

Prevalence of HPV Infection Among Females in the United States

Eileen F. Dunne, MD, MPH

Elizabeth R. Unger, PhD, MD

Maya Sternberg, PhD

Geraldine McQuillan, PhD

David C. Swan, PhD

Sonya S. Patel, BS

Lauri E. Markowitz, MD

HUMAN PAPILLOMAVIRUS (HPV) is estimated to be the most common sexually transmitted infection in the United States.¹ HPV prevalence has been found to be highest among young persons within the first few years after sexual debut.²⁻⁶ However, there are no data on the prevalence of HPV among women across a broad age range and representative of the US population.

Genital HPV types are categorized according to their epidemiological association with cervical cancer. Infections with low-risk types, such as HPV types 6 and 11, can cause benign or low-grade changes in cells of the cervix, genital warts, and recurrent respiratory papillomatosis. High-risk HPV types can cause cervical, anal, and other genital cancers. High-risk HPV types are detected in 99% of cervical cancers, and worldwide approximately 70% of cervical cancers are due to HPV types 16 and 18.^{7,8} Although HPV infection is common, studies suggest approximately 90% of infections clear within 2 years.^{9,10}

A highly efficacious prophylactic vaccine against HPV types 6, 11, 16, and 18 was licensed in June 2006 and recommended for routine use in

See also p 876 and Patient Page.

Context Human papillomavirus (HPV) infection is estimated to be the most common sexually transmitted infection. Baseline population prevalence data for HPV infection in the United States before widespread availability of a prophylactic HPV vaccine would be useful.

Objective To determine the prevalence of HPV among females in the United States.

Design, Setting, and Participants The National Health and Nutrition Examination Survey (NHANES) uses a representative sample of the US noninstitutionalized civilian population. Females aged 14 to 59 years who were interviewed at home for NHANES 2003-2004 were examined in a mobile examination center and provided a self-collected vaginal swab specimen. Swabs were analyzed for HPV DNA by L1 consensus polymerase chain reaction followed by type-specific hybridization. Demographic and sexual behavior information was obtained from all participants.

Main Outcome Measures HPV prevalence by polymerase chain reaction.

Results The overall HPV prevalence was 26.8% (95% confidence interval [CI], 23.3%-30.9%) among US females aged 14 to 59 years (n=1921). HPV prevalence was 24.5% (95% CI, 19.6%-30.5%) among females aged 14 to 19 years, 44.8% (95% CI, 36.3%-55.3%) among women aged 20 to 24 years, 27.4% (95% CI, 21.9%-34.2%) among women aged 25 to 29 years, 27.5% (95% CI, 20.8%-36.4%) among women aged 30 to 39 years, 25.2% (95% CI, 19.7%-32.2%) among women aged 40 to 49 years, and 19.6% (95% CI, 14.3%-26.8%) among women aged 50 to 59 years. There was a statistically significant trend for increasing HPV prevalence with each year of age from 14 to 24 years ($P<.001$), followed by a gradual decline in prevalence through 59 years ($P=.06$). HPV vaccine types 6 and 11 (low-risk types) and 16 and 18 (high-risk types) were detected in 3.4% of female participants; HPV-6 was detected in 1.3% (95% CI, 0.8%-2.3%), HPV-11 in 0.1% (95% CI, 0.03%-0.3%), HPV-16 in 1.5% (95% CI, 0.9%-2.6%), and HPV-18 in 0.8% (95% CI, 0.4%-1.5%) of female participants. Independent risk factors for HPV detection were age, marital status, and increasing numbers of lifetime and recent sex partners.

Conclusions HPV is common among females in the United States. Our data indicate that the burden of prevalent HPV infection among females was greater than previous estimates and was highest among those aged 20 to 24 years. However, the prevalence of HPV vaccine types was relatively low.

JAMA. 2007;297:813-819

www.jama.com

females aged 11 to 12 years in the United States.¹¹⁻¹⁴ Clinical studies of the quadrivalent HPV vaccine demonstrated close to 100% efficacy in preventing infection and disease (cervical cancer precursors, genital lesions) associated with types included in the vaccine in analyses restricted to those women who were naive to HPV types 6, 11, 16, or 18 (either by HPV DNA or HPV antibodies).

Representative data on type-specific prevalence of HPV DNA detection in the United States could provide a baseline estimate to mea-

Author Affiliations: Centers for Disease Control and Prevention, Atlanta, Ga (Drs Unger and Swan, and Ms Patel); and National Center for Health Statistics, Centers for Disease Control and Prevention, Bethesda, Md (Dr McQuillan).

Corresponding Author: Eileen F. Dunne, MD, MPH, Centers for Disease Control and Prevention, 1600 Clifton Rd, MS E-02, Atlanta, GA 30333 (dde9@cdc.gov).

sure the wide-scale impact of the vaccine for reducing infection and could help guide models evaluating impact and cost effectiveness. With widespread implementation of the prophylactic HPV vaccine, decreases in the prevalence of vaccine HPV types would be expected. To determine a prevaccine population-based prevalence of cervicovaginal HPV in the United States, we performed HPV DNA testing on self-collected vaginal swabs among females participating in the National Health and Nutrition Examination Survey (NHANES) 2003-2004.

