

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

Prevalence and Sociodemographic Covariates of Human Papillomavirus in the United States:
Findings from the 2008 to 2020 National Inpatient Sample Database

Emily Dantes¹, Mohammad Alfrad Nobel Bhuiyan¹, Margaret Bourg¹, Corey D. Smith², Md.
Shenuarin Bhuiyan², Deborah G Smith³

¹Department of Medicine, Louisiana State University Health Sciences Shreveport, Shreveport,
LA, USA

²Department of Pathology and Translational Pathobiology, Louisiana State University Health
Sciences Shreveport, Shreveport, LA, USA

³School of Allied Health Professions, Louisiana State University Health Sciences Shreveport,
Shreveport, LA, USA

Corresponding Author:
eam001@lsuhs.edu (ED)

24 **Abstract**

25 Human papillomavirus (HPV) infection is widespread worldwide, leading to a significant burden
26 of HPV-related cancers. This study aimed to identify the prevalence of HPV infection in the United
27 States based on a large nationwide inpatient sample database. We used the Nationwide Inpatient
28 Sample (NIS) database from 2008 to 2020 and identified HPV cases using the International
29 Classification of Diseases, Ninth (ICD-9) and Tenth Revision (ICD-10) codes. Pearson's chi-
30 square and independent-sample t-test were used for the analysis. The prevalence ratio was
31 calculated using logistic regression models and the Cochran Armitage trend test to examine trends.
32 A total of 47,506 HPV infections were identified from 2008 to 2020 among adults aged >18 years.
33 Most patients were female (99.72%), with a higher prevalence among the 26-40 age group and
34 white individuals. HPV infections were more prominent among low-income individuals and in the
35 southern region of the United States. From 2008 to 2020, HPV increased from 6.76% to 10.91%.
36 This study provides insights into the epidemiology of HPV infection in the United States.
37 Understanding the trends and characteristics of HPV infection can contribute to public health
38 interventions aimed at prevention and early detection.

39

40 **Introduction**

41 Human papillomavirus (HPV) infection is widespread worldwide, leading to a significant burden
42 of HPV-related cancers (1, 2). In the United States, HPV is the most common sexually
43 transmitted infection, and it is estimated that 14 million Americans were newly infected with
44 HPV in 2018 (3, 4, 5). Nearly all sexually active adults are infected at some point during their
45 lives, although they may be unaware of their infection (5). While most HPV infections are
46 asymptomatic and resolve spontaneously, long-lasting infections can progress into genital warts,

47 as well as oropharyngeal and anogenital cancers (including anal, penile, vaginal, vulvar, and
48 cervical) (2, 5).

49 More than 200 distinct types of HPV have been identified, with types falling into a low-risk or
50 high-risk group (4). Low-risk types rarely lead to symptoms or clinical disease. High-risk or
51 oncogenic HPV types, similarly asymptomatic at the onset of infection, may develop into cancer
52 years later if undetected and untreated. According to the National Cancer Institute, high-risk
53 HPV types cause 3% of all cancers in women and 2% in men in the United States (6). From 2015
54 to 2019, about 47,199 new HPV-associated cancers occurred annually in the United States (5, 7).

55 Among women, cervical cancer is the most common HPV-associated cancer (5). It is currently
56 the fourth most common cancer in women worldwide (2). Among men, oropharyngeal cancers
57 (cancers of the back of the throat, including the base of the tongue and tonsils) are the most
58 common (5, 8, 9). The incidence of oropharyngeal cancer in the United States increased by
59 approximately 1.94% per year from 2000-2015 and has been projected to increase continuously
60 over the next thirty years (10).

61 Many risk factors for HPV infection have been identified, including having multiple sexual
62 partners, sexual intercourse at a younger age (<15), long-term use of oral contraceptives,
63 smoking, low socioeconomic status, age, and genetic factors (5). Risk factors related to lifestyle
64 behaviors can be addressed with public health intervention, which is necessary to prevent new
65 infections and cancer progression (11). One important intervention is HPV vaccination, which
66 has been demonstrated to prevent up to 90% of HPV-related cancers (6). The American Cancer
67 Society recommends routine HPV vaccination for all persons aged 9 to 12 years old, with catch-
68 up vaccination for those under 26 who are not adequately vaccinated (12). However,
69 immunization rates in the US are much lower than for other vaccines, with only 59% of

70 adolescents aged 13 to 17 years old being up to date with HPV vaccination as of 2020 (12). In
71 addition, vaccination does not protect against all types of HPV, nor does it protect against
72 already-established HPV infection (12). Another important intervention is early detection
73 through screening; however, not all types of HPV-associated diseases and cancer have screening
74 recommendations. For instance, while pap tests and HPV co-tests are available to screen for
75 HPV-related cervical cancer, there is no method recommended by the American Cancer Society
76 or the U.S. Preventative Services Task Force to screen for other HPV-related cancers like
77 oropharyngeal, anal, penile, vulvar, or vaginal cancer (13). Due to these limitations, it is
78 important to understand the current epidemiology of HPV infection in the United States. This
79 study aims to understand the trends and identify the prevalence of HPV infection in the United
80 States from 2008 to 2020 based on NIS data.

81

82 **Material & Methods**

83 We used the Healthcare Cost and Utilization Project (HCUP)'s Nationwide Inpatient Sample
84 (NIS) database from 2008 to 2020 (accessed December 19, 2022) and identified HPV -related
85 hospital discharge cases using the International Classification of Diseases, ninth (ICD-9) and
86 tenth Revision (ICD-10) codes to conduct a retrospective study. NIS's database is the largest
87 inpatient database, which covers information from seven million hospital stays yearly from 45
88 states in the United States. This database contains data from 20% of hospital discharges from all
89 hospitals in HCUP (14).

90 To improve the sample's representativeness, accuracy, and stability of weighted national
91 estimates, the NIS was changed in 2012 to include 20% of all hospital discharges in HCUP.

92 Before 2012, the NIS included all discharges from 20% of hospitals in the HCUP (15). The NIS

93 data are organized so that each observation in the sample relates to a unique hospitalization. The
94 NIS database provides information on more than 100 clinical characteristics, including patient
95 demographics (such as age, sex, race, and the median income for ZIP code), hospital
96 characteristics (such as ownership, size, teaching status, census region, and division), primary
97 and up to 24 secondary diagnoses, 15 procedures as administrative codes, and codes for disease
98 severity. Additionally, because it contains data on several million hospitalizations each year, it
99 enables the analysis of healthcare utilization for issues related to rare diseases treated primarily
100 in an inpatient setting.

