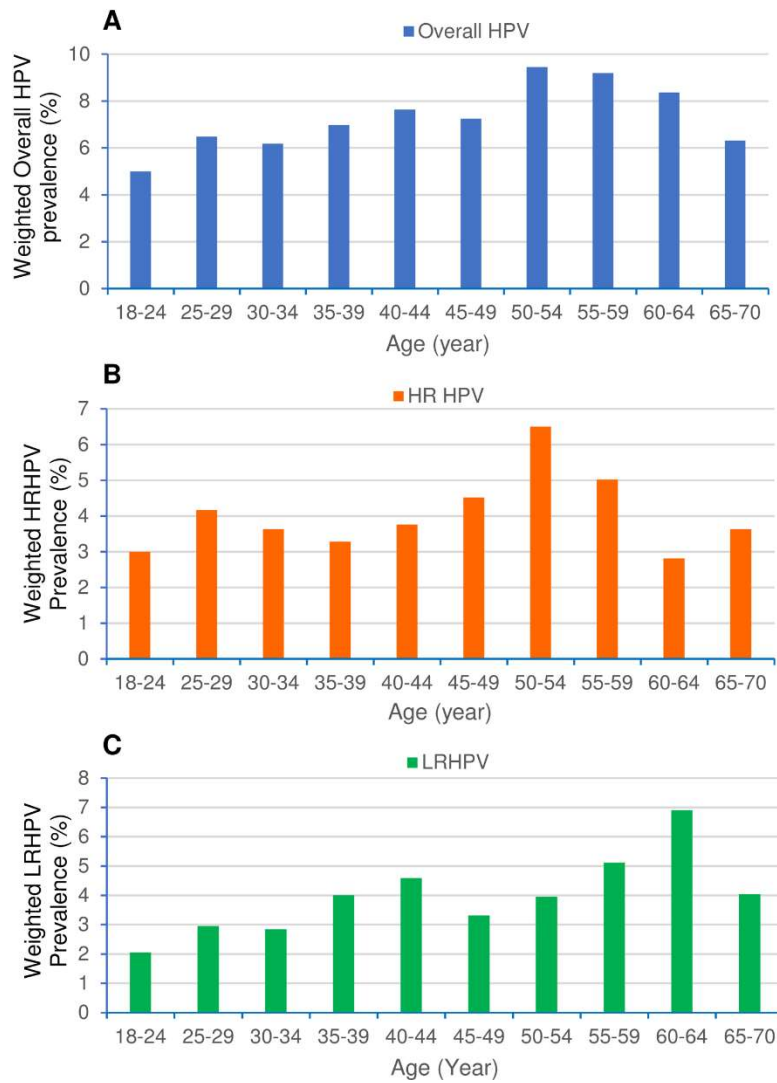


Figure 3 Weighted oral HPV prevalence by age group
 (A) Overall Oral HPV prevalence any types; (B) Oral HPV prevalence of high-risk types; (C) Oral HPV prevalence of low-risk types.

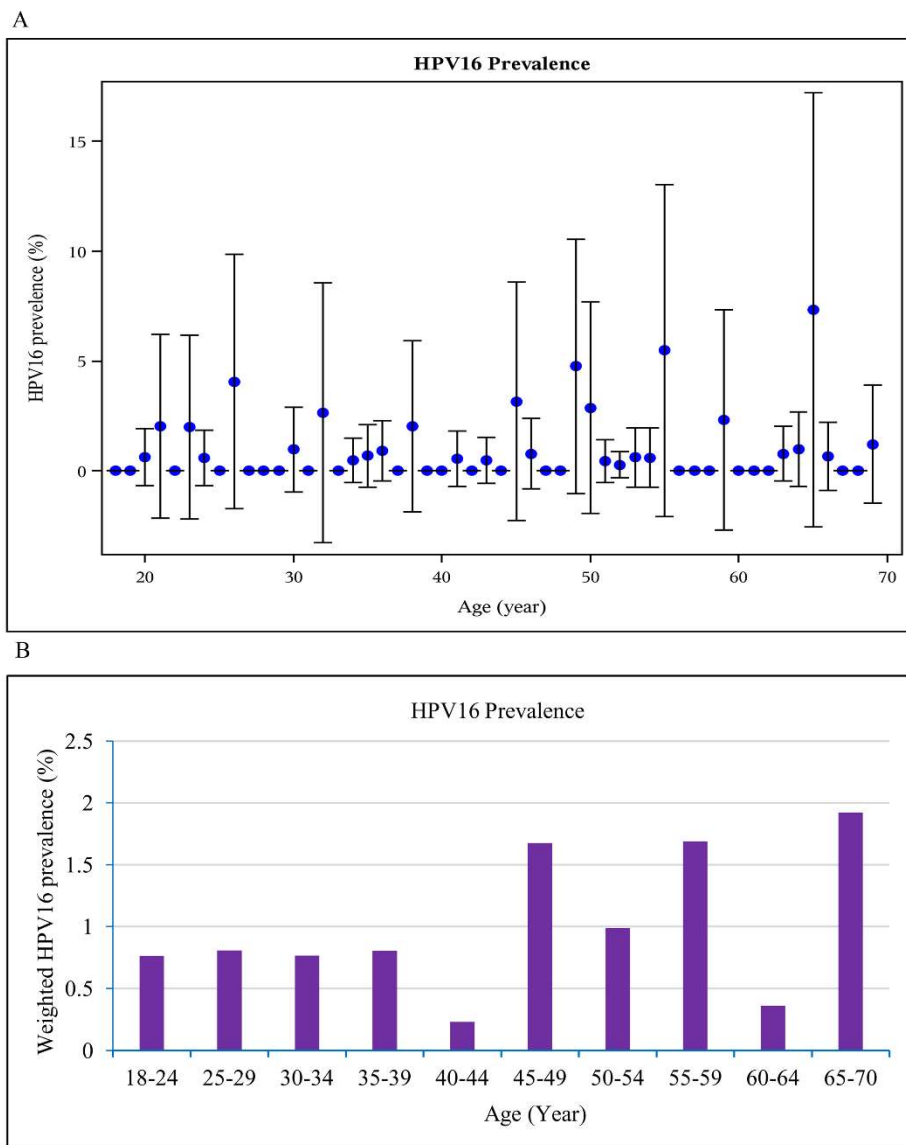


Figure 4 Weighted prevalence of HPV16 by age

3.3 Oral HPV Prevalence by Sexual History

The oral HPV prevalence classified by sexual behavior and sexual transmitted diseases was shown in Table 4. The distribution of any HPV infection was significantly different for 6 of the 10 characteristics ($p < 0.05$) including ever had vaginal/oral/anal sex, age of sex debut, number of lifetime opposite or same sex partners. HRHPV distribution differed significantly across the

categories of four factors, ever had vaginal/oral/anal sex, ever had anal sex, age of sex debut, and number of lifetime opposite sex partners. For LRHPV, only 3 factors, ever had anal sex, age of sex debut, and number of lifetime opposite sex partners, showed significant difference in the infection distribution across categories of these factors (Table 4).

The prevalence of Oral HPV (any types, HRHPV, and LRHPV) was higher in people who had sex experience (vaginal, oral and anal), compared to those who reported no sex experience. People who started sex before 18 years of age had higher overall HPV prevalence (9.8%; 95%CI, 7.6%-11.9%) than those who started sex at 18 years or older (3.7%; 95%CI, 2.7%-4.7%). Among those who had active sexual activities, the HPV prevalence was correlated with number of opposite sex partners. People who had more than 10 partners had oral HPV prevalence of 14.6% (95%CI, 9.9%-19.3%), 9.5% (95%CI, 6.4%-12.6%), and 6.8% (95%CI, 4.4%-9.1%) for any type, HR-, and LR- HPV, respectively (Table 4). Interestingly, people who had same sex partners showed a different pattern. The highest prevalence was observed in people who had 5-10 same sex partners (32.7%; 95%CI, 8.2%-57.2%) for any HPV, 16.2% (95%CI, 5.9%-38.4%) for HRHPV, and 16.4 % (95%CI, -4.7%-37.7%) for LRHPV (Table 4). We observed relative high prevalence in people who reported having sex once in past 12 months (12.1 % for any HPV, 6.4% for HRHPV, and 6.2% for LRHPV). For people who had sex between 2 to 103 times in past 12 months had lower prevalence compared to those who had sex only once or more frequently (>103 times) (Table 4). Whether a person had new partners in past 12 months seemed to have little impact on HPV infection. Sexual transmitted disease status informed by doctors had mixed results of HPV prevalence (Table 4). HPV prevalence was higher in people who had chlamydia infection (9.4%) than those who did not. Prevalence of HPV

infection among people who did not respond (missing) to questions about their sexual history was reported as well to reflect the overall prevalence in the population (Table 4).