METHODS

Study Population and Study Design

NHANES is conducted by the National Center for Health Statistics, Centers for Disease Control and Prevention, and uses a representative sample of the US noninstitutionalized civilian population. The representative sample is obtained by using a complex, stratified, multistage probability sample design with unequal probabilities of selection to obtain a nationally representative sample.¹⁵ Certain subgroups of people, such as adolescents, non-Hispanic blacks, and Mexican Americans are oversampled. All females aged 14 to 59 years selected for NHANES 2003-2004 were eligible for participation in this study. Of 2482 females aged 14 to 59 years who were interviewed at home for the 2003-2004 cycle, 2387 (96.2%) were examined in a mobile examination center. A total of 2026 females (81.6%) submitted cervicovaginal swab specimens. Four hundred sixty-six females (23.0%) were considered nonresponders because they either submitted an inadequate swab specimen (n=105) or they did not submit a swab specimen (n=361). There were significant differences between nonresponders and responders on some demographic and behavioral variables (nonresponders were significantly more likely than responders to be of other race/ethnicity, to be younger

[<40 years], to be born outside the United States or Mexico, and to have never had sex). Written informed consent was obtained from all participants, and parental permission for those females younger than 18 years. This human subjects research was approved by the Centers for Disease Control and Prevention institutional review board.

Demographic and Behavioral Data

Race and ethnicity were self-reported into categories, including non-Hispanic black, non-Hispanic white, and Mexican American. Poverty index ratio was calculated by dividing total family income by the poverty threshold index, adjusted for family size at year of interview.¹⁶ Estimates of the total number of cases were generated by multiplying the appropriate population size from the January 2004 monthly postcensal civilian noninstitutionalized population¹⁷ by the weighted prevalence estimate among females in the appropriate age category.

Sex was defined as vaginal, oral, or anal sex. For those females who had at least 1 lifetime sex partner, additional questions were asked on age of first sex and lifetime number of sex partners. For those females aged 14 to 17 years who were sexually active, additional questions about sexual behavior were asked, including age at first sex, number of lifetime partners, and ever use of a condom. For women aged 18 years or older who were sexually active, additional questions were asked on number and gender of sex partners in the last 12 months and lifetime sex partners, and past history of sexually transmitted infections.

Specimen Collection and Processing

Females aged 14 to 59 years were asked to self-collect a cervicovaginal sample in the mobile examination center. In brief, each female was given a collection device, which was a small foam swab on a plastic handle packaged in an individual reclosable plastic

sleeve (Catch-All Sample Collection Swabs Epicenter, Madison, Wis). Participants were instructed to wash their hands before opening the swab, to hold the swab by the end of the handle, to insert the foam swab into the vagina similar to inserting a tampon, to gently turn the swab during a count of 10, and to replace the swab in the plastic sleeve, avoiding contact with the external genitalia. Participants took swabs and instructions into a bathroom and collected the samples in privacy. Swabs were given to NHANES personnel, stored at room temperature, and mailed to the Centers for Disease Control and Prevention laboratory for processing.

DNA Isolation

DNA was extracted using slight modifications of the QIAmp Mini Kit protocol (Qiagen, Valencia, Calif) within 1 month of sample collection. Briefly, swabs were incubated at 56°C for at least 12 hours in proteinase K lysis solution. One half of each sample was added to an equal volume of 100% ethanol and applied to each of 2 QIAmp Mini columns for DNA isolation. The eluates were collected, washed, and concentrated in Microcon 100 concentrators (Fisher Scientific, Pittsburgh, Pa), adjusting the final volume to 100 μ L with deionized distilled water. Samples were tested immediately or stored at -20°C. For every 40 samples, a water blank was processed through all steps of extraction to serve as a contamination control.

HPV Genotyping Test

HPV detection and typing was performed by using the Roche prototype line blot assay (reagents provided as a gift from Roche Molecular Systems Inc, Pleasanton, Calif). This assay uses HPV L1 consensus polymerase chain reaction with biotinylated PGMY09/11 primer sets and β -globin as an internal control for sample amplification.^{18,19} Five μ L of the DNA was used in the 100- μ L polymerase chain reaction. Amplicons (10 μ L) were evaluated for β -globin and HPV bands with

1.5% agarose gel electrophoresis stained with ethidium bromide, and those amplicons with an HPV band were hybridized to the typing strips. The first generation strip used from January 2003 to April 2004 included probes for 27 HPV types (6, 11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 45, 51, 52, 53, 54, 55, 56, 57, 58, 59, 66, 68, 73, 82, 83, and 84). The second generation strip used after April 2004 included additional types (61, 62, 64, 67, 69, 70, 71, 72, 81, 89, and IS39) but omitted HPV-57.²⁰ Samples that did not hybridize the strip were sequenced as previously described to determine HPV type.²¹ Samples negative for both β -globin and HPV ($n=105$, 5.2%) were considered inadequate for interpretation and were omitted from further analysis. We considered low-risk HPV types as HPV type 6, 11, 32, 40, 42, 44, 54, 55, 61, 62, 64, 71, 72, 74, 81, 83, 84, 87, 89, and 91; and high-risk HPV types as HPV type 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 67, 68, 69, 70, 73, 82, 85, and IS39.