101 Because the NIS database is deidentified, the Institutional Review Board (IRB) waived the IRB
102 application. Authors did not have access to information that could identify individual participants
103 during or after data collection. We included adult patients (age: >18 years and older) of male and
104 female sex with a primary or secondary diagnosis of HPV based on the International
105 Classification of Diseases, Ninth (ICD-9) and Tenth revision (ICD-10) diagnosis code (ICD-9 =
106 795.05, 795.09, 796.75, 796.79, ICD-10 = R87.810, R87.820, R85.81, R85.82) (Appendix A).

107 When analyzing NIS hospital discharge data, we adhered to AHRQ's suggestions for producing
108 national estimates from NIS data. Data was weighted with care to account for the change in
109 database structure that occurred in 2012. The hospital discharge trend weight (TRENDWT) was
110 used to weight data for 2012 and earlier, while regular discharge weight (DISCWT) was used to
111 weight data for years after 2012. Descriptive statistics were used to characterize participant
112 details. Continuous variables were summed for weighted and unweighted data, while mean and
113 categorical variables were presented as frequencies, percentages, and 95% Confidence Intervals
114 (CI). Design-based t-tests were used for continuous values, and design-based Chi-squared tests
115 were applied for categorical values. The prevalence ratio was calculated using binomial logistic

116 regression models. Linear regression models and the Cochran Armitage trend test were utilized
 117 to examine trends and contrast two groups. A p-value <0.05 was utilized to determine the
 118 statistical significance of the variables. All statistical tests were conducted using the free and
 119 open-source software R. (version 4.2.1).

120

121 Results

122 We analyzed 395 million weighted hospital-discharged patients aged >18 or older from 2008-
 123 2020 NIS data. Overall, we found 47,506 hospital encounters with a primary or secondary
 124 diagnosis of HPV. Regarding the characteristics of the patients, most were females, 99.72%,
 125 61.2% were between 26-40 years old, and 41.59% were Whites. HPV-related hospital discharges
 126 were more prominent among low-income individuals (35.33%) than high-income individuals
 127 (17.78%). The prevalence was higher in the southern United States (44.72%). In addition, among
 128 those patients with HPV, 19.3% had anemia, 15.4% were obese, and 10.2% were tobacco users
 129 (Table 1).

130 **Table 1.** Demographic characteristics of patients with high-risk and low-risk HPV in the United
 131 States, 2008 - 2020.

Variables	HPV diagnosis (n=47,506)		High risk HPV (n=31,492)		Low risk HPV (n=16,014)	
	n (%)	95% confidence interval	n (%)	95% confidence interval	n (%)	95% confidence interval
	Weighted frequency (%)					
NIS data year						
2008	3,212 (6.76%)	6.54 - 6.99	1,947 (6.18%)	5.92 - 6.45	1,274 (7.92%)	7.51 - 8.33
2009	2,123 (4.47%)	4.28 - 4.65	1,366 (4.34%)	4.11 - 4.56	762 (4.74%)	4.41 - 5.07
2010	2,605 (5.48%)	5.28 - 5.69	1,542 (4.90%)	4.66 - 5.13	1,068 (6.64%)	6.26 - 7.02
2011	3,201 (6.74%)	6.51 - 6.96	1,067 (3.39%)	3.19 - 3.59	2,144 (13.34%)	12.81 - 14.87
2012	2,770 (5.83%)	5.62 - 6.04	1,505 (4.78%)	4.54 - 5.01	1,265 (7.87%)	7.45 - 8.29

2013	2,970 (6.25%)	6.03 - 6.47	1,575 (5.0%)	4.76 - 5.24	1,395 (8.68%)	8.24
2014	3,495 (7.36%)	7.12-7.59	2,065 (6.56%)	6.28 - 6.83	1,435 (8.93%)	8.48
2015	3,740 (7.87%)	7.63 - 8.11	2,440 (7.75%)	7.45 - 8.04	1,310 (8.15%)	7.73
2016	3,785 (7.97%)	7.72 - 8.21	2,740 (8.70%)	8.39 - 9.01	1,045 (6.50%)	6.12
2017	4,610 (9.70%)	9.44 - 9.97	3,500 (11.11%)	10.77 - 11.46	1,115 (6.94%)	6.54
2018	5,035 (10.60%)	10.32 - 10.88	4,135 (13.13%)	12.76 - 13.50	9,10 (5.66%)	5.3
2019	4,775 (10.05%)	9.78 - 10.32	3,570 (11.34%)	10.99 - 11.69	1,205 (7.49%)	7.09
2020	5,185 (10.91%)	10.63 - 11.19	4,040 (12.83%)	12.46 - 13.20	1,150 (7.15%)	6.75
Sex						
Male	114 (0.24%)	0.20 - 0.28	89 (0.28%)	0.22 - 0.34	25 (0.16%)	0.09
Female	47,387 (99.76%)	99.72 - 99.80	31,402 (99.72%)	99.66 - 99.78	16,048 (99.84%)	99.7
Age at discharge, years						
18-25 years	13,037 (27.44%)	27.04 - 27.84	8,515 (27.04%)	26.55 - 27.53	4,557 (28.35%)	27.6
26-40 years	29,074 (61.20%)	60.76 - 61.64	19,492 (61.90%)	61.36 - 62.43	9,605 (59.74%)	58.9
41-64 years	4,672 (9.83%)	9.57 - 10.10	3,056 (9.71%)	9.38 - 10.03	1,621 (10.08%)	9.61
≥65 years	723 (1.52%)	1.41 - 1.63	428 (1.36%)	1.23 - 1.49	295 (1.83%)	1.63
Race						
White	18,994 (41.59%)	41.14 - 42.04	12,318 (40.65%)	40.10- 41.20	6,696 (43.39%)	42.6
Black	10,754 (23.54%)	23.16 - 23.93	7,790 (25.71%)	25.22 - 26.20	2,977 (19.29%)	18.6
Hispanic	11,855 (25.96%)	25.55 - 26.36	7,417 (24.48%)	23.99 - 24.96	4,467 (28.94%)	28.2
Asian or Pacific Islander	1,657 (3.63%)	3.46 - 3.80	1,166 (3.85%)	3.63 - 4.07	490 (3.18%)	2.90
Native American	353 (0.77%)	0.69 - 0.85	258 (0.85%)	0.75-0.95	96 (0.62%)	0.50
Other	2,060 (4.51%)	4.32 - 4.70	1,353 (4.46%)	4.23 - 4.70	708 (4.59%)	4.26
Primary Payer						
Medicare	1,554 (3.28%)	3.12 - 3.44	1,038 (3.31%)	3.11 - 3.50	516 (3.22%)	2.94
Medicaid	25,168 (53.11%)	52.66 - 53.56	16,876 (53.76%)	53.21 - 54.31	8,317 (51.80%)	51.0
Private Insurance	17,688 (37.33%)	36.89 - 37.76	11,499 (36.63%)	36.1 - 37.16	6,213 (38.69%)	37.9
Self-pay	1,571 (3.32%)	3.15 - 3.48	1,061 (3.38%)	3.18 - 3.58	515 (3.21%)	2.94
No charge	216 (0.45%)	0.39 - 0.52	114 (0.36%)	0.30 - 0.43	102 (0.63%)	0.51
Other	1,190 (2.51%)	2.37 - 2.65	805 (2.56%)	2.39 - 2.74	395 (2.46%)	2.22
Median household Income						