Table 4 Weighted prevalence of oral HPV among U.S. adults aged 18-69 years by sexual behavior, National Health and Nutrition Examination Survey, 2013-2014

Characteristic	Any HPV			HRHPV (with or without LRHPV)		LRHPV (with or without HRHPV)	
	Sample Size	Prevalence (95%CI)	P value ^a	Prevalence (95%CI)	P value ^a	Prevalence (95%CI)	P value ^a
Total	4724	7.1(5.9-8.4)		4(3-4.9)		3.8(3-4.5)	
Ever had vaginal/anal/oral sex			*	*			
Yes	4115	7.2(5.8-8.7)	0.003	4.1(3-5.2)	<.0001	3.8(3-4.6)	0.1298
No	239	2.5(0-5)		0.4(-0.4-1.3)		2(0-4.2)	
Missing	370	8.5(4.4-12.6)		3.7(0.6-6.9)		4.7(2.1-7.4)	
Ever had vag. sex							
Yes	4030	7.1(5.7-8.5)	0.5906	4(3-5.1)	0.3859	3.7(2.9-4.6)	0.9796
No	324	5.9(1.3-10.6)		3(0.1-5.8)		3.7(0.4-6.9)	
Missing	370	8.4(4.2-12.6)		3.6(0.4-6.8)		4.7(2-7.4)	
Ever had oral sex							
Yes	3362	7.6(6-9.1)	0.0261	4.3(3.2-5.5)	0.0575	4(3.1-5)	0.0197
No	982	4.8(2.5-7.1)		2.6(0.9-4.4)		2.2(0.9-3.5)	
Missing	380	7.4(4.5-10.3)		2.8(0.4-5.1)		4.6(2-7.3)	
Ever had anal sex			*		*		
Yes	1580	9(7.1-10.9)	0.0042	5.4(3.8-7.1)	0.0098	4.5(3.5-5.5)	0.0524
No	2763	5.9(4.1-7.6)		3.1(1.9-4.4)		3.2(2.1-4.3)	
Missing	381	7.5(4.6-10.4)		2.8(0.4-5.2)		4.7(2-7.3)	
Age of sex debut			*		*		*
<18	2403	9.8(7.6-11.9)	<.0001	5.5(3.6-7.4)	0.0014	5.2(4.1-6.2)	<.0001
≥18	1704	3.7(2.7-4.7)		2.3(1.6-2.9)		1.8(1.1-2.5)	
Missing	617	5.7(3.9-7.5)		1.9(0.4-3.5)		3.7(2.2-5.3)	
No. Life time heterosexual partners			*		*		*
0 partner	272	5.5(0.4-10.6)	<.0001	2.2(-0.8-5.2)	<.0001	3.5(0-7)	0.0005
1 partner	642	1.8(0.2-3.3)		0.6(-0.1-1.4)		1.1(0-2.3)	
>1 to ≤5	1448	3.6(2.3-4.9)		1.7(1-2.5)		2.1(1-3.2)	
>5 to ≤10	903	7.2(4.9-9.6)		3.6(2-5.3)		4.3(2.5-6.1)	
>10	1101	14.6(9.9-19.3)		9.5(6.4-12.6)		6.8(4.4-9.1)	
Missing	358	7.7(4.6-10.9)		2.8(0.3-5.3)		4.9(2.1-7.7)	
No. of Life time homosexual partners			*				
1 partner	101	7.6(1.4-13.9)	0.0427	3.5(-1.4-8.6)	0.1014	4(-0.8-9)	0.3464
>1 to ≤5	142	5.6(1.4-9.9)		3.2(0.3-6.1)		2.9(-0.5-6.3)	
>5 to ≤10	23	32.7(8.2-57.2)		16.2(-5.9-38.4)		16.4(-4.7-37.7)	
>10	35	28.5(6.7-50.4)		23.2(1.5-45)		6.2(-0.5-13)	
Missing	4423	6.8(5.6-8)		3.7(2.9-4.6)		3.7(2.9-4.5)	
Age of 1st oral sex							
<18	529	5(1.6-8.3)	0.2854	3.7(0.7-6.7)	0.1544	1.5(0.1-2.9)	0.5957
≥18	995	3.3(2-4.6)		1.6(0.9-2.3)		2(0.9-3.1)	
Missing	3200	8.9(7.3-10.4)		4.8(3.7-5.9)		4.9(3.8-5.9)	
New sex partner in past year							
Yes	575	9.9(6.4-13.4)	0.0741	6.4(3.4-9.3)	0.1042	4.7(2.7-6.6)	0.2164
No	2353	6.9(4.6-9.2)		4(2.6-5.5)		3.3(1.9-4.6)	

Missing	1796	6.4(4.4-8.5)		2.9(1.9-3.9)		4.3(2.3-6.3)	
Frequency of vaginal/anal sex in past year							
Never	89	6.4(1-11.8)	0.3554	4.4(0.9-7.8)	0.6663	3.6(-0.1-7.5)	0.7701
Once	129	12.1(4.1-20.1)		6.4(-0.3-13.2)		6.2(0.1-12.2)	
2-11 times	683	7(3-11)		4.2(1-7.3)		3.4(1.1-5.7)	
12-51 times	984	7(4.1-9.9)		3.9(1.5-6.4)		3.6(2.4-4.8)	
52-103 times	612	6.6(3.6-9.5)		4.3(2-6.6)		2.9(1.2-4.7)	
104-364 times	370	9.1(4.7-13.4)		5.7(1.9-9.4)		3.8(1.4-6.3)	
≥365 times	39	19.8(3.7-36)		12.8(-2.1-27.8)		7(-2.7-16.8)	
Missing	1818	6.4(4.4-8.5)		2.9(2-3.9)		4.2(2.2-6.2)	
Herpes status by doctor					*		
Yes	124	5.1(0.9-9.2)	0.2321	1.7(-0.4-3.9)	0.0226	3.3(-0.4-7.1)	0.9533
No	3228	7.3(5.3-9.3)		4.5(3.1-5.9)		3.4(2.4-4.4)	
Missing	1372	6.8(4.3-9.4)		2.7(1.5-3.9)		4.9(2.4-7.5)	
Genital warts status by doctor			*		*		
Yes	124	3.5(1.2-5.8)	0.0003	2.1(0.4-3.9)	0.0183	1.9(-0.1-4)	0.0979
No	3229	7.4(5.4-9.4)		4.5(3.1-5.8)		3.5(2.4-4.5)	
Missing	1371	6.8(4.3-9.4)		2.7(1.5-3.9)		4.9(2.4-7.5)	
Chlamydia status by doctor							
Yes	45	9.4(1.1-17.6)	0.5374	4.5(-0.3-9.4)	0.9435	6.3(-1.1-13.8)	0.3953
No	3307	7.2(5.3-9.1)		4.4(3-5.7)		3.4(2.4-4.4)	
Missing	1372	6.8(4.3-9.4)		2.7(1.5-3.9)		4.9(2.4-7.5)	

Note: CI: Confidence Interval.

^aRao-Scott χ^2 test P values for distribution of prevalence among characteristics, missing values were excluded from the analysis.

* P<0.05.

3.4 Univariate Analysis of Factors Associated with HPV Infection

The association between any type HPV infection and 26 demographic and behavioral characteristics was evaluated with unadjusted odds ratio (OR) from univariate analysis (Table 5). Of the 26 characteristics, 16 were associated with HPV infection. Factors with p value less than 0.0001 included age, gender, race, education level, age of sex debut, and number of life time sex partners. Categories or levels within each associated characteristic conferred various degrees of potential risk on HPV infection. The unadjusted ORs were correlated with prevalence of HPV infection presented in Table 3. The factors conferring high risks included males, heavy alcohol

and tobacco use, having sex experience, high number of life time sexual partners, and high frequency of sex.

Table 5 Univariate analysis of characteristics associated with oral HPV infection (any type) in U.S. population aged 18-69 years

Characteristic	Unadjusted OR (95%CI)	P value
Age		
18-24	1[reference]	
25-29	1.32(0.69-2.52)	<0.0001
30-34	1.25(0.71-2.19)	
35-39	1.42(1.01-2.02)	
40-44	1.57(0.89-2.76)	
45-49	1.48(0.82-2.67)	
50-54	1.98(1.36-2.9)	
55-59	1.92(0.95-3.89)	
60-64	1.73(0.65-4.63)	
65-70	1.28(0.56-2.89)	
Gender		
	0(0-0)	
M	3.55(2.72-4.65)	<0.0001
F	1[reference]	
Country of birth		
US	1.5(1.03-2.17)	0.0321
Others	1[reference]	
Race		
Mexican American	0.86(0.54-1.36)	<.0001
Other Hispanic	0.66(0.43-1)	
NHW	1[reference]	
NHB	1.32(1.07-1.62)	
NH Asian	0.37(0.25-0.55)	
Other	0.87(0.47-1.6)	
Education		
Less than high school	1[reference]	
High school or equivalent	1.52(0.9-2.56)	0.0001
College or above	0.91(0.53-1.54)	
Inpoivr		
<1.0	1[reference]	
≥1.0 to <2.0	0.84(0.54-1.31)	0.2682
≥2.0 to <3.0	0.76(0.44-1.31)	
≥3.0	0.75(0.56-1.02)	
Marital status		
Married/living with partners	0.75(0.6-0.94)	0.0166
Widow/divorced/separated	1.16(0.81-1.67)	
Never married	1[reference]	
No. Alcohol/day		
No alcohol	1[reference]	
≥2 to <4/day	1.28(0.79-2.08)	0.0027
≥4 to <8/day	2.16(1.11-4.2)	
>8/day	2.82(1.38-5.77)	
More than 4 per day		
Yes	2.08(1.32-3.25)	0.0014
No	1[reference]	
Currently smoking		
Yes	1.91(1.35-2.71)	0.0002
No	1[reference]	