Statistical Analysis

Females who submitted an adequate swab for HPV evaluation were included in the final analysis ($n=1921$). Statistical analyses were conducted using SAS version 9.1²² and SAS callable SUDAAN.²³ Variance estimates were calculated by using a Taylor series linearization that incorporated the complex sample design of the survey.²⁴ All estimates were weighted using the 2003-2004 medical examination weights provided by National Center for Health Statistics to account for the unequal probabilities of selection and adjustment for nonresponse. The weighting methodology has been described previously.²⁵ Because there were some missing laboratory specimens, we investigated whether any additional nonresponse adjustments to the original NHANES weights were needed. We found that using weights with an additional nonresponse and poststratification adjustment always provided prevalence estimates within the 95% confidence interval (CI) based on the

original NHANES weight; therefore, no additional adjustments to the NHANES weights were made.

We considered a prevalence estimate unreliable if the relative SE was more than 30% of the prevalence estimate; these estimates are not presented. Confidence intervals were calculated by using a log transformation with the SE of the log prevalence based on the delta method and applying SUDAAN estimated SEs.²⁶ Tests of association between HPV and the demographic or behavioral characteristics were based on the Wald χ^2 statistic. To compare the prevalence between HPV types, we applied a version of the McNemar test for complex surveys. No adjustments were made to the P values for multiple comparisons.

To explore the association with age and overall HPV prevalence, age was categorized into 4-year intervals (14-19, 20-24, 25-29, etc) and plotted against the log odds of HPV. The plot revealed a non-linear association between age and HPV prevalence. Prevalence increased up to age 20 to 24 years and then decreased. To test for the presence of an increasing linear trend among those females younger than 24 years and a decreasing linear trend for those females older than 24 years, 2 separate logistic regressions were used. Each logistic regression treated age as a continuous variable and was restricted to the age range under investigation. A trend was considered statistically significant if the β coefficient for the independent variable was non-zero at $P<.05$, using the Satterthwaite adjusted F test.

Demographic and behavioral characteristics of participants were evaluated. Characteristics significant at the $P<.20$ level based on a Wald χ^2 statistic were considered in a multivariate model. There are only 15 df available to develop a multivariate model, due to the complex survey design in the 2-year cycle of NHANES. To limit the df used when including age in the model, we collapsed age into 4 categories (18-19, 20-24, 25-29, and 30-59 years), because the prevalences for age categories between ages 30 and 59

years did not differ significantly. The multivariate model was limited to females aged 18 to 59 years because all questions regarding sexual exposures were asked of this group. We used SUDAAN for logistic regression to model independent associations between prevalence of any HPV and demographic and behavioral variables among sexually active females aged 18 to 59 years. We eliminated variables in a backward fashion that did not meet the criteria of $P<.05$ by Satterthwaite adjusted F test at each step. Any participants with missing data on variables included in the multivariate analysis were excluded. Once all variables in the model were statistically significant at the $P<.05$ level, all pairwise interactions were evaluated and retained only if the overall P value for the interaction was $<.05$. Goodness of fit for the final step of the model was assessed using the Hosmer-Lemeshow Satterthwaite adjusted F test.

RESULTS

HPV DNA Prevalence Overall and by Age

From 2003-2004, 2026 vaginal swab specimens from female NHANES participants aged 14 to 59 years were collected. Of the 1921 adequate specimens, 26.8% (95% CI, 23.3%-30.9%) were positive for any HPV DNA. Using 2000 Census data, this corresponds to 24.9 million females in this age range with prevalent HPV infection.

Prevalence of HPV infection was highest among females aged 20 to 24 years (44.8%; 95% CI, 36.3%-55.3%); overall prevalence among females aged 14 to 24 years was 33.8% (95% CI, 28.6%-40.0%) (TABLE 1). There was a statistically significant trend for increasing HPV prevalence with each year of age from 14 to 24 years ($P<.001$), which was followed by a nonsignificant gradual decline in HPV prevalence through 59 years ($P=.06$).