<50,000	16,571 (35.33%)	34.9 - 35.77	11,281 (36.29%)	35.75 - 36.82	5,324 (33.54%)	32.8
50,000-64,999	11,328 (24.15%)	23.77 - 24.54	7,505 (24.14%)	23.67 - 24.62	3,827 (24.11%)	23.4
65,000-85,999	10,658 (22.73%)	22.35 - 23.11	7,081 (22.78%)	22.31 - 23.24	3,587 (22.60%)	21.9
>86,000	8,340 (17.78%)	17.44 - 18.13	5,221 (16.79%)	16.38 - 17.21	3,134 (19.75%)	19.1
US Region						
Northeast	7,230 (15.22%)	14.9 - 15.54	4,294 (13.63%)	13.26 - 14.01	2,950 (18.35%)	17.7
Midwest	7,631 (16.06%)	15.73 - 16.39	5,416 (17.2%)	16.78 - 17.62	2,225 (13.84%)	13.3
South	21,246 (44.72%)	44.28 - 45.17	14,486 (46.0%)	45.45 - 46.55	6,780 (42.17%)	41.4
West	11,399 (23.99%)	23.61 - 24.38	7,296 (23.17%)	22.7 - 23.63	4,123 (25.65%)	24.9
Length of Hospital stay						
0 - 3 days	38,949 (81.99%)	81.64 - 82.33	25,949 (82.40%)	81.98 - 82.82	13,054 (81.19%)	80.5
4 - 6 days	6,657 (14.01%)	13.7 - 14.33	4,366 (13.87%)	13.48 - 14.25	2,301 (14.31%)	13.7
7 - 9 days	878 (1.85%)	1.73 - 1.97	550 (1.75%)	1.6 - 1.89	328 (2.04%)	1.82
10 - 12 days	303 (0.64%)	0.57 - 0.71	180 (0.57%)	0.49 - 0.66	123 (0.76%)	0.63
>13 or more	718 (1.51%)	1.4 - 1.62	446 (1.42%)	1.29 - 1.55	272 (1.69%)	1.49
Died during hospitalization	45 (0.1%)	0.07 - 0.12	40 (0.13%)	0.09 - 0.17	5 (0.03%)	0.0
Comorbidities						
Drug	2,468 (5.55%)	5.34 - 5.76	1,702 (5.79%)	5.52 - 6.05	766 (5.06%)	4.71
Alcohol	489 (1.10%)	1.00 - 1.20	293 (1.0%)	0.88 - 1.11	196 (1.30%)	1.12
Tobacco	4,834 (10.18%)	9.90 - 10.45	3,275 (10.40%)	10.06 - 10.74	1,565 (9.73%)	9.27
Anemia	7,570 (19.26%)	18.87 - 19.6	4,609 (18.16%)	17.69 - 18.64	2,971 (21.26%)	20.5
Arthritis	434 (1.10%)	1.00 - 1.21	293 (1.16%)	1.02 - 1.29	141 (1.01%)	0.84
Chronic pulmonary disease	4476 (10.06%)	9.78 - 10.34	3062 (10.41%)	10.06 - 10.76	1429 (9.45%)	8.98
Congestive heart failure	165 (0.42%)	0.36 - 0.48	115 (0.45%)	0.37 - 0.54	50 (0.36%)	0.26
Coagulopathy	946 (2.41%)	2.26 - 2.56	592 (2.33%)	2.15 - 2.52	354 (2.54%)	2.27
Depression	3,487 (7.84%)	7.59 - 8.09	2,391 (8.13%)	7.82 - 8.44	1,096 (7.25%)	6.83
Diabetes	2,104 (4.73%)	4.53 - 4.93	1,463 (4.97%)	4.72 - 5.22	641 (4.24%)	3.92
Hypertension	3,537 (7.95%)	7.70 - 8.20	2,320 (7.89%)	7.58 - 8.19	1,217 (8.05%)	7.61
Liver disease	368 (0.94%)	0.84 - 1.03	213 (0.84%)	0.73 - 0.95	155 (1.11%)	0.94

Obesity	6,863 (15.43%)	15.09 - 15.77	4,791 (16.29%)	15.86 -16.71	2,072 (13.7%)	13.1
Peripheral vascular disorders	170 (0.38%)	0.32 - 0.44	124 (0.42%)	0.35 - 0.50	45 (0.30%)	0.21
Pulmonary circulation disorders	60 (0.15%)	0.12 - 0.19	25 (0.10%)	0.06 - 0.14	35 (0.25%)	0.17
Renal failure	325 (0.83%)	0.74 - 0.92	199 (0.79%)	0.68 - 0.89	125 (0.90%)	0.74

132
 133 Patients discharged from hospitals in the United States were 84 times more likely to have a
 134 primary or secondary diagnosis of HPV in 2020 compared to 2008 (95% CI: 1.67-2.03). The
 135 prevalence ratio of HPV was higher among females than males (PR: 295.27; 95% CI: 201.28 -
 136 457.90). While the number of hospital discharges with an HPV-associated diagnosis was greatest
 137 among White individuals, the prevalence ratio of HPV was 398 times higher among Hispanics
 138 than White individuals (95% CI: 3.78-4.18), followed by Blacks (PR: 228; 95% CI: 2.45-2.72).
 139 The prevalence ratio was also higher for the patients with Medicaid (PR: 45.14; 95% CI: 40.35 –
 140 50.70). With the regions, the prevalence of HPV was 45 times higher in the South compared to
 141 the Northeast (95% CI: 1.36-1.53) and 57 times higher in the West compared to the Northeast
 142 (95% CI: 1.47-1.68) (Table 2).

143 **Table 2.** Prevalence ratio for HPV-related patient discharges and sociodemographic covariates in
 144 the United States, 2008 – 2020.