Age starting smoking regularly		
Never smoke	1[reference]	
≤18	0.99(0.42-2.34)	0.4369
19-30	1.36(0.53-3.44)	
>30	2.2(0.69-7.02)	
No. Cigarettes/day		
0	0(0-0)	
≤1	1[reference]	
≥2 to <10/day	1.22(0.44-3.33)	0.0741
≥10 to <20/day	2(0.8-5.04)	
≥20/day	2.04(0.86-4.8)	
Ever had vaginal/anal/oral sex		
Yes	3.03(1.14-8.01)	0.0254
No	1[reference]	
Ever had vaginal sex		
Yes	1.21(0.56-2.6)	0.6182
No	1[reference]	
Ever had oral sex		
Yes	1.62(0.99-2.65)	0.0531
No	1[reference]	
Ever had anal sex		
Yes	1.57(1.14-2.17)	0.0055
No	1[reference]	
Age of sex debut		
Yes	2.78(2.01-3.85)	<.0001
No	1[reference]	
No. Life time heterosexual partners		
0 partner	1[reference]	
1 partner	0.32(0.11-0.91)	<.0001
>1 to ≤5	0.64(0.23-1.81)	
>5 to ≤10	1.33(0.43-4.07)	
>10	2.93(1.28-6.72)	
No. of Life time homosexual partners		
1 partner	1[reference]	
>1 to ≤5	0.72(0.24-2.17)	<.0001
>5 to ≤10	5.85(1.28-26.63)	
>10	4.81(1.24-18.55)	
Age of 1st oral sex		
<18	1.52(0.77-2.99)	0.2202
≥18	1[reference]	
New sex partner in past year		
Yes	1.47(0.99-2.18)	0.0547
No	1[reference]	
Frequency of vaginal/anal sex in past year		
never	1[reference]	
once	2.01(0.69-5.81)	0.0125
2-11 times	1.1(0.34-3.5)	
12-51 times	1.1(0.45-2.68)	
52-103 times	1.03(0.42-2.53)	
104-364 times	1.46(0.62-3.44)	
≥365 times	3.62(0.89-14.61)	
HPV status by doctor		
Yes	1.37(0.52-3.6)	0.5207
No	1[reference]	

Herpes status by doctor			
	Yes	1(1.06-0.3)	0.3011
	No	1[reference]	
Genital warts status by doctor			
	Yes	0.46(0.27-0.78)	0.0044
	No	1[reference]	
Chlamydia status by doctor			
	Yes	1.32(0.61-2.85)	0.4657
	No	1[reference]	

3.5 Multivariable Analysis of Factors Associated with Oral HPV Infection

In multivariable analysis, ORs for HPV infection were adjusted for 9 characteristics, age, gender, country of birth, race, education, marital status, number of alcohols consumed per day, whether or not smoke at the time of interview, and age of sex debut. Categories in each characteristic showed different risks for HPV infection (any types, HRHPV, and LRHPV) (Table 6). Four characteristics (age, gender, whether or not smoke at the time of interview, and age of sex debut) were independently associated with any HPV infection ($p < 0.05$). Five factors (age, gender, country of birth, whether or not smoke at the time of interview, and age of sex debut) were associated with HRHPV infection ($p < 0.05$). LRHPV had six independently associated factors (age, gender, race, education level, marital status, and age of sex debut) ($p < 0.05$). Three characteristics, age, gender, and age of sex debut, were significantly associated with HPV infection of all types, HRHPV and LRHPV. Two factors, country of birth, number of alcohols used per day seemed to have little independent risks. Other factors exhibited mixed results (Table 6). People who started sex before 18 years old were at least 3 times more likely to be HPV positive than those who started sex after 18 years old (adjusted OR, 3.95 for overall HPV, 3.65 for HRHPV, and 3.68 for LRHPV). Interestingly, use of tobacco had risks for overall and

HR HPV but not for LRHPV, while race, education and marital status were associated with LRHP infection only (Table 6).

Among the age groups, people aged 65-70 years had the highest risk of being infected by any HPV or HRHPV (adjusted OR, 3.01 for all types and 2.14 for HRHPV). People aged 30-34 years were less likely to be infected by HPV (adjusted OR, 0.69 for all types and 0.22 for HRHPV) (Table 6). Persons aged 25 and older were more likely to had LRHPV infection (adjusted OR range: 2.33-9.54) compared to those of 18-24 years of age. Notably, people aged 60-64 were 9.54 times were more likely to be infected by LRHPV than those in any other group (Table 6).

Table 6 Multivariable analysis of characteristics independently associated with HPV infection among people aged 18-69 years

Characteristic	Any HPV		HRHPV		LRHPV	
	Adj. OR (95%CI)	P value	Adj. OR (95%CI)	p value	Adj. OR (95%CI)	P value
Age						
18-24	1[Reference]	0.0004	1[Reference]	0.0007	1[Reference]	<.0001
25-29	1.38(0.42-4.55)		1.28(0.37-4.44)		2.33(0.32-16.61)	
30-34	0.69(0.18-2.59)		0.22(0.04-1.1)		2.82(0.44-17.89)	
35-39	1.7(0.57-5.01)		0.68(0.17-2.61)		4.99(0.71-34.76)	
40-44	2.23(0.64-7.81)		0.97(0.32-2.88)		6.32(0.52-76.07)	
45-49	2.34(0.8-6.78)		1.42(0.27-7.3)		5.53(1.29-23.7)	
50-54	2.53(1.09-5.89)		1.49(0.56-3.9)		6.58(1.36-31.82)	
55-59	2.62(0.59-11.63)		1.12(0.19-6.39)		7.77(1.15-52.51)	
60-64	2.65(0.55-12.64)		0.76(0.17-3.39)		9.54(1.2-75.34)	
65-70	3.01(0.93-9.76)		2.14(0.63-7.2)		5.94(0.86-40.91)	
Gender						
M	3.41(1.77-6.54)	0.0002	3.26(1.36-7.79)	0.0079	2.98(1.59-5.57)	0.0006
F	1[Reference]		1[Reference]		1[Reference]	
Country of birth						
US	1.06(0.3-3.64)	0.927	2.88(1.18-7.04)	0.0197	0.49(0.09-2.71)	0.4208
Others	1[Reference]		1[Reference]		1[Reference]	
Race						
Mexican American	0.87(0.26-2.9)	0.784	1.28(0.37-4.39)	0.5142	0.75(0.15-3.76)	0.0471
Other Hispanic	0.59(0.18-1.96)		1.52(0.6-3.82)		0.19(0.03-1.04)	
NHW	1[Reference]		1[Reference]		1[Reference]	
NHB	1.22(0.84-1.78)		0.9(0.45-1.78)		1.38(0.88-2.16)	
NH Asian	0.82(0.25-2.62)		2.75(0.58-12.92)		0.17(0.02-1.26)	
Other	0.56(0.13-2.37)		0.53(0.12-2.23)		0.92(0.15-5.36)	
Education						
Less than high school	1[Reference]	0.0892	1[Reference]	0.6569	1[Reference]	0.0071
High school or equivalent	3.49(0.98-12.44)		2.21(0.4-12.21)		6.65(1.47-30.01)	
College or above	2.71(0.73-10.01)		2.22(0.38-12.95)		3.56(0.75-16.76)	

Marital status						
Married/living with partners	0.87(0.52-1.45)	0.7908	1.95(0.66-5.75)	0.2714	0.42(0.21-0.82)	0.0416
Widow/divorced/separated	1.1(0.54-2.22)		2.62(0.8-8.56)		0.64(0.28-1.5)	
Never married	1[Reference]		1[Reference]		1[Reference]	
No. Alcohol/day						
No alcohol	1[Reference]	0.7144	1[Reference]	0.5197	1[Reference]	0.0727
≥2 to <4/day	0.97(0.53-1.79)		1.45(0.83-2.5)		0.63(0.33-1.22)	
≥4 to <8/day	1.05(0.5-2.19)		1.22(0.49-3.01)		1.27(0.5-3.2)	
>8/day	1.59(0.59-4.31)		1.52(0.45-5.09)		1.6(0.53-4.81)	
Currently smoking						
Yes	1.88(1.29-2.74)	0.0009	3.7(1.83-7.49)	0.0003	0.96(0.48-1.92)	0.9088
No	1[Reference]		1[Reference]		1[Reference]	
Age of sex debut						
<18	3.95(2.33-6.7)	<.0001	3.65(1.56-8.55)	0.0028	3.68(1.63-8.31)	0.0017
≥18	1[Reference]		1[Reference]		1[Reference]	