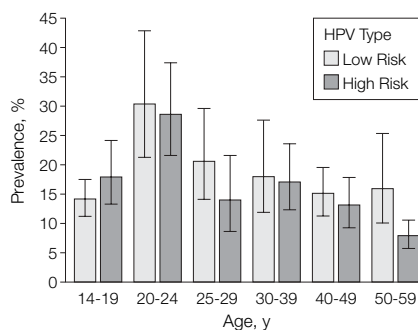
When the analysis was restricted to sexually active females, the prevalence of HPV was still highest among those aged 20 to 24 years. Among sexu-

Table 1. Prevalence of Any HPV Infection by Demographic Characteristics

Demographics	Sample Size	Prevalence, % (95% Confidence Interval)	P Value*
Overall (aged 14-59 years)	1921	26.8 (23.3-30.9)	
Age, y			
14-19	652	24.5 (19.6-30.5)	.003
20-24	189	44.8 (36.3-55.3)	
25-29	174	27.4 (21.9-34.2)	
30-39	328	27.5 (20.8-36.4)	
40-49	324	25.2 (19.7-32.2)	
50-59	254	19.6 (14.3-26.8)	
Race†			
Non-Hispanic white	837	24.2 (20.5-28.6)	<.001
Non-Hispanic black	533	39.2 (31.0-49.4)	
Mexican American	442	24.3 (19.3-30.6)	
Marital status			
Married	676	17.3 (14.0-21.5)	<.001
Widowed, divorced, separated	231	41.2 (32.3-52.4)	
Never married	882	31.1 (28.1-34.5)	
Living with partner	132	46.1 (35.2-60.4)	
Education‡			
<High school	383	35.0 (29.4-41.7)	.006
High school and GED	380	29.7 (23.4-37.6)	
>High school	754	24.7 (20.9-29.1)	
Poverty index§			
Below poverty	503	37.5 (29.9-47.1)	.005
At or above poverty	1322	24.4 (21.1-28.4)	
Country of birth			
United States	1580	26.8 (22.8-31.6)	.30
Mexico	206	19.6 (11.4-33.6)	
Other	135	30.8 (18.9-50.3)	

Abbreviations: GED, general equivalency diploma; HPV, human papillomavirus.
 *By Wald χ^2 statistic.
 †Race and ethnicity were self-reported; other race (n = 109) is not shown.
 ‡This analysis was limited to females older than 17 years.
 §Does not sum to 1921 because some of the responses are missing.
 ||Other includes any country other than the United States or Mexico.

Figure 1. Prevalence of Low-Risk and High-Risk HPV Types Among Females Aged 14 to 59 Years, NHANES 2003-2004



HPV indicates human papillomavirus; NHANES, National Health and Nutrition Examination Survey. Error bars indicate 95% confidence intervals. Both low-risk and high-risk HPV types were detected in some females. Low-risk HPV types are defined as HPV type 6, 11, 32, 40, 42, 44, 54, 55, 61, 62, 64, 71, 72, 74, 81, 83, 84, 87, 89, and 91; and high-risk HPV types as HPV type 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 67, 68, 69, 70, 73, 82, 85, and IS39.

ally active females, HPV prevalence was 39.6% (95% CI, 32.9%-47.8%) for 14 to 19 years, 49.3% (95% CI, 40.7%-59.6%) for 20 to 24 years, 27.8% (95% CI, 21.7%-35.7%) for 25 to 29 years, 27.3% (95% CI, 20.2%-36.8%) for 30 to 39 years, 23.9% (95% CI, 18.8%-30.5%) for 40 to 49 years, and 20.2% (95% CI, 14.6%-27.8%) for 50 to 59 years.

Prevalence of High-Risk, Low-Risk, and Specific Types

The overall prevalence of high- and low-risk HPV types was 15.2% and 17.8%, respectively. Prevalences of both low-risk and high-risk HPV types were highest in females aged 20 to 24 years (FIGURE 1). There was a statistically significant difference between low- and high-risk HPV types among females aged 14 to 19 years and 50 to 59 years.

Prevalence of high-risk types decreased after 20 to 29 years, and prevalence of low-risk types plateaued after 30 to 39 years.

The most common HPV types detected were HPV-62 (3.3%; 95% CI, 2.2%-5.1%) and HPV-84 (3.3%; 95% CI, 2.2%-5.1%), HPV-53 (2.8%; 95% CI, 2.1%-3.7%), and HPV-89 (2.4%; 95% CI, 1.4%-4.3%) and HPV-61 (2.4%; 95% CI, 1.6%-3.8%) (FIGURE 2). HPV-16 was detected in 1.5% (95% CI, 0.9%-2.6%) of females aged 14 to 59 years. There was no statistically significant difference in the prevalence of HPV-16 and the 13 more commonly detected types, except for HPV-84 and HPV-62. HPV-6 was detected in 1.3% (95% CI, 0.8%-2.3%), HPV-11 in 0.1% (95% CI, 0.03%-0.3%; relative SE $\geq 30\%$), and HPV-18 in 0.8% (95% CI, 0.4%-1.5%) of female participants. Most participants infected with HPV (60.1%) had only 1 HPV type detected (95% CI, 53.2%-67.9%); however, 23.9% had 2 types (95% CI, 18.3%-31.3%) and 16% had 3 or more types detected (95% CI, 12.0%-21.2%). Overall, HPV types 6, 11, 16, or 18 were detected in 3.4% of the study participants, corresponding with 3.1 million females with prevalent infection with HPV types included in the quadrivalent HPV vaccine. Few participants (0.10%) had both HPV types 16 and 18 and none had all 4 HPV vaccine types. At least 1 of these 4 HPV types was detected in 6.2% (95% CI, 3.8%-10.3%) of females aged 14 to 19 years.