Variables	Prevalence Ratio (Unadjusted)	95% confidence interval		P-value
		Lower	Upper	
NIS data year				< 0.000
2008	Reference			
2009	0.68	0.61	0.77	
2010	0.8	0.72	0.90	
2011	1	0.90	1.11	
2012	0.89	0.79	0.99	

2013	0.97	0.87	1.09	
2014	1.15	1.04	1.28	
2015	1.22	1.10	1.35	
2016	1.23	1.11	1.37	
2017	1.49	1.35	1.64	
2018	1.64	1.49	1.80	
2019	1.55	1.41	1.71	
2020	1.84	1.67	2.03	
Sex				< 0.000
Male	Reference			
Female	295.27	201.28	457.90	
Age (years)				
18 – 25 years	Reference			
26 – 40 years	0.94	0.90	0.98	
41-64 years	0.08	0.08	0.09	
≥65 years	0.01	0.01	0.02	
Race				<0.000
White	Reference			
Black	2.59	2.45	2.72	
Hispanic	3.98	3.78	4.18	
Asian or Pacific Islander	2.29	2.05	2.56	
Native American	1.93	1.51	2.42	
Other	2.45	2.21	2.70	
Primary Payer				<0.000
Medicare	Reference			
Medicaid	45.14	40.35	50.70	
Private Insurance	18.62	16.62	20.94	
Self-pay	10.14	8.68	11.85	
No charge	13.96	10.05	18.91	
Other	11.46	9.69	13.55	
Median household Income, percentile				<0.000
0-25th	Reference			
26th-50th	0.78	0.74	0.82	
51th-75th	0.81	0.76	0.85	
76th-100th	0.74	0.70	0.78	
US Region				<0.000
Northeast	Reference			

Midwest	0.89	0.83	0.95	
South	1.45	1.36	1.53	
West	1.57	1.47	1.68	
Length of Hospital stay (days)				<0.000
0 - 3 days	Reference			
4 – 6 days	0.40	0.38	0.42	
7 – 9 days	0.14	0.12	0.17	
10 – 12 days	0.11	0.09	0.14	
>13 days or more	0.17	0.14	0.20	
Comorbidities				<0.000
Drug	1.20	1.10	1.32	
Alcohol	0.22	0.18	0.26)	
Tobacco	0.42	0.39	0.44	
Anemia	1.36	1.28	1.43	
Arthritis	0.40	0.32	0.49	
Chronic pulmonary disease	0.48	0.45	0.52	
Congestive heart failure	0.04	0.03	0.06	
Coagulopathy	0.54	0.47	0.62	
Depression	0.67	0.62	0.73	
Diabetes	0.16	0.14	0.17	
Hypertension	0.09	0.08	0.09	
Liver disease	0.30	0.24	0.38	
Obesity	1.22	1.15	1.29	
Peripheral vascular disorders	0.06	0.04	0.09	
Pulmonary circulation disorders	0.08	0.04	0.13	
Renal failure	0.06	0.05	0.08	

145

146 **Hospital Discharges with a Primary or Secondary HPV Diagnosis**

147 From 2008 to 2020, hospital discharges with a primary or secondary diagnosis of HPV increased
148 from 6.76% to 10.91% (Figure 1). Our study also revealed notable differences in HPV trends
149 across races, with white individuals having the highest number of HPV infections as compared to
150 individuals from other races ($p < 0.001$) (Figure 2). We observed a rising pattern among
151 individuals aged between 26 and 40 years ($p < 0.001$) (Figure 3). A greater number of HPV

152 diagnoses was found in hospital encounters among females compared to males ($p < 0.001$) (Figure
153 4).

154 Figure 1. Trends in HPV-associated hospital discharge in the United States from 2008-2020

155 Figure 2. Trends in HPV-associated hospital discharge in the United States by race from 2008-
156 2020

157 Figure 3. Trends in HPV-associated hospital discharge in the United States by age group from
158 2008-2020

159 Figure 4. Trends in HPV-associated hospital discharge in the United States by sex from 2008-
160 2020

162 **Discussion**

163 Using hospital discharge registry data from NIS, we observed a significant increase in the
164 incidence of a primary or secondary diagnosis of HPV in hospital discharges from 2008 to 2020
165 in the United States despite the introduction of HPV vaccination recommendations in 2006.
166 These increases were observed among women of younger age groups and the increase from 2008
167 to 2020 was highest in the 26 to 40 age group. These findings could indicate a greater ability to
168 detect HPV and a higher emphasis being placed on HPV testing over time; however, the
169 introduction of vaccination should theoretically have shown a decrease in the incidence of HPV-
170 associated diagnoses over the course of 12 years despite increased detection.

171 These results follow the American Cancer Society's findings that HPV vaccination rates remain
172 low in the United States (12). This could be due to insufficient education of patients by
173 healthcare providers about the risks of HPV infection or about the availability of HPV
174 vaccination. This could also indicate a lack of public priority on annual wellness visits where
175 these types of discussions may occur or distrust in the vaccine's safety and efficacy. This study
176 further supports the need to adopt evidence-based HPV vaccination programs, such as those
177 listed by the National Institute of Health as part of their evidence-based cancer control programs,
178 that promote HPV vaccination education in various communities (16).

179 Our findings differed from a prior data analysis from the 2013-2014 National Health and
180 Nutrition Examination Survey that showed HPV prevalence of any and high-risk genital HPV for
181 adults aged 18–59 were 45.2% and 25.1% in men and 39.9% and 20.4% in women, respectively
182 (17). Our results, conversely, showed a greater HPV infection prevalence among women than

183 men. This could be explained by differences in the length of time studied or due to obtaining
184 information from hospital discharges versus an examination survey (18, 19).

185 Differences in HPV diagnoses were also observed among various racial groups. Whites had the
186 highest number of hospitalizations associated with an HPV diagnosis, followed by Hispanic
187 individuals, then Blacks. Our estimate was similar to the study by Pierre-Victor, Trepka (20),
188 which found a higher HPV prevalence among non-Hispanic White females in the United States
189 between 14-59 years old. Some prior studies, however, revealed higher HPV infection rates
190 among non-Hispanic Black adults in the United States. For instance, one study showed that
191 Black women had twice the prevalence of HPV infection as other racial groups (21). The
192 difference in our results could be that other racial groups are more likely to go to the hospital
193 than Blacks. Due to our information being obtained from hospital discharges, this would have
194 resulted in an underreported percentage of HPV infection among Blacks. The fact that the
195 prevalence ratios of Blacks and Hispanics were higher than Whites in our results could be a
196 further testament to this concept.