3.6 Factors Associated with Oral HPV Infection by Gender

Males were 3 times more likely to be infected by HPV than females (adjusted OR, 3.41 for any HPV infection, 3.26 for HRHPV, and 2.98 for LRHPV) (Table 6). Consistent with this, the weighted overall prevalence of any oral HPV was higher in males than that in females at all age groups. The prevalence of any oral HPV exhibited a trend to increase by age in men, but remained flat in women (Figure 5), 17.1% of men between 55-59 were HPV DNA positive, while the prevalence was only 1.1% in women at this age group (Figure 5). This data revealed a synergistic effect of gender and age on oral HPV prevalence. Consistent with this idea, gender and age showed interactions on the effect of any type HPV infection in a logistic model with these two variables ($p < 0.001$). Notably, the oncogenic HPV16 was more than 7 times higher in males (1.73%) compared to that in females (0.23%). Without considering other factors, men were 7.6 times more likely to be infected by HPV16 than women (unadjusted OR: 7.6) (Figure 6). To further understand the difference of risks between men and women, logistic models were constructed for men and women with several covariates including age, race, education, number

of alcohol consumed per day, current smoking, and age of sex debut. As shown in Table 7, Men and women exhibited profound differences in HPV infection risks across the characteristics. Gender showed interactions with age and race for any oral HPV infection of any type ($p < 0.05$). For HRHPV, in addition to interaction between gender and age, gender and number of alcohol used per day were identified to have significant interaction ($p < 0.001$), while gender and education showed marginal interaction ($p = 0.03$) (Table 7). The risks (measured as adjusted OR) for any HPV type infection by age showed distinct patterns between male and female. The trend of risk increased by age for men but was flat for women (Figure 7). Consistent with prevalence data in Figure 5, the highest risk for men was at age of 55-59 (adjusted OR: 7.13 for men and 0.22 for women). For HRHPV, the risk of infection in men were generally low during 18-44 years of age and increased at age of 45-59. The highest risk was observed at age of 65-70 (Figure 8). Notably, women were more likely to have HRHPV infection when they were at age of 40 to 44 (Adjusted OR=5.3) compared to other ages and men (Figure 8). Several characteristics, including education and use of cigarette, seemed to play a role in modifying the association between HPV infection and gender. Compared to men with education below high school, men with high school or college level education had higher risks for any HPV type infection (adjusted OR: 4.93 for high school and 4.13 for college and above; $p = 0.0396$), as well as for HRHPV (adjusted OR: 5.58 for high school and 7.21 for college and above; $p = 0.0517$) (Table 7 and Figure 9). By contrast, women's risk for HPV infection was not associated with education (p value, 0.44 for any HPV and 0.31 for HRHPV). Smoking of cigarette seemed to increase risks in men (OR=4.77; $p = 0.011$) significantly but moderately in women (OR=2.68; $p = 0.32$) (Figure 10).

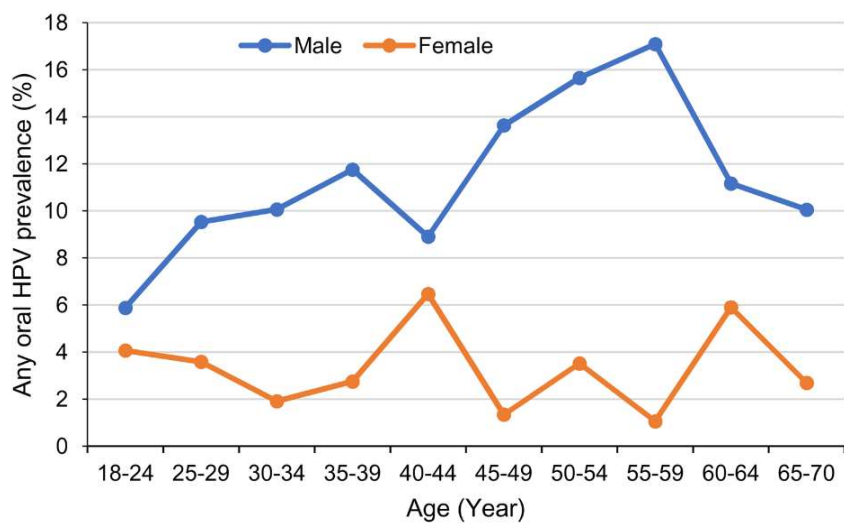


Figure 5 Weighted prevalence of oral HPV (any type) among U.S. populations by age and gender

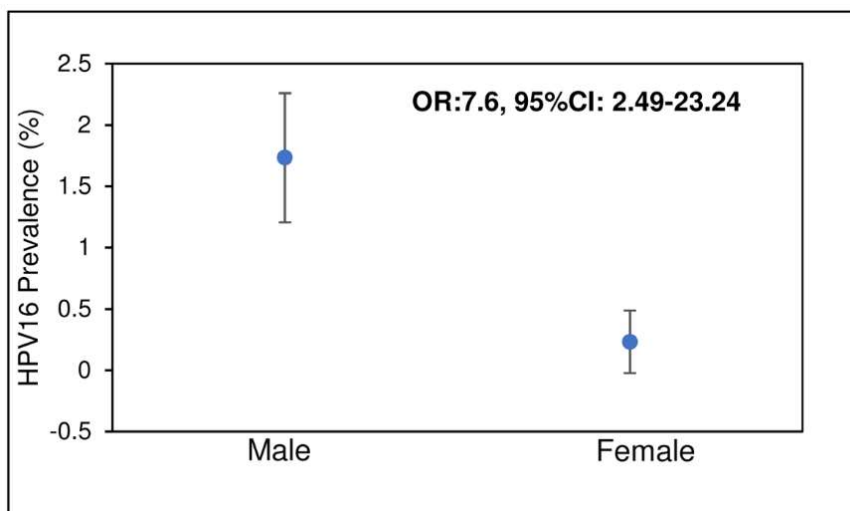


Figure 6 Weighted prevalence of oral HPV16 infection among U.S. adults by gender

Table 7 Adjusted Odds Ratio for HPV infection by gender in U.S. population aged 18-69 years

Characteristic	Overall HPV Adjusted OR (95%CI)		P value for interaction with gender	HRHPV Adjusted OR (95%CI)		P value for interaction with gender
	Men	Women		Men	Women	
Age						
18-24	1[Reference]	1[Reference]	<.0001	1[Reference]	1[Reference]	<.0001
25-29	2.86(0.6-13.51)	0.25(0.05-1.22)		2.37(0.5-11.1)	0.26(0.02-2.58)	
30-34	1.37(0.38-4.92)	0.13(0.01-1.26)		0.39(0.07-2.13)	0.42(0.02-8.54)	
35-39	2.84(0.85-9.49)	0.69(0.15-3.12)		1.1(0.32-3.77)	0.81(0.03-18.03)	
40-44	2.78(0.76-10.08)	1.73(0.66-4.51)		0.74(0.14-3.97)	5.3(1.28-21.87)	
45-49	6.11(1.51-24.63)	0.14(0.02-1.05)		3.66(0.84-15.98)	0.53(0.02-9.97)	
50-54	4.78(2.22-10.28)	0.83(0.18-3.8)		3.72(1.43-9.65)	1.1(0.05-22.73)	
55-59	7.13(1.68-30.15)	0.22(0.03-1.53)		3.23(0.73-14.25)	0.54(0.04-6.89)	
60-64	3.9(0.89-17.14)	1.71(0.27-10.79)		1.65(0.45-6.04)	1.67(0.31-9.06)	
65-70	6.82(1.81-25.61)	0.29(0.06-1.37)		6.12(1.8-20.8)	0.36(0.03-4.2)	
p value	<.0001	<.0001	<.0001	0.0007		
Race						
Mexican American	0.82(0.3-2.22)	1.18(0.14-9.69)	0.01	1.11(0.38-3.22)	1.39(0.13-14.5)	0.2676
Other Hispanic	0.54(0.28-1.03)	0.68(0.21-2.22)		0.99(0.52-1.9)	0.98(0.21-4.56)	
NHW	1[Reference]	1[Reference]		1[Reference]	1[Reference]	
NHB	1.16(0.72-1.87)	1.29(0.51-3.27)		0.89(0.41-1.92)	0.21(0.02-2.15)	
NH Asian	0.35(0.1-1.23)	2.53(0.43-14.78)		0.5(0.05-4.31)	6.49(0.72-58.17)	
Other	0.74(0.14-3.71)	0.27(0.02-2.51)		0.72(0.13-3.88)	0.54(0.04-5.99)	
p value	0.0372	0.3239		0.966	0.2953	
Education						
Less than high school	1[Reference]	1[Reference]	0.1353	1[Reference]	1[Reference]	0.0309
High school or equivalent	4.93(1.43-17.03)	1.18(0.23-6)		5.58(1.21-25.62)	0.82(0.13-4.99)	
College or above	4.13(1.18-14.41)	0.77(0.12-4.69)		7.21(1.46-35.5)	0.4(0.07-2.36)	
p value	0.0396	0.444		0.0517	0.3059	
No. Alcohol/day						
No alcohol	1[Reference]	1[Reference]	0.7238	1[Reference]	1[Reference]	<.0001
≥2 to <4/day	0.85(0.35-2.01)	1.26(0.57-2.76)		1.14(0.52-2.5)	5.98(1.46-24.33)	
≥4 to <8/day	0.83(0.32-2.17)	1.98(0.62-6.32)		0.61(0.21-1.75)	20.5(6.05-69.42)	
>8/day	1.42(0.36-5.53)	2.28(0.33-15.64)		1.24(0.34-4.46)	8.11(2.44-26.92)	
p value	0.762	0.7174		0.6573	<.0001	
Currently smoking						
Yes	2.09(1.24-3.53)	1.68(0.76-3.67)	0.5362	4.77(1.87-12.16)	2.68(0.38-18.61)	0.6522
No	1[Reference]	1[Reference]		1[Reference]	1[Reference]	
p value	0.0055	0.1943		0.0011	0.3181	
Age of sex debut						
<18	3.57(1.75-7.3)	7.02(1.86-26.46)	0.2296	3.5(1.23-9.97)	5.71(0.86-37.7)	0.1736
≥18	1[Reference]	1[Reference]		1[Reference]	1[Reference]	
p value	0.0005	0.004		0.0185	0.0701	

