



Published in final edited form as:

*Int J Cancer*. 2020 June 15; 146(12): 3320–3328. doi:10.1002/ijc.32693.

## Epidemiological evidence that common HPV types may be common because of their ability to evade immune surveillance: Results from the Women’s Interagency HIV study

Philip E. Castle<sup>1</sup>, Robert D. Burk<sup>1</sup>, Leslie S. Massad<sup>2</sup>, Isam-Eldin Eltoum<sup>3</sup>, Charles B. Hall<sup>1</sup>, Nancy A. Hessol<sup>4</sup>, Kathryn Anastos<sup>1</sup>, Xianhong Xie<sup>1</sup>, Howard Minkoff<sup>5</sup>, Xiaonan Xue<sup>1</sup>, Gypsyamber D’Souza<sup>6</sup>, Lisa Flowers<sup>7</sup>, Christine Colie<sup>8</sup>, Lisa Rahangdale<sup>9</sup>, Margaret A. Fischl<sup>10</sup>, Joel M. Palefsky<sup>4</sup>, Howard D. Strickler<sup>1</sup>

<sup>1</sup>Albert Einstein College of Medicine, Bronx, NY

<sup>2</sup>Washington University School of Medicine, St. Louis, MO

<sup>3</sup>University of Alabama, Birmingham, AB

<sup>4</sup>University of California, San Francisco, CA

<sup>5</sup>Maimonides Medical Center, Brooklyn, NY

<sup>6</sup>Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

<sup>7</sup>Emory University School of Medicine, Atlanta, GA

<sup>8</sup>Georgetown University Medical Center, Washington, DC

<sup>9</sup>University of North Carolina School of Medicine, Chapel Hill, NC

<sup>10</sup>University of Miami Miller School of Medicine, Miami, FL

### Abstract

Infection by human papillomavirus (HPV) type 16, the most oncogenic HPV type, was found to be the least affected by HIV-status and CD4 count of any of the approximately 13 oncogenic HPV types. This relative independence from host immune status has been interpreted as evidence that HPV16 may have an innate ability to avoid the effects of immunosurveillance. However, the impact of immune status on other individual HPV types has not been carefully assessed. We studied type-specific HPV infection in a cohort of 2,470 HIV-positive (HIV[+]) and 895 HIV-negative (HIV[-]) women. Semi-annually collected cervicovaginal lavages were tested for >40 HPV types. HPV type-specific prevalence ratios (PRs), incidence and clearance hazard ratios (HRs), were calculated by contrasting HPV types detected in HIV[+] women with CD4 < 200 to

**Correspondence to:** Howard D. Strickler, howard.strickler@einstein.yu.edu.

#### Disclosures

PEC has received commercial HPV tests for research at a reduced or no cost from Roche, Cepheid, BD and Arbor Vita Corporation. JMP has been compensated financially as a member of a Merck data and safety monitoring board for HPV vaccines; as a consultant for Vir Biotechnologies received fees and stock options; Antiva Biosences provided grant support; with Ubiome participated in a scientific advisory board and received stock options; with Vaccitech participated in a scientific advisory board and received stock options; with Janssen Pharmaceuticals was an invited speaker. KA has received grant funding from several institutes at NIH. HDS has received free-blinded testing using HPV E6/E7 protein assays by Arbor Vista, p16/Ki67 cytology by MTM Laboratories/Ventura-Roche, and MCM-2/TOP2A cytology by BD Diagnostics.

HIV[−] women. HPV71 and HPV16 prevalence had the weakest associations with HIV-status/CD4 count of any HPV, according to PRs. No correlations between PRs and HPV phylogeny or oncogenicity were observed. Instead, higher HPV type-specific prevalence in HIV[−] women correlated with lower PRs ( $\rho = -0.59$ ;  $p = 0.0001$ ). An alternative (quadratic model) statistical approach ( $P_{\text{HIV}+} = a * P_{\text{HIV}-} + b * P_{\text{HIV}-}^2$ ;  $R^2 = 0.894$ ) found similar associations ( $p = 0.0005$ ). In summary, the most prevalent HPV types in HIV[−] women were the types most independent from host immune status. These results suggest that common HPV types in HIV[−] women may have a greater ability to avoid immune surveillance than other types, which may help explain why they are common.

## Introduction

Persistent cervical infection by an oncogenic human papillomavirus (HPV) genotype (type) is the primary cause of cervical cancer and its immediate precursor lesions.<sup>1</sup> HIV-positive (HIV[+]) women have significantly elevated the risk of cervical cancer/precancer<sup>2</sup> as well as infection by HPV, relative to the general population, and these risks increase with diminishing CD4+ T-cell (CD4) count. Quantifiable decrements in host immune status reflected by CD4 count in HIV[+] women makes the study of cervical disease and HPV in this population highly informative regarding the relationship between immunity and the natural history of HPV infection.