197 Disparities associated with HPV infection rates were seen in lower socioeconomic groups.
198 Individuals with lower household incomes had a higher prevalence of HPV infections, while
199 higher incomes were associated with the lowest number of infections diagnosed. An association
200 was also found in the primary insurance of HPV-infected individuals, with most of those
201 diagnosed with HPV infection having Medicaid insurance as the primary payer. This correlates
202 with other studies that found low socioeconomic status as an associated risk factor for HPV
203 infection. When looking at a group of low-income, diverse young women, Shikary, Bernstein
204 (22) found the prevalence of HPV infection to be 68%, much higher than the national average of

205 45.2%. These results illustrate that patients with lower socioeconomic status are at an increased
206 risk of HPV infection and should, therefore have increased resources for HPV education,
207 vaccination, and screening.

208 Geographically, more infections were observed in the southern regions of the United States.
209 Currently, the South has one of the lowest HPV vaccination rates in the USA (23). Moreover, the
210 high-risk type of HPV infection is higher in the South than in other regions. (20). Benard,
211 Thomas (24) found the highest levels of HPV-associated cervical cancer in Arkansas, Oklahoma,
212 Alabama, Tennessee, and Texas. These results indicate that additional efforts are needed to
213 address HPV prevention and early detection in the southern United States.

214 Vaccination is currently the most effective method to address the increasing incidence of HPV
215 infection in the United States, and early screening is needed to detect HPV-related neoplastic
216 transformation at its early stages (6). Additional research is needed to identify barriers to HPV
217 vaccination and to investigate why HPV vaccination rates remain lower than other routine
218 vaccinations. In addition, clinical trials are needed to determine how to achieve efficacious early
219 screening for non-cervical, HPV-related cancers. Future studies should also investigate the most
220 effective methods of disseminating information about HPV infection, vaccination, and screening
221 to those with significant risk factors who are unlikely to attend regular wellness visits for early
222 screening and vaccination.

223 **Limitations**

224 A few significant elements in this study's design and analysis could result in misunderstanding.
225 No longitudinal data are available due to NIS's inability to identify specific patients; ergo,
226 numerous recordings may be made because of the same patient's repeated hospitalizations. Thus,

227 it is important to understand that the epidemiologic data presented shows the trends and
228 characteristics leading to hospitalization associated with a primary or secondary diagnosis of
229 HPV, not the trends and characteristics of becoming infected with HPV. Furthermore, NIS data
230 does not consider outpatient interactions and observation-only stays, so there may be an
231 underrepresentation of patients infected with HPV who are being monitored only in an outpatient
232 setting. In addition, the sampling methodology forbids the estimation of volumes for subsets (for
233 example, state-level estimates cannot be derived by applying discharge weights to all sampled
234 hospitals/patients within a state). Finally, the administrative, diagnostic, and procedural codes
235 utilized by NIS for the linked hospitalizations are supplied on the billing record. They are not
236 individually validated due to changes in methodology in 2012 that prevented hospital-level
237 estimates from being produced.

238

239 **Conclusion**

240 Overall, the prevalence of HPV has increased in the United States in the past 12 years.
241 Understanding the trends and characteristics of HPV infection can contribute to public health
242 interventions aimed at prevention and early detection. Efforts to improve HPV vaccination rates
243 and address risk factors associated with HPV infection are crucial to reducing the burden of HPV-
244 related diseases. Further research can be conducted among those at higher risk to investigate the
245 underlying factors responsible for the disparities and increasing rates of HPV observed in this
246 study.

247

248 **Appendices**

249 **Appendix A: ICD-9 and ICD-10 code description for HPV:**

250 ICD-9 codes

251 795.05 = cervical high risk human papillomavirus; DNA test positive

252 795.09 = other abnormal Papanicolaou smear of cervix and cervical human papillomavirus;
253 cervical low-risk human papillomavirus; DNA test positive.

254 796.75 = anal high risk human papillomavirus; DNA test positive

255 796.79 = other abnormal Papanicolaou smear of anus and anal HPV; anal low-risk human
256 papillomavirus; DNA test positive.

257 ICD-10 codes

258 R87.810 = high-risk human papillomavirus DNA test positive

259 R87.820 = cervical low-risk human papillomavirus DNA test positive

260 R85.81= high-risk human papillomavirus DNA test positive from female genital organs

261 R85.82 = low-risk human papillomavirus DNA test positive from female genital organs

262

263 **References**

264 1. Arbyn M, Weiderpass E, Bruni L, de Sanjosé S, Saraiya M, Ferlay J, et al. Estimates of
265 incidence and mortality of cervical cancer in 2018: a worldwide analysis. *The Lancet Global*
266 *Health*. 2020;8(2):e191-e203.