Factors Associated With HPV DNA Detection

The variables associated with HPV DNA detection that were significant in the bivariate analysis were age, race, poverty index, education, marital status, and sexual behavior (Table 1 and TABLE 2).

HPV DNA was detected in 5.2% of females who reported never having had sex (Table 2). In an unweighted analysis, 88% of these females were between 14 to 19 years, and the remaining 12% were 40 to 49 years. Two of these participants were also positive for

herpes simplex type 2 by type-specific antibody testing,²⁷ and 1 female was positive for chlamydia by nucleic acid amplification testing.

The final multivariate model demonstrated that age younger than 25 years, marital status, and increasing numbers of recent or lifetime sex partners were independently associated with HPV detection (TABLE 3). Race/ethnicity was not significant in the multivariate model. Increasing numbers of recent sex partners was more strongly associated with HPV detection than was lifetime sex partners. No pairwise interactions among the variables in the final model were statistically significant below the $P < .05$ level.

COMMENT

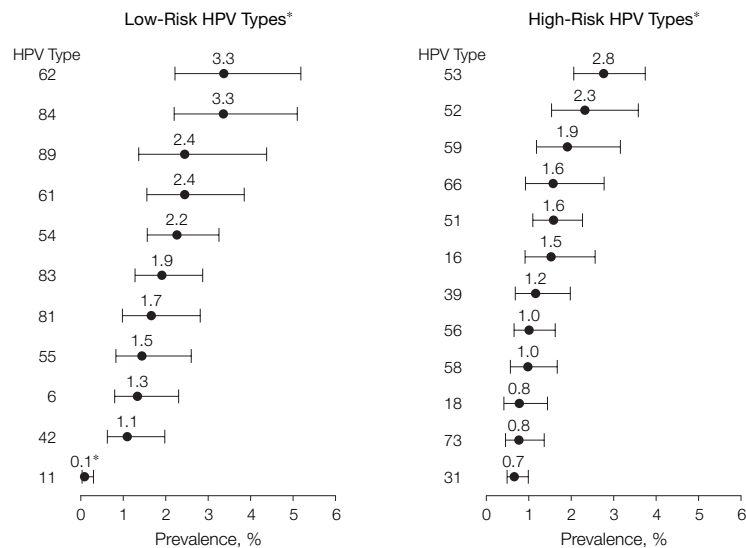
Prevalence of HPV DNA in a representative sample of US females aged 14 to 59 years was 26.8%, with the highest prevalence (44.8%) among women aged 20 to 24 years. The overall prevalence of HPV among females aged 14 to 24 years was 33.8%. This prevalence corresponds with 7.5 million females with HPV infection, which is higher than the previous estimate of 4.6 million prevalent HPV infections among females in this same age group in the United States.¹ The overall prevalences of high-risk HPV types 16 and 18 were 1.5% (95% CI, 0.9%-2.6%) and 0.8% (95% CI, 0.4%-1.5%), respectively.

We found that the prevalence of HPV infection increased from 14 years through 24 years, and then decreased at older ages. Most evaluations in the United States have found that the prevalence of any HPV was highest in the younger age groups (<20 years). Manhart et al⁶ in a population-based assessment of sexually active 18- to 25-year-old women using a urine sample found the prevalence of HPV-6 or HPV-11 to be 2.2%, and the prevalence of HPV-16 or HPV-18 to be 7.8%. Among sexually active 18- to 25-year-old women in our study, the prevalence of HPV-6 or HPV-11 (2.2%; 95% CI, 1.0%-5.0%; relative SE >30%) was similar to that found by Manhart et al,⁶ and the prevalence of

HPV-16 or HPV-18 (3.5%; 95% CI, 1.8%-6.7%; relative SE >30%) was lower. These differences may be due

to differences in the type of test used for the HPV detection, differences in the study population, or both.

Figure 2. Prevalence of HPV Types Among Females Aged 14 to 59 Years, NHANES 2003-2004



HPV indicates human papillomavirus; NHANES, National Health and Nutrition Examination Survey. Error bars indicate 95% confidence intervals.

*HPV types with a relative SE of more than 30% are not presented, except for HPV-11, since it is a vaccine type.