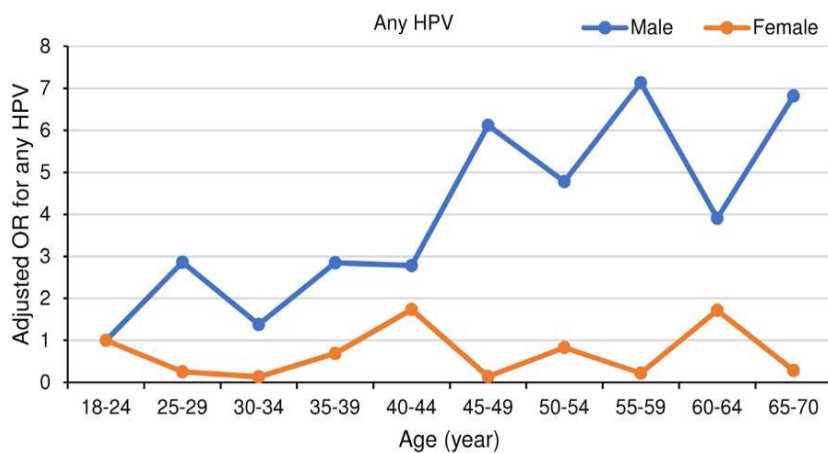


Figure 7 Adjusted Odds Ratio for oral HPV (any type) infection among U.S. adults by age and gender

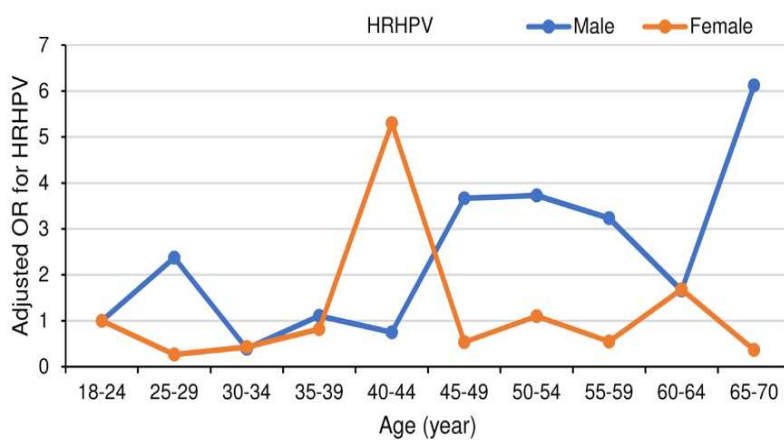


Figure 8 Adjusted Odds Ratio for oral HRHPV infection among U.S. adults by age and gender

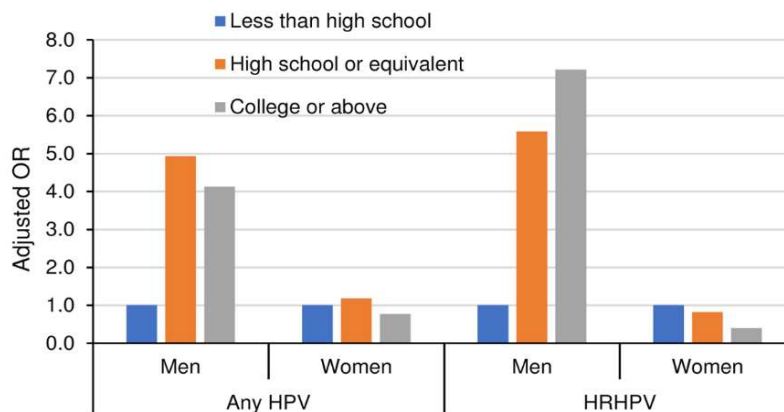


Figure 9 Adjusted Odds Ratio for oral HPV (any type) infection among U.S. population aged 18-69 years by education and gender

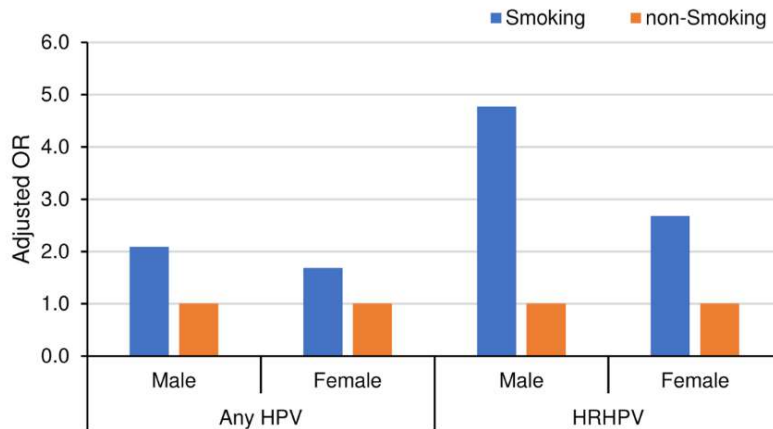


Figure 10 Adjusted Odds Ratio for oral HPV infection among U.S. population aged 18-69 years by tobacco use and gender

Despite the lower oral HPV prevalence in women compared in men, some group of women appeared to have be at high risk of being infected, particularly for HRHPV. As shown in Table 7 and Figure 11, women who drank had much higher risks for HPV infection than men. Strikingly, women who had more than 2 alcohols per day were much more likely to be infected by HRHPV than those who did not drink (Adjusted OR range: 5.98-20.5). By contrast, drinking alcohols appeared to have little effects in men (Figure 11). Interestingly, among race groups, Asian males had the lowest risk for HPV infection (adjusted OR, 0.35 for any HPV and 0.5 for HRHPV), but Asian female were highly vulnerable to HPV infection, particularly to HRHPV (adjusted OR, 6.49) (Table 7 and Figure 12). Black women had the lowest risk for HRHPV (adjusted OR, 0.21). Age of sex debut was another risk factors for HPV infection. Both men and women who started sex younger than 18 years old were more likely to have HPV infection than those who started sex after 18 years old (Table 7). Women seemed to be more vulnerable to HPV infection when they started sex before 18 years of age (adjusted OR: 7.02 in women vs 3.57 in men for any HPV, 5.71 in women vs 3.51 in men for HRHPV infection) (Figure 13).

Race and number of life time opposite sex partners showed significant interactions in the pair wise test. The effect of the number of life time opposite sex partners on oral HPV prevalence differed by race. Figure 14 showed the weighted prevalence of any HPV types over life time sex partners by race. The HPV prevalence increased in all race groups as the number of sex partners increased from 2-5 to more than 10. It increased more sharply in Non-Hispanic White and Black than in Mexican Americans and Asians. For people with less than 2 partners, Mexican and Black had relative high prevalence (Figure 14).

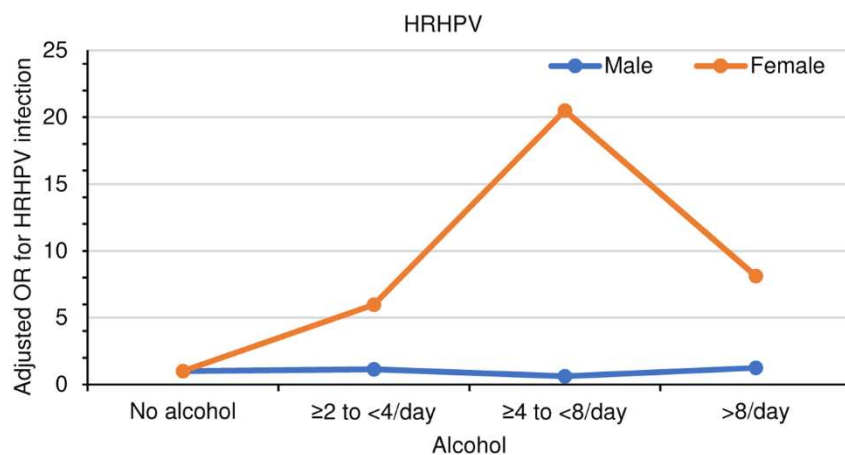


Figure 11 Adjusted Odds Ratio for HRHPV infection among U.S. population aged 18-69 years by gender and alcohol use