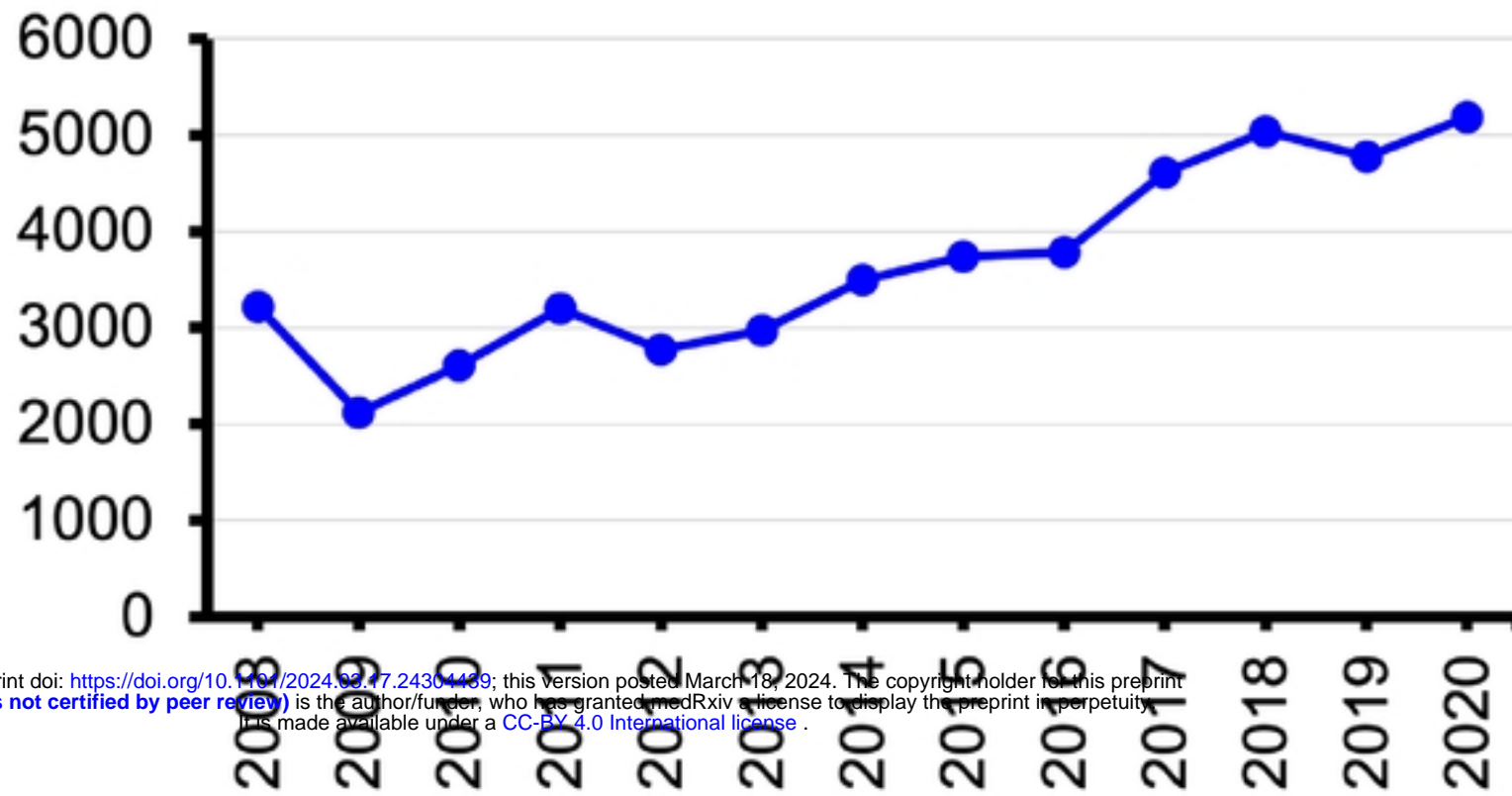
267 2. Bruni L, Diaz M, Castellsagué M, Ferrer E, Bosch FX, de Sanjosé S. Cervical human
268 papillomavirus prevalence in 5 continents: meta-analysis of 1 million women with normal
269 cytological findings. *Journal of Infectious Diseases*. 2010;202(12):1789-99.

270 3. Roman BR, Aragones A. Epidemiology and incidence of HPV-related cancers of the
271 head and neck. *Journal of Surgical Oncology*. 2021;124(6):920-2.

- 272 4. Meites E, Szilagyi PG, Chesson HW, Unger ER, Romero JR, Markowitz LE. Human
273 Papillomavirus Vaccination for Adults: Updated Recommendations of the Advisory Committee
274 on Immunization Practices. *MMWR Morb Mortal Wkly Rep.* 2019;68(32):698-702.
- 275 5. Centers for Disease Control and Prevention. Cancers Associated with Human
276 Papillomavirus, United States, 2015–2019. Atlanta, GA: Centers for Disease Control and
277 Prevention, US Department of Health and Human Services.; 2022 October 24, 2023. Contract
278 No.: USCS Data Brief,.
- 279 6. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA: a cancer journal*
280 *for clinicians.* 2023;73(1):17-48.
- 281 7. Winer RL, Lee S-K, Hughes JP, Adam DE, Kiviat NB, Koutsky LA. Genital human
282 papillomavirus infection: incidence and risk factors in a cohort of female university students.
283 *American journal of epidemiology.* 2003;157(3):218-26.
- 284 8. Senkomago V, Henley SJ, Thomas CC, Mix JM, Markowitz LE, Saraiya M. Human
285 papillomavirus–attributable cancers—United States, 2012–2016. *Morbidity and Mortality*
286 *Weekly Report.* 2019;68(33):724.
- 287 9. Steinau M, Saraiya M, Goodman MT, Peters ES, Watson M, Cleveland JL, et al. Human
288 papillomavirus prevalence in oropharyngeal cancer before vaccine introduction, United States.
289 2014.
- 290 10. Xu L, Dahlstrom KR, Lairson DR, Sturgis EM. Projected oropharyngeal carcinoma
291 incidence among middle-aged US men. *Head & Neck.* 2019;41(9):3226-34.

- 292 11. World Health Organization. Cervical Cancer. 2022 February 22.
- 293 12. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA: a cancer journal*
294 *for clinicians*. 2022;72(1):7-33.
- 295 13. Ghebre R, Berry-Lawhorn JM, D'Souza G. State of the science: screening, surveillance,
296 and epidemiology of HPV-related malignancies. *American Society of Clinical Oncology*
297 *Educational Book*. 2021;41:377-88.
- 298 14. Agency for Healthcare Research and Quality. HCUP using multiple years of data tutorial
299 2023 [Available from: https://hcup-us.ahrq.gov/tech_assist/trends/508/508course_2019.jsp.
- 300 15. Houchens R R, D, Elixhauser A, Jiang J. . Nationwide Inpatient Sample (NIS) Redesign
301 Final Report. 2014. HCUP Methods Series Report # 2014-04 ONLINE. . U.S. Agency for
302 Healthcare Research and Quality.; 2014 April 4, 2014.
- 303 16. NIH. Evidence-Based Cancer Control Programs (EBCCP), HPV Vaccination Evidence-
304 Based Programs Listing. 2023 [Available from:
305 <https://ebccp.cancercontrol.cancer.gov/topicPrograms.do?topicId=22626661&choice=defaultThi>
306 [s](#)
- 307 17. McQuillan GM, Kruszon-Moran D, Markowitz LE, Unger ER, Paulose-Ram R.
308 Prevalence of HPV in adults aged 18-69: United States, 2011-2014: US Department of Health
309 and Human Services, Centers for Disease Control and ...; 2017.