Table 2. Prevalence of Any HPV Infection by Behavioral Characteristics

Behavioral	Sample Size	Prevalence, % (95% Confidence Interval)	P Value*
Ever had sex			
Yes	1477	28.1 (24.6-32.1)	<.001
No	283	5.2 (2.4-11.2)	
Age at first sexual intercourse, y†			.047
<16	519	33.6 (27.5-41.1)	
≥16	953	26.2 (22.6-30.3)	
No. of lifetime sex partners‡			<.001
0	283	5.2 (2.4-11.2)	
1	349	11.5 (7.8-16.9)	
2	185	24.3 (16.5-35.7)	
3-5	430	32.0 (26.9-38.2)	
≥6	499	35.3 (29.7-42.0)	
Total No. of sex partners last year‡‡			<.001
0	137	18.7 (11.5-30.3)	
1	930	23.6 (20.7-27.0)	
2	115	51.0 (37.1-70.1)	
3-5	89	58.6 (42.7-80.4)	
≥6	30	55.6 (30.7-100.0)	
Health care professional ever told you had genital warts?††			.20
Yes	76	33.7 (26.0-43.7)	
No	1226	27.3 (23.4-31.8)	

Abbreviation: HPV, human papillomavirus.

*By Wald χ^2 statistic.

†Question was asked only to those females who had ever had sex.

‡Question was asked only to those females older than 18 years.

Table 3. Multivariate Analysis of Factors Associated With HPV Infection Among Sexually Active Females Aged 18 to 59 Years, NHANES 2003-2004*

Characteristic	Odds Ratio (95% Confidence Interval)	P Value†
Age, y		
18-19	1.41 (0.78-2.55)	.01
20-24	2.20 (1.33-3.62)	
25-29	0.81 (0.54-1.23)	
30-59	1 [Referent]	
Marital status		
Married	1 [Referent]	.003
Widowed, divorced, separated	3.09 (1.79-5.34)	
Never married	2.20 (1.39-3.49)	
Living with partner	3.37 (1.92-5.91)	
No. of lifetime sex partners		
1	1 [Referent]	.01
2	2.07 (0.98-4.38)	
≥3	2.74 (1.52-4.94)	
Total No. of sex partners in past year‡		
0	1 [Referent]	.02
1	2.05 (1.04-4.06)	
2	4.04 (1.79-9.11)	
≥3	4.12 (1.72-9.85)	

Abbreviations: HPV, human papillomavirus; NHANES, National Health and Nutrition Examination Survey.

*Includes 1288 participants who were 18 to 59 years, reported ever having sex, and who had complete information.

†Goodness of fit based on Hosmer-Lemeshow Satterthwaite adjusted $F_5 = 2.0$, $P = .13$.

‡By Hosmer-Lemeshow Satterthwaite adjusted F test.

§Question was asked only to those females older than 18 years.

We found the overall prevalence of HPV-16 to be low and that other HPV types were more prevalent. In most other studies, HPV-16 has been found to be the most prevalent type, although prevalence varied based on the population evaluated. Population-based studies outside the United States have found lower prevalence of HPV-16 than clinic-based studies have.²⁸ It is possible that our cervicovaginal assessment using the self-collected vaginal swabs was more likely to detect HPV types not related to cervical infection. Castle et al²⁸ hypothesized that there is tropism of some phylogenetic groupings (A3/A4/A15) to the vagina. In our assessment, the most common types detected were HPV-84 and HPV-62, both in the A3 phylogenetic grouping. Manhart et al⁶ found HPV types 84 and 62 were also frequently detected in urine samples among females aged 18 to 25 years.

Independent risk factors for HPV DNA detection in our analysis included sexual behavior (number of sex partners in the last year, number of life-

time sex partners) and demographic variables, including young age and marital status, consistent with risk factors for HPV detection found in other studies.

HPV DNA was detected in approximately 5% of women in our study who reported never having had sex. In an unweighted analysis of these women, we found that most were young in age and some had other sexually transmitted infections, suggesting their self-reported sexual history may not be accurate. Genital HPV is primarily associated with sexual intercourse; however, 1 study³ found that nonpenetrative sexual contact, such as genital-genital contact, could also result in HPV transmission. A detailed sexual history was not collected, therefore, we could not evaluate specific types of sexual contact.

The 84.9% response rate (number of collected swabs/number of eligible swabs=2026/2387) for vaginal swab collection in our study suggests that the self-collected swabs were a feasible and effective method for HPV DNA detec-

tion in a large survey. Previous assessments of the acceptability of self-collected vaginal swabs found this collection method to be more acceptable than specimens collected by health care professionals.^{29,30} Evaluations have also demonstrated a high correlation of HPV DNA detection, both high-risk and low-risk types, in self-collected swabs compared with swabs collected by health care professionals.³¹⁻³⁵ There are few studies comparing correlation of specific HPV types. Available studies on type-specific correlations suggest different types may be detected from these specimens, although a strong correlation, at least in 1 study, existed for HPV types 6, 11, 16, and 18.³⁵ There are no evaluations in the general US population using self-collected swabs precluding a direct comparison of our data with those from other studies.