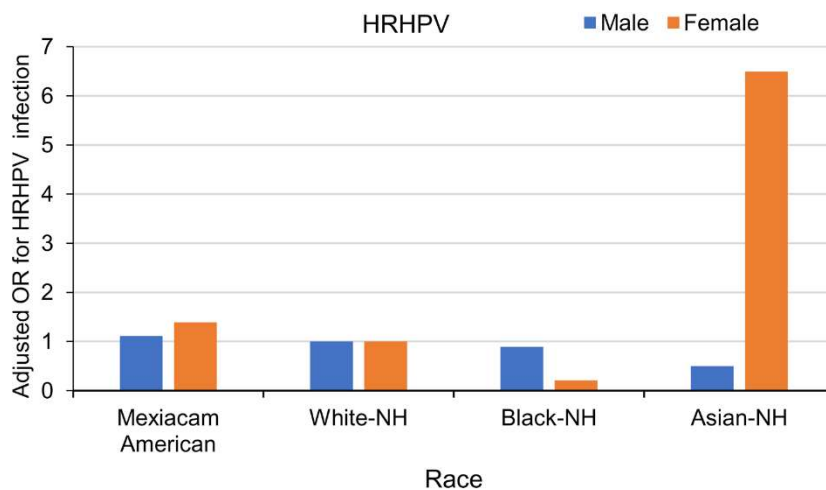


Figure 12 Adjusted Odds Ratio for HRHPV infection among U.S. population aged 18-69 years by gender and race

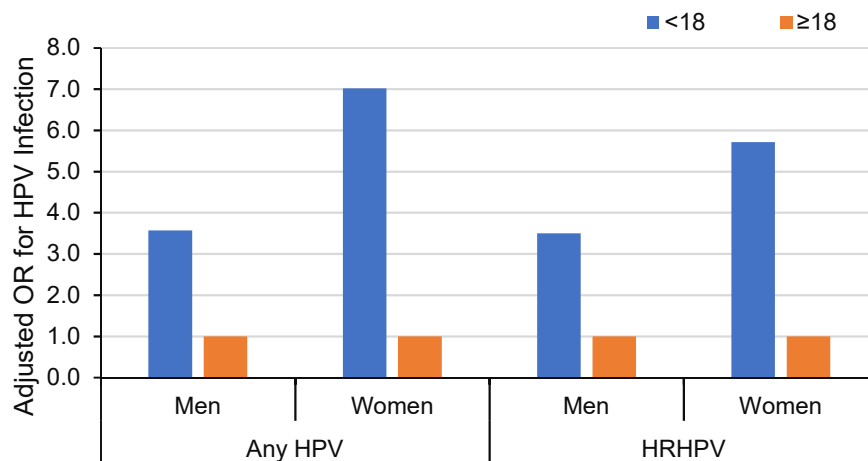


Figure 13 Association of gender and age of sex debut with oral HPV infection among U.S. population aged 18-69 years

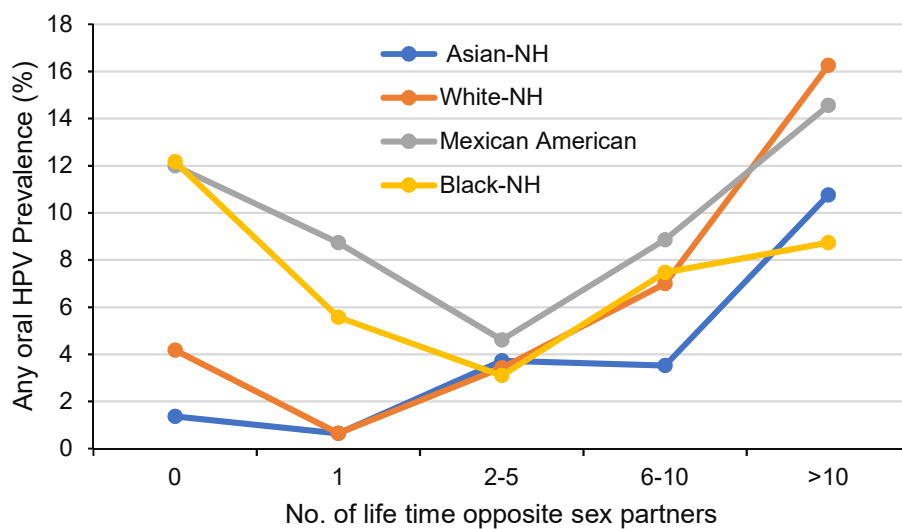


Figure 14 Weighted prevalence of oral HPV (any type) infection among U.S. adults by race and number of lifetime opposite sex partners

4 DISCUSSION

In this study, we used nationally representative sample of U.S. population collected through NHANES to study oral HPV infection. The prevalence oral HPV among U.S. population aged 18-69 years was estimated as 7.18%, 4.5% and 3.5% for any HPV types, HRHPV, and LRHPV, respectively. These results were comparable with previously reported oral HPV prevalence of 6.9%, 3.7% and 3.1% for any types, HRHPV and LRHPV, respectively [13]. The oral HPV prevalence was substantially lower than that of genital HPV infection in men or women [20, 21, 50]. Thirty-six HPV types were detected in the samples. The most prevalent oral HPV type was HPV16 (0.97%). HPV infection was associated with sexual history and several demographic characteristics, including age, gender, and race. Alcohol and tobacco use were identified as potential risk modifiers for HPV infection in men and women.

4.1 Overall Prevalence of Oral HPV Infection

Prevalence of any HPV types was the lowest among 18-24 years old people (5.0%), increased to a peak (9.4%) among people at age of 50-54, and decreased to 6.3% in 65-70 years of age. The prevalence of HRHPV and LRHPV over age each followed a bimodal pattern with peaks at 25-29 years and 50-54 years for HRHPV, and with peaks at 40-44 and 60-64 for LRHPV. This pattern for HRHPV infection was similar with that in a previous report [13]. However, the age when prevalence peaks occurred for HRHPV and LRHPV were different from that of previous report in which the bimodal pattern was observed with peaks at 30-34 and 60-64 years of age for both HRHPV and LRHPV [13]. Similarly, prevalence of genital HPV in men follow a bimodal pattern with peaks at 28-32 years and 58-59 years [21]. The reason for these age-related changes in prevalence are not clear, but are likely resulting from the difference of samples sets, or related to factors including clearance of virus over time, alternative transmission

route for oral HPV, changes of incidence caused by sexual behavior changes, or changes in immune systems [20].

4.2 Prevalence of Oral HPV Infection Varies by Characteristics

Oral HPV prevalence differed across categories of demographic and behavioral characteristics including gender, country of birth, race, education, marital status, as well as alcohol and tobacco use. This is similar with previous observation for both oral and genital HPV infection [13, 20, 21]. An analysis of the effects of oral health on HPV infection, including decayed teeth, gum disease and oral hygiene, was not possible because of low number of participants who answered “No” to these questions. Prevalence of oral HPV infection (any HPV type, HRHPV, and LRHPV) was higher in individuals with sexual experience and increased with number of life time sex partners. Prevalence of oral HPV among people who started sex before 18 years of age was 2-3 time higher than those who started sex after 18 years old. People who had multiple same sex partners had the highest prevalence (32.7%). Oral HPV prevalence were higher among people who had new partners in past 12 months. These data suggest that HPV infection is strongly associated with sexual behavior, supporting the opinion that sexual contact was the major route for oral HPV transmission [13]. It should be pointed out that assessment of associations between any single sexual behavior and HPV infection is difficult due to the high collinearity among sexual behaviors [13]. Of note, sexual history and sexually transmitted disease status were self-reported and there were considerable number of nonresponses, which may not be accurate due to misclassification and the sensitivity of the topics [20, 21].

4.3 Factors Associated with Oral HPV Infection

Univariate analysis identified 16 characteristics that potentially contribute to oral HPV infection ($p < 0.05$). 9 significant characteristics were selected for multivariable regression

analysis. Some of the behavioral characteristics, including sexual behavior as well as alcohol and tobacco use, had collinearity and were analyzed separately. This analysis revealed four risk factors that were independently associated with oral HPV infection of any type including age, gender whether or not smoke at time of interview, and age of sex debut. These results were consistent with previous reports [13]. Furthermore, we identified country of birth, in addition to the four factors for any types, to be associated with HRHPV. We found that LRHPV infection was associated with six factors including age, gender, race, education, marital status, and age of sex debut. Surprisingly, whether or not smoke at time of interview was not associated with LRHPV infection ($p=0.9$). All the 3 HPV categories (any type, HRHPV and LRHPV) included 3 independently associated factors, age, gender, and age of sex debut, which were reported as risk factors for genital HPV infection as well [20, 21]. The association of sexual behavior with HPV infection was consistent with previous report that oral HPV was predominately sexually transmitted [13, 38]. In the dataset used in this study, people age 65-70 years had the highest risk for any types and HRHPV, possibly due to changes in hosts' immune system as they aged, though the reason was unknown. People aged 30-34 years had lowest risk for HRHPV, which followed the first HRHPV prevalence peak at 25-29. This may reflect the active immune responses for virus clearance after a wave of HPV infection.