- 310 18. Gerend MA, Shepherd MA, Lustria MLA. Increasing human papillomavirus vaccine
311 acceptability by tailoring messages to young adult women's perceived barriers. Sexually
312 transmitted diseases. 2013;40(5):401-5.
- 313 19. Beavis AL, Levinson KL. Preventing cervical cancer in the United States: barriers and
314 resolutions for HPV vaccination. *Frontiers in Oncology*. 2016;6:19.
- 315 20. Pierre-Victor D, Trepka MJ, Stephens D, Li T, Page T, Madhivanan P. Human
316 papillomavirus infections in the Southern and other United States regions. *Journal of Health
317 Disparities Research and Practice*. 2018;11(2):1.
- 318 21. Shi R, Devarakonda S, Liu L, Taylor H, Mills G. Factors associated with genital human
319 papillomavirus infection among adult females in the United States, NHANES 2007–2010. *BMC
320 research notes*. 2014;7(1):1-8.
- 321 22. Shikary T, Bernstein DI, Jin Y, Zimet GD, Rosenthal SL, Kahn JA. Epidemiology and
322 risk factors for human papillomavirus infection in a diverse sample of low-income young
323 women. *Journal of Clinical Virology*. 2009;46(2):107-11.
- 324 23. Adjei Boakye E, Babatunde OA, Wang M, Osazuwa-Peters N, Jenkins W, Lee M, et al.
325 Geographic Variation in Human Papillomavirus Vaccination Initiation and Completion Among
326 Young Adults in the U.S. *Am J Prev Med*. 2021;60(3):387-96.
- 327 24. Benard VB, Thomas CC, King J, Massetti GM, Doria-Rose VP, Saraiya M. Vital signs:
328 cervical cancer incidence, mortality, and screening—United States, 2007–2012. *Morbidity and
329 mortality weekly report*. 2014;63(44):1004.
- 330



medRxiv preprint doi: <https://doi.org/10.1101/2024.03.17.24304439>; this version posted March 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Figure 1. Trends in HPV-associated hospital discharge in the United States from 2008-2020

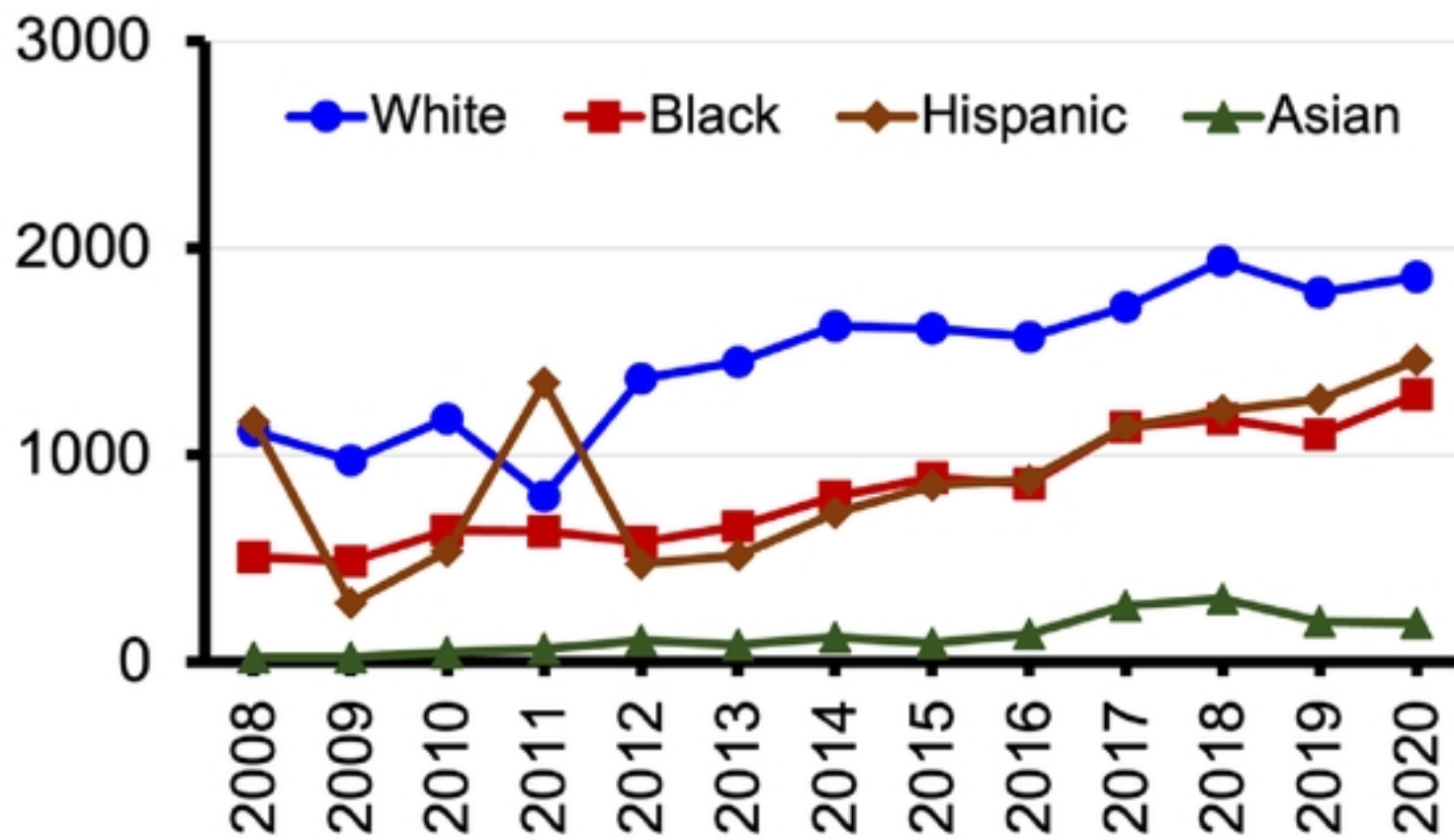
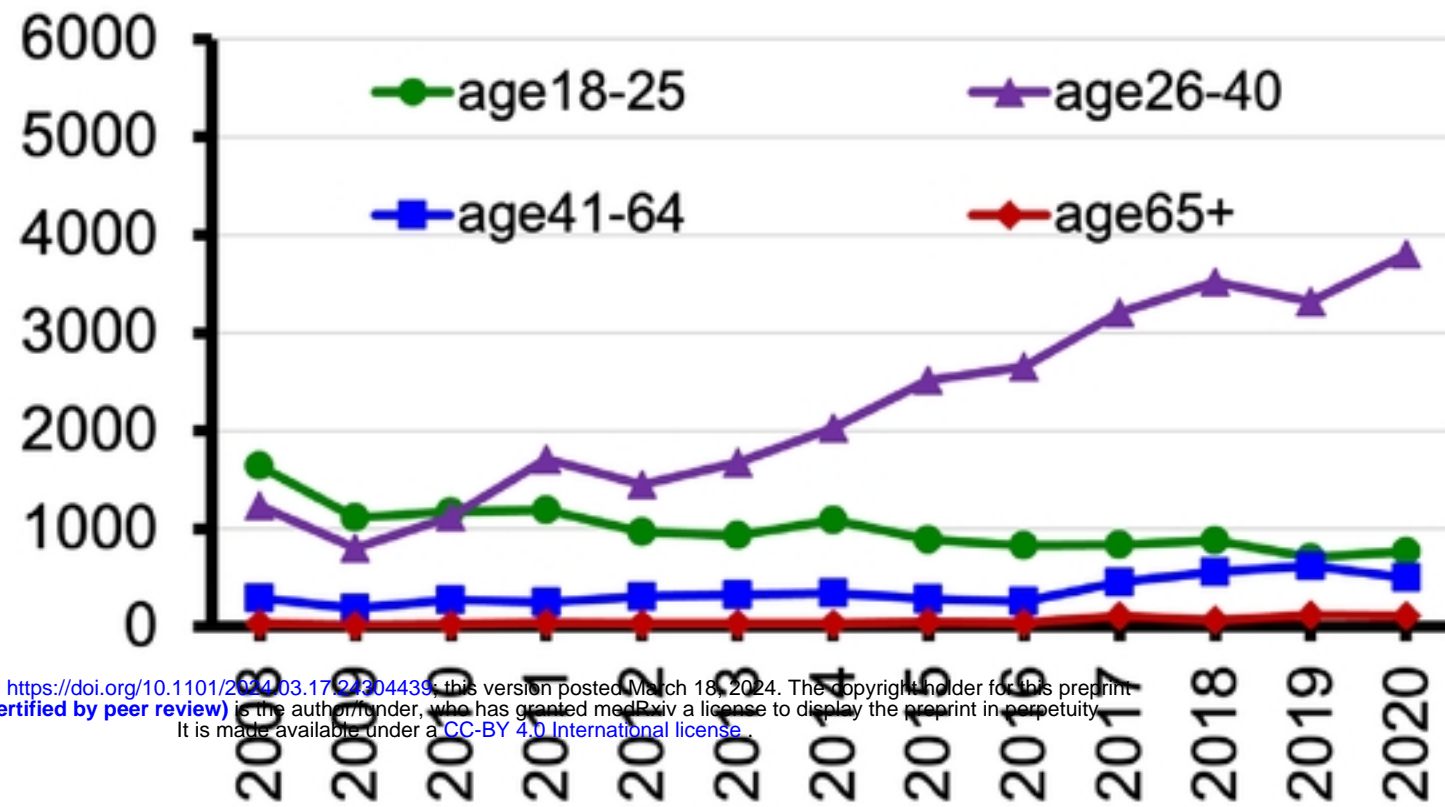