There were several limitations to our study. Nonresponders were significantly different from responders by certain demographic characteristics (race, age, country of birth, and ever had sex) and this could bias prevalence estimates; however, as noted, evaluations of the weighted prevalence estimates taking the nonresponders into consideration did not substantially change our point estimates. A self-collected cervicovaginal sample was obtained, which may not detect the same HPV types as cervical mucosa samples obtained by health care professionals. Also, HPV DNA point prevalence will most certainly underestimate cumulative incidence as many infections clear^{9,10}; this assessment only measures current infection and does not indicate past exposure to HPV. A study in a previous NHANES sample found overall HPV-16 seroprevalence in females to be 17.9%, with the prevalence peaking in 20- to 29-year-olds.³⁶ Seroprevalence provides a better estimate of cumulative exposure. The difference between the seroprevalence found in that study and the current DNA prevalence in our study reflects the high clearance of HPV infections. This assessment, as all

assessments of HPV DNA, could not determine if the HPV detected was from the participant or a partner, or if it represented active infection. Finally, we had only 2 years of data; the distribution of HPV types may change with additional years of data and a larger sample. We did not present demographic and behavioral factors associated with HPV vaccine type infection because the analysis was limited to subgroups with prevalence estimates in which the relative SE was 30% or less; a relative SE of more than 30% means that the SE is quite large relative to the estimate and hence considered unreliable.

CONCLUSION

Our study provides the first national estimate of prevalent HPV infection among females aged 14 to 59 years in the United States. Overall, HPV prevalence was high (26.8%), and prevalence was highest among females aged 20 to 24 years. Our data indicate that the burden of prevalent HPV infection among women was higher than previous estimates. However, the prevalence of HPV vaccine types was relatively low.

Author Contributions: Dr Dunne had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Unger, McQuillan, Markowitz.

Acquisition of data: Unger, McQuillan, Swan, Patel, Markowitz.

Analysis and interpretation of data: Dunne, Unger, Sternberg, McQuillan, Patel, Markowitz.

Drafting of the manuscript: Dunne, Patel, Markowitz.
Critical revision of the manuscript for important intellectual content: Dunne, Unger, Sternberg, McQuillan, Swan.

Statistical analysis: Sternberg, McQuillan.

Administrative, technical, or material support: Unger, McQuillan, Swan.

Study supervision: Unger, Markowitz.

Financial Disclosures: None reported.

Funding/Support: This study was supported by the Division of STD Prevention, Centers for Disease Control and Prevention.

Role of the Sponsor: The funding organization, National Center for Health Statistics, Centers for Disease Control and Prevention, assisted with the conduct of the study, in the collection and management of the data, and in the preparation and review of the manuscript.