4.4 Differences of Oral HPV infection in Males and Females

It has been reported that HPV prevalence is different between male and female [13, 38]. Our data provided additional evidence for this point. We found that male and female differed profoundly regarding to the sensitivity of oral HPV infection in response to risk factors. Oral HPV prevalence was higher in men than that in women at all age stages. The prevalence of oral HPV tended to increase by age in men. The highest prevalence (17.1%) of oral HPV was found

in men aged 55-59 years. One explanation for this observation was decreased immune response to natural infection in men as they aged [51, 52]. We identified an interaction between gender and age on infection of any HPV type. HPV prevalence (any type) in men at age of 55-59 years were 17 times higher than women at the same age (adjusted OR: 7 times higher). To further assess the clinical consequences of HPV infection, we evaluated the risk factors for HRHPV. In addition to interactions between age and gender, number of alcohol consumed per day interacted to gender to exert effect on HRHPV infection. Men had nearly 4 times higher risk for HRHPV infection at their age of 45-59, compared to women in the same age. In addition, the prevalence of the oncogenic HPV16 was higher in males than females. These data may provide some explanations for higher prevalence of OPCs observed in men [34, 53]. Age of sex debut of younger than 18 can increase risk for oral HPV infection more than 3 times. Other factors including education and whether or not smoke seemed to be risk modifiers for HPV infection in men.

We found that certain group of women were at particularly high risk for HPV infection. First, women at age of 40-44 were 5 times more likely to have HRHPV infection than those of other ages. Second, women who drank more than 2 alcohols per day were at much higher risk (5-20 times) than those who did not drank and men. Third, Asian women seemed to be more vulnerable to HRHPV infection (adjusted OR, 6.49) than men and women of any other ethnic groups. Finally, women were twice more likely to be infected by HPV than men if their age of sex debut were younger than 18 and had 5 to 7 times more risk compared to women who started sex after 18 years old. These data identified subgroups of women who were at higher risk for HPV infections at certain conditions and would be helpful for development of tailored

prevention programs. An evaluation of HPV16 by subgroups was restricted by the low number of HPV16 positive subjects in the dataset.

4.5 Limitations

This study has several limitations. First, the prevalence of HPV infection could be underestimated. This study relied on HPV DNA positivity detected by Roche LA, which only covered 37 types out of more than 200 known HPV types. Those types that were not on LA test were not reported even if they were present in the samples. New technologies such as Next Generation Sequencing (NGS) allow for simultaneously testing for all known HPV types [54]. The implementation of NGS-based method into survey data would provide more accurate estimation for HPV prevalence. Second, the data are cross-sectional and only provide a snapshot of HPV infection. Therefore, the study cannot distinguish between incidences from prevalence. We cannot calculate accumulative infections. This could result in incomplete pictures for HPV infection because HPV infection is dynamic and could be transient as a result of immune system clearance. Therefore, the prevalence estimate does not reflect the lifetime infection of HPV. Serologic testing that detects antibodies against HPV specific proteins provides better estimate for cumulative infection [55]. Third, self-reported information on characteristics are incomplete and might be inaccurate due to misclassifications or involving sensitive topics that participants are not willing to answer. Finally, estimates for prevalence of HRHPV and LRHPV are subjected to changes according to HPV classification. HPV risk classification is evolving as more data available on the oncogenicity of a HPV type and natural history of infection [20, 56]. It should be pointed out that some of the types classified as HRHPV in this study only had limited evidence to cause cancer [56]. The classification used in this study was selected to enable comparison with previous studies.

5 CONCLUSIONS

This study provided a national representative estimate for prevalence of oral HPV infection among population aged 18-69 years in the United States using the most update data in NHANES. The overall prevalence of any oral HPV types infection was estimated as 7.18% and the oncogenic high-risk HPV prevalence was 4.5%. The distribution of oral HPV infection varied by demographic and behavioral characteristics. Oral HPV infection was highly associated with sexual history. Age, gender, and sexual behavior were common risk factors for HPV infection. Prevalence of oral HPV infection were significantly higher in males than that in females. Our data suggested that factors including use of alcohol and tobacco were risk modifiers for HPV infection in males. We identified several subgroups of women who were at high risk for oral HPV infection including women aged 40-44 years, women who drank more than 2 alcohols daily, women who started sex before 18 years old, and women of Asian group. In conclusion, our study provided the most update estimate for oral HPV infection at population level and identified factors conferring significant risks for oral HPV infection. Our data provided additional evidence that sexual contact was a major route for oral HPV transmission. The finding in this study has important implications on oral HPV surveillance and epidemiological study.

REFERENCES

1. Hagensee ME, Yaegashi N, Galloway DA: **Self-assembly of human papillomavirus type 1 capsids by expression of the L1 protein alone or by coexpression of the L1 and L2 capsid proteins.** *J Virol* 1993, **67**:315-322.
2. Kirnbauer R, Booy F, Cheng N, Lowy DR, Schiller JT: **Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic.** *Proc Natl Acad Sci U S A* 1992, **89**:12180-12184.
3. Bzhalava D, Guan P, Franceschi S, Dillner J, Clifford G: **A systematic review of the prevalence of mucosal and cutaneous human papillomavirus types.** *Virology* 2013, **445**:224-231.
4. de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H: **Classification of papillomaviruses.** *Virology* 2004, **324**:17-27.
5. Weinstock H, Berman S, Cates W, Jr.: **Sexually transmitted diseases among American youth: incidence and prevalence estimates, 2000.** *Perspect Sex Reprod Health* 2004, **36**:6-10.
6. Bernard HU: **The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses.** *J Clin Virol* 2005, **32 Suppl 1**:S1-6.
7. Satterwhite CL, Torrone E, Meites E, Dunne EF, Mahajan R, Ocfemia MC, Su J, Xu F, Weinstock H: **Sexually transmitted infections among US women and men: prevalence and incidence estimates, 2008.** *Sex Transm Dis* 2013, **40**:187-193.
8. zur Hausen H: **Papillomaviruses and cancer: from basic studies to clinical application.** *Nat Rev Cancer* 2002, **2**:342-350.

9. Forman D, de Martel C, Lacey CJ, Soerjomataram I, Lortet-Tieulent J, Bruni L, Vignat J, Ferlay J, Bray F, Plummer M, Franceschi S: **Global burden of human papillomavirus and related diseases.** *Vaccine* 2012, **30 Suppl 5**:F12-23.
10. Saraiya M, Unger ER, Thompson TD, Lynch CF, Hernandez BY, Lyu CW, Steinau M, Watson M, Wilkinson EJ, Hopenhayn C, et al: **US assessment of HPV types in cancers: implications for current and 9-valent HPV vaccines.** *J Natl Cancer Inst* 2015, **107**:dju086.
11. Arbyn M, Castellsague X, de Sanjose S, Bruni L, Saraiya M, Bray F, Ferlay J: **Worldwide burden of cervical cancer in 2008.** *Ann Oncol* 2011, **22**:2675-2686.
12. Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah KV, Snijders PJ, Meijer CJ, International Agency for Research on Cancer Multicenter Cervical Cancer Study G: **Epidemiologic classification of human papillomavirus types associated with cervical cancer.** *N Engl J Med* 2003, **348**:518-527.
13. Gillison ML, Broutian T, Pickard RK, Tong ZY, Xiao W, Kahle L, Graubard BI, Chaturvedi AK: **Prevalence of oral HPV infection in the United States, 2009-2010.** *JAMA* 2012, **307**:693-703.
14. Jacobs MV, de Roda Husman AM, van den Brule AJ, Snijders PJ, Meijer CJ, Walboomers JM: **Group-specific differentiation between high- and low-risk human papillomavirus genotypes by general primer-mediated PCR and two cocktails of oligonucleotide probes.** *J Clin Microbiol* 1995, **33**:901-905.
15. Parkin DM, Bray F: **Chapter 2: The burden of HPV-related cancers.** *Vaccine* 2006, **24 Suppl 3**:S3/11-25.

16. Smith EM, Ritchie JM, Summersgill KF, Klussmann JP, Lee JH, Wang D, Haugen TH, Turek LP: **Age, sexual behavior and human papillomavirus infection in oral cavity and oropharyngeal cancers.** *Int J Cancer* 2004, **108**:766-772.
17. Koutsky L: **Epidemiology of genital human papillomavirus infection.** *Am J Med* 1997, **102**:3-8.
18. Burd EM: **Human papillomavirus and cervical cancer.** *Clin Microbiol Rev* 2003, **16**:1-17.
19. Bosch FX, de Sanjose S: **Chapter 1: Human papillomavirus and cervical cancer--burden and assessment of causality.** *J Natl Cancer Inst Monogr* 2003:3-13.
20. Hariri S, Unger ER, Sternberg M, Dunne EF, Swan D, Patel S, Markowitz LE: **Prevalence of genital human papillomavirus among females in the United States, the National Health And Nutrition Examination Survey, 2003-2006.** *J Infect Dis* 2011, **204**:566-573.
21. Han JJ, Beltran TH, Song JW, Klaric J, Choi YS: **Prevalence of Genital Human Papillomavirus Infection and Human Papillomavirus Vaccination Rates Among US Adult Men: National Health and Nutrition Examination Survey (NHANES) 2013-2014.** *JAMA Oncol* 2017, **3**:810-816.
22. Castellsague X, Ghaffari A, Daniel RW, Bosch FX, Munoz N, Shah KV: **Prevalence of penile human papillomavirus DNA in husbands of women with and without cervical neoplasia: a study in Spain and Colombia.** *J Infect Dis* 1997, **176**:353-361.
23. Kjaer SK, Chackerian B, van den Brule AJ, Svare EI, Paull G, Walbomers JM, Schiller JT, Bock JE, Sherman ME, Lowy DR, Meijer CL: **High-risk human papillomavirus is sexually transmitted: evidence from a follow-up study of virgins starting sexual activity (intercourse).** *Cancer Epidemiol Biomarkers Prev* 2001, **10**:101-106.