Figure 2. Trends in HPV-associated hospital discharge in the United States by race from 2008-2020



medRxiv preprint doi: <https://doi.org/10.1101/2024.03.17.2404439>; this version posted March 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Figure 3. Trends in HPV-associated hospital discharge in the United States by age group from 2008-2020

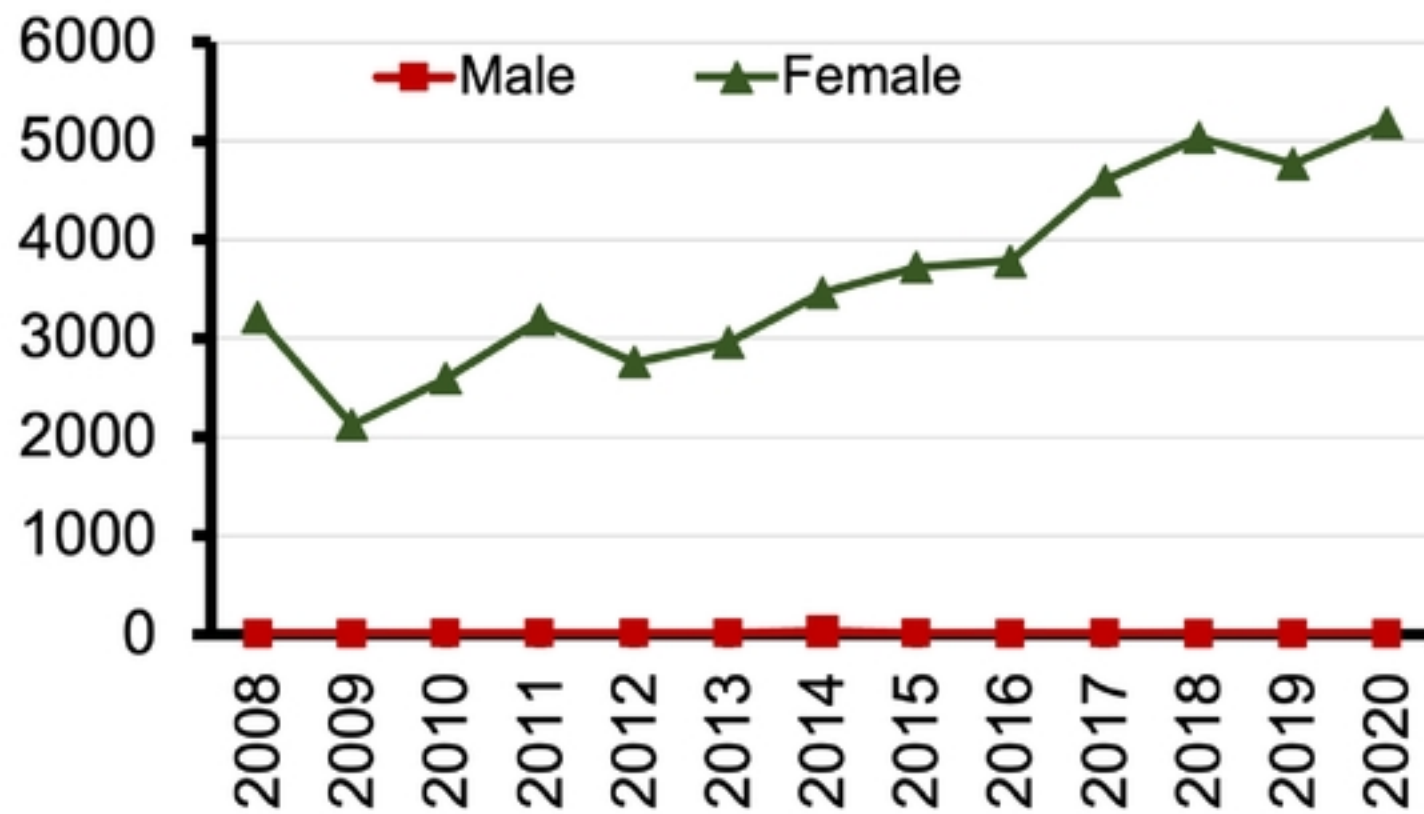


Figure 4. Trends in HPV-associated hospital discharge in the United States by sex from 2008-2020