REFERENCES

- Weinstock H, Berman S, Cates W. Sexually transmitted infections in American youth: incidence and prevalence estimates, 2000. *Perspect Sex Reprod Health.* 2004;36:6-10.
- Revzina NV, Diclemente RJ. Prevalence and incidence of human papillomavirus infection in women in the USA: a systematic review. *Int J STD AIDS.* 2005;16:528-537.
- Tarkowski TA, Koumans EH, Sawyer M, et al. Epidemiology of human papillomavirus infection and abnormal cytologic test results in a cohort of female adolescent population. *J Infect Dis.* 2004;189:46-50.
- Winer RL, Lee SK, Hughes JP, Adam DE, Kiviat NB, Koutsky LA. Genital human papillomavirus: infection incidence and risk factors in a cohort of female university students. *Am J Epidemiol.* 2003;157:218-226.
- Trottier H, Franco EL. The epidemiology of genital human papillomavirus infection. *Vaccine.* 2006;24(suppl 1):S1-S15.
- Manhart LE, Holmes KK, Koutsky LA, et al. Human papillomavirus infection among sexually active young women in the United States: implications for developing a vaccination strategy. *Sex Transm Dis.* 2006;33:502-508.
- Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol.* 1999;189:12-19.
- Bosch FX, de Sanjose S. Chapter 1: human papillomavirus and cervical cancer-burden and assessment of causality. *J Natl Cancer Inst Monogr.* 2003; (31):3-13.
- Moscicki AB, Shiboski S, Broering J, et al. The natural history of human papillomavirus infection as measured by repeated DNA testing in adolescent and young women. *J Pediatr.* 1998;132:277-284.
- Franco EL, Villa LL, Sobrinho JP, et al. Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer. *J Infect Dis.* 1999;180:1415-1423.
- Food and Drug Administration. Gardasil (quadrivalent human papillomavirus [types 6, 11, 16, 18] recombinant vaccine) [package insert]. Whitehouse Station, NJ: Merck & Co; 2006. <http://www.fda.gov/cber/label/HPVmer060806LB.pdf>. Accessed February 8, 2007.
- Koutsky LA, Ault KA, Wheeler CM, et al; Proof of Principle Study Investigators. A controlled trial of a human papillomavirus type 16 vaccine. *N Engl J Med.* 2002;347:1645-1651.
- Villa LL, Costa RL, Petta CA, et al. Prophylactic quadrivalent human papillomavirus (types 6, 11, 16, and 18) L1 virus-like particle vaccine in young women: a randomised double-blind placebo-controlled multicentre phase II efficacy trial. *Lancet Oncol.* 2005;6:271-278.
- US Centers for Disease Control and Prevention. ACIP provisional recommendations for the use of quadrivalent HPV vaccine. http://www.cdc.gov/nip/recs/provisional_rec/hpv.pdf. Accessed February 8, 2007.
- National Center for Health Statistics. NHANES 2003-2004. <http://www.cdc.gov/nchs/about/major/nhanes/nhanes>. Accessed February 23, 2006.
- US Census Bureau. Ratio of income to poverty level. <http://www.census.gov/hhes/income/defs/ratio.html>. Accessed April 27, 2006.
- US Census Bureau Web site. National estimates: monthly population estimates, 2000 to 2004. http://www.census.gov/popest/national/asrh/2004_nat_detail.html. Accessibility verified January 25, 2007.
- Gravitt PE, Peyton CL, Apple RJ, Wheeler CM. Genotyping of 27 human papillomavirus types by using L1 consensus PCR products by a single-hybridization, reverse line blot detection method. *J Clin Microbiol.* 1998;36:3020-3027.
- Gravitt PE, Peyton CL, Alessi TQ, et al. Improved amplification of genital human papillomaviruses. *J Clin Microbiol.* 2000;38:357-361.
- Peyton CL, Gravitt PE, Hunt WC, et al. Determinants of genital human papillomavirus detection in a US population. *J Infect Dis.* 2001;183:1554-1564.
- Vernon SD, Unger ER, Williams D. Comparison of human papillomavirus detection and typing by cycle sequencing, line blotting, and hybrid capture. *J Clin Microbiol.* 2000;38:651-655.
- SAS Institute Inc. *SAS Companion for the Microsoft Window Environment, Version 9*. Cary, NC: SAS Institute Inc; 2002.
- Shah BV, Barnwell BG, Bieler GS. *SUDAAN User's Manual: Release 9.0*. Research Triangle Park, NC: Research Triangle Institute; 2004.
- Korn E, Graubard B. *Analysis of Health Surveys*. New York, NY: Wiley; 1999.
- National Center for Health Statistics. Design and estimation for the National Health Interview Survey, 1995-2004. *Vital Health Stat 2.* 2000;130:1-31.
- Casella G, Berger R. *Statistical Inference*. Pacific Grove, Calif: Wadsworth & Brooks/Col; 1990.
- Xu F, Sternberg MR, Kottiri BJ, et al. Trends in herpes simplex virus type 1 and type 2 seroprevalence in the United States. *JAMA.* 2006;296:964-973.
- Castle PE, Schiffman M, Bratti MC, et al. A population-based study of vaginal human papillomavirus infection in hysterectomized women. *J Infect Dis.* 2004;190:458-467.
- Dzuba IG, Diaz EY, Allen B, et al. The acceptability of self-collected samples for HPV testing vs. the pap test as alternatives in cervical cancer screening. *J Womens Health Gen Based Med.* 2002;11:265-275.
- Kahn JA, Bernstein DI, Rosenthal SL, et al. Acceptability of human papillomavirus self testing in female adolescents. *Sex Transm Infect.* 2005;81:408-414.
- Sellors JW, Lorincz AT, Mahony JB, et al. Comparison of self-collected vaginal, vulvar and urine samples with physician-collected cervical samples for human papillomavirus testing to detect high-grade squamous intraepithelial lesions. *CMAJ.* 2000;163:513-518.
- Gravitt PE, Lacey JV Jr, Brinton LA, et al. Evaluation of self-collected cervicovaginal cell samples for human papillomavirus testing by polymerase chain reaction. *Cancer Epidemiol Biomarkers Prev.* 2001;10:95-100.
- Ogilvie GS, Patrick DM, Schulzer M, et al. Diagnostic accuracy of self collected vaginal specimens for human papillomavirus compared to clinician collected human papillomavirus specimens: a meta-analysis. *Sex Transm Infect.* 2005;81:207-212.
- Lorenzato FR, Singer A, Ho L, et al. Human papillomavirus detection for cervical cancer prevention with polymerase chain reaction in self-collected samples. *Am J Obstet Gynecol.* 2002;186:962-968.
- Baldwin S, Santos C, Mendez Brown E, et al. Comparison of type-specific human papillomavirus data from self and clinician directed sampling. *Gynecol Oncol.* 2005;97:612-617.
- Stone KM, Karem KL, Sternberg MR, et al. Seroprevalence of human papillomavirus type 16 infection in the United States. *J Infect Dis.* 2002;186:1396-1402.