24. Moscicki AB, Hills N, Shiboski S, Powell K, Jay N, Hanson E, Miller S, Clayton L, Farhat S, Broering J, et al: **Risks for incident human papillomavirus infection and low-grade squamous intraepithelial lesion development in young females.** *JAMA* 2001, **285**:2995-3002.
25. Munoz N, Castellsague X, Bosch FX, Tafur L, de Sanjose S, Aristizabal N, Ghaffari AM, Shah KV: **Difficulty in elucidating the male role in cervical cancer in Colombia, a high-risk area for the disease.** *J Natl Cancer Inst* 1996, **88**:1068-1075.
26. De Vincenzo R, Ricci C, Conte C, Scambia G: **HPV vaccine cross-protection: Highlights on additional clinical benefit.** *Gynecol Oncol* 2013, **130**:642-651.
27. Centers for Disease Control and Prevention: **FDA licensure of quadrivalent human papillomavirus vaccine (HPV4, Gardasil) for use in males and guidance from the Advisory Committee on Immunization Practices (ACIP).** *MMWR Morb Mortal Wkly Rep* 2010, **59**:630-632.
28. Centers for Disease Control and Prevention: **Recommendations on the use of quadrivalent human papillomavirus vaccine in males--Advisory Committee on Immunization Practices (ACIP), 2011.** *MMWR Morb Mortal Wkly Rep* 2011, **60**:1705-1708.
29. Centers for Disease Control and Prevention: **CDC recommends only two HPV shots for younger adolescents.** <https://www.cdc.gov/media/releases/2016/p1020-hpv-shots.html>. Updated October 20, 2016 2016.
30. Steinau M, Hariri S, Gillison ML, Broutian TR, Dunne EF, Tong ZY, Markowitz LE, Unger ER: **Prevalence of cervical and oral human papillomavirus infections among US women.** *J Infect Dis* 2014, **209**:1739-1743.

31. Steinau M, Saraiya M, Goodman MT, Peters ES, Watson M, Cleveland JL, Lynch CF, Wilkinson EJ, Hernandez BY, Copeland G, et al: **Human papillomavirus prevalence in oropharyngeal cancer before vaccine introduction, United States.** *Emerg Infect Dis* 2014, **20**:822-828.
32. Gillison ML: **Human papillomavirus-associated head and neck cancer is a distinct epidemiologic, clinical, and molecular entity.** *Semin Oncol* 2004, **31**:744-754.
33. Viens LJ, Henley SJ, Watson M, Markowitz LE, Thomas CC, Thompson TD, Razzaghi H, Saraiya M: **Human Papillomavirus-Associated Cancers - United States, 2008-2012.** *MMWR Morb Mortal Wkly Rep* 2016, **65**:661-666.
34. Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, Jiang B, Goodman MT, Sibug-Saber M, Cozen W, et al: **Human papillomavirus and rising oropharyngeal cancer incidence in the United States.** *J Clin Oncol* 2011, **29**:4294-4301.
35. Hong AM, Grulich AE, Jones D, Lee CS, Garland SM, Dobbins TA, Clark JR, Harnett GB, Milross CG, O'Brien CJ, Rose BR: **Squamous cell carcinoma of the oropharynx in Australian males induced by human papillomavirus vaccine targets.** *Vaccine* 2010, **28**:3269-3272.
36. Nasman A, Attner P, Hammarstedt L, Du J, Eriksson M, Giraud G, Ahrlund-Richter S, Marklund L, Romanitan M, Lindquist D, et al: **Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: an epidemic of viral-induced carcinoma?** *Int J Cancer* 2009, **125**:362-366.
37. Jemal A, Simard EP, Dorell C, Noone AM, Markowitz LE, Kohler B, Ehemann C, Saraiya M, Bandi P, Saslow D, et al: **Annual Report to the Nation on the Status of Cancer, 1975-**

- 2009, featuring the burden and trends in human papillomavirus(HPV)-associated cancers and HPV vaccination coverage levels. *J Natl Cancer Inst* 2013, **105**:175-201.**
38. Sonawane K, Suk R, Chiao EY, Chhatwal J, Qiu P, Wilkin T, Nyitray AG, Sikora AG, Deshmukh AA: **Oral Human Papillomavirus Infection: Differences in Prevalence Between Sexes and Concordance With Genital Human Papillomavirus Infection, NHANES 2011 to 2014. *Ann Intern Med* 2017, **167**:714-724.**
39. Gillison ML, D'Souza G, Westra W, Sugar E, Xiao W, Begum S, Viscidi R: **Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. *J Natl Cancer Inst* 2008, **100**:407-420.**
40. 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**:3065-3069.**
41. Centers for Disease Control and Prevention: **National Health and Nutrition Examination Survey. <http://www.cdc.gov/nchs/nhanes.htm>, Accessed February 20, 2018.**
42. Mirel LB, Mohadjer LK, Dohrmann SM, Clark J, Burt VL, Johnson CL, Curtin LR: **National Health and Nutrition Examination Survey: estimation procedures, 2007-2010. *Vital Health Stat 2* 2013:1-17.**
43. NHANES 2013-2104 Overview: **Centers for Disease Control and Prevention. <https://www.cdc.gov/nchs/nhanes/continuousnhanes/Overview.aspx?BeginYear=2013>, Accessed February 22, 2018.**
44. Centers for Disease Control and Prevention: **MEC Laboratory Procedures Manual. http://www.cdc.gov/nchs/data/nhanes/nhanes_13_14/2013_MEC_Laboratory_Procedures_Manual.pdf 2013.**

45. Broutian TR, He X, Gillison ML: **Automated high throughput DNA isolation for detection of human papillomavirus in oral rinse samples.** *J Clin Virol* 2011, **50**:270-275.
46. Gravitt PE, Peyton CL, Alessi TQ, Wheeler CM, Coutlee F, Hildesheim A, Schiffman MH, Scott DR, Apple RJ: **Improved amplification of genital human papillomaviruses.** *J Clin Microbiol* 2000, **38**:357-361.
47. Marks M, Gupta SB, Liaw KL, Kim E, Tadesse A, Coutlee F, Sriplienchan S, Celentano DD, Gravitt PE: **Confirmation and quantitation of human papillomavirus type 52 by Roche Linear Array using HPV52-specific TaqMan E6/E7 quantitative real-time PCR.** *J Virol Methods* 2009, **156**:152-156.
48. Rao JNKaS, A. J. : **The Analysis of Categorical Data from Complex Surveys: Chi-Squared Tests for Goodness of Fit and Independence in Two-Way Tables.** *Journal of the American Statistical Association* 1981, **76**:221-230.
49. Rao JNKaS, A. J.: **On Chi-Squared Tests for Multiway Contingency Tables with Cell Properties Estimated from Survey Data.** *Annals of Statistics* 1984, **12**:46-60.
50. Dunne EF, Unger ER, Sternberg M, McQuillan G, Swan DC, Patel SS, Markowitz LE: **Prevalence of HPV infection among females in the United States.** *JAMA* 2007, **297**:813-819.
51. Giuliano AR, Lu B, Nielson CM, Flores R, Papenfuss MR, Lee JH, Abrahamsen M, Harris RB: **Age-specific prevalence, incidence, and duration of human papillomavirus infections in a cohort of 290 US men.** *J Infect Dis* 2008, **198**:827-835.
52. Giuliano AR, Palefsky JM, Goldstone S, Moreira ED, Jr., Penny ME, Aranda C, Vardas E, Moi H, Jessen H, Hillman R, et al: **Efficacy of quadrivalent HPV vaccine against HPV Infection and disease in males.** *N Engl J Med* 2011, **364**:401-411.

53. Chaturvedi AK, Engels EA, Anderson WF, Gillison ML: **Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States.** *J Clin Oncol* 2008, **26**:612-619.
54. Li T, Unger ER, Batra D, Sheth M, Steinau M, Jasinski J, Jones J, Rajeevan MS: **Universal Human Papillomavirus Typing Assay: Whole-Genome Sequencing following Target Enrichment.** *J Clin Microbiol* 2017, **55**:811-823.
55. Markowitz LE, Sternberg M, Dunne EF, McQuillan G, Unger ER: **Seroprevalence of human papillomavirus types 6, 11, 16, and 18 in the United States: National Health and Nutrition Examination Survey 2003-2004.** *J Infect Dis* 2009, **200**:1059-1067.
56. Paavonen J, Naud P, Salmeron J, Wheeler CM, Chow SN, Apter D, Kitchener H, Castellsague X, Teixeira JC, Skinner SR, et al: **Efficacy of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types (PATRICIA): final analysis of a double-blind, randomised study in young women.** *Lancet* 2009, **374**:301-314.