Prior data from our group and others have provided evidence-based on HIV[+] cohorts that there may be type-specific differences in the effects of host immune status on the natural history of HPV infection. For example, the prevalence of HPV16, which causes approximately 55–60% of all cervical cancers in the general population, was found to be the least affected by HIV status and CD4 count of any oncogenic HPV, based on prevalence ratios (PR) in multiplicative multivariate models.<sup>3</sup> This relative independence from host immune status has been interpreted as evidence that HPV16 may have an innate ability to avoid the effects of immunosurveillance, which could partly explain the high prevalence of HPV16 in the general population and its unique carcinogenicity.<sup>4–6</sup>

More recently, the clinical and biological significance of these results were demonstrated by a study showing that the prevalence of HPV16 was significantly lower in cervical precancer diagnosed in HIV[+] than HIV-negative (HIV[−]) women.<sup>4</sup> Subsequent meta-analyses reported similar results, finding that HIV[+] women had reduced HPV16 in cervical precancer (using data from around the world)<sup>5</sup> and in invasive cervical cancer (using data from Africa),<sup>6</sup> compared to HIV[−] women. Additional studies found lower HPV16 prevalence in invasive anal cancers among HIV[+] relative to HIV[−] women and men.<sup>7</sup>

We therefore hypothesized that there is variation in the effects of immune status on HPV infection by type, reflecting differences in the innate capacity of HPVs to avoid immunosurveillance. However, the impact of immune status on other HPV types other than HPV16 has not been carefully assessed at the individual level. To study these possible type-specific differences the current investigation involved HPV DNA data from a much larger number of subjects and more than double the person-visits of observation than in the earlier study focused on HPV16.<sup>3</sup>

## Methods

### Population

Women's Interagency HIV Study (WIHS) is an ongoing prospective cohort of HIV[+] and at risk HIV[-] women enrolled through similar sources at six sites (Bronx, NY; Brooklyn, NY; Chicago; Los Angeles; San Francisco; Washington, DC). Enrollment was conducted between October 1994 and November 1995 ( $n = 2,054$  HIV[+],  $n = 569$  HIV[-]), and a second group was similarly enrolled between October 2001 and September 2002 ( $n = 737$  HIV[+],  $n = 406$  HIV[-]). Overall, there were 2,791 HIV[+] and 975 HIV[-] women, involving 15,220 person-years of observation (11,498 in HIV[+] and 3,722 in HIV[-] WIHS women, more than twice that in an earlier study focused on HPV16.<sup>3</sup>

Details of the WIHS data collection and recruitment methods have been previously reported.<sup>8</sup> In brief, participants underwent semiannual visits that involved a gynecologic examination with specimen collection, including a Pap test and cervicovaginal lavage (CVL) for HPV DNA testing. Written informed consent was obtained from all participants, and the study was approved by each local institutional review board.

### Laboratory testing

CVLs were tested for >40 HPV genotypes using a sensitive and specific MY09/MY11 PCR assay for HPV,<sup>9–15</sup> as previously described.<sup>8–10</sup> Primer set PC04/GH20, which amplifies a 268-base pair cellular  $\beta$ -globin DNA fragment, was included in each assay as an internal control to assess the adequacy of amplification. Consistent with International Agency for Research on Cancer recommendations, HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68, were classified as oncogenic and all other HPV types were classified as nononcogenic types.<sup>16</sup> In subanalyses, HPV genotypes were grouped by phylogeny; for example,  $\alpha$ 9 (HPV16-related) vs.  $\alpha$ 7 (HPV18-related) species.<sup>17</sup>

### Statistical analysis

**Prevalence ratios.**—Contingency tables were generated to assess baseline patient characteristics. The prevalence of specific HPV types was based on the percentage of women with adequate HPV test results (i.e., women whose CVL specimens were positive for  $\beta$ -globin amplification) within each HIV/CD4 count stratum by clinical visit. The cutoffs for each HIV/CD4 stratum, a measure of host immune status, were determined based on prior studies of HPV DNA data in WIHS,<sup>3,11,13</sup> namely (i) HIV[-], and HIV[+] with (ii) CD4 > 500, (iii) CD4 200–500 and (iv) CD4 < 200 T-cells/mm<sup>3</sup>. Analyses of these data were conducted using multivariate Poisson regression models that incorporated generalized estimating equations (GEE) as previously reported.<sup>18</sup> These models estimated summary PRs for each HPV type by contrasting HIV[+] women with CD4 < 200 to those with CD4 > 500 and to HIV[-] women, while accounting for both repeated observations of the same women over time and the possibility of multiple concurrent infections by different HPV types. HPV type-specific prevalence itself was similarly estimated using GEE models by individual HIV and CD4 strata, adjusted for age. The presence of a biologic gradient of HPV prevalence across the four HIV/CD4 strata was assessed by the use of an ordinal variable with four levels in the GEE multivariate Poisson regression models (*p*-trend). While initial models

adjusted only for age, additional models adjusted for age, ethnicity, smoking, injection drug use, the lifetime number of male sexual partners at baseline, the number of male sexual partners in past 6 months, cervical treatment in past 6 months and enrollment period. All variables were time-updated except ethnicity, enrollment period and lifetime number of male sexual partners. The clinical and biological relevance of multiplicative models (and PRs in particular) for these analyses have been previously demonstrated.<sup>3–6</sup>

**Studying correlations and the possibility of a statistical tautology.**—Spearman correlation coefficients were used to characterize associations between the HPV type-specific PRs (above) and the prevalence of each HPV type in HIV[–] women. However, we also addressed the possibility that a very low type-specific HPV prevalence in HIV[–] women ( $P_{\text{HIV-}}$ ) could result in an inflated or unstable PR and, thereby, influence the observed correlation (leading to a tautology). Specifically, the possibility of a tautology was addressed by modeling  $P_{\text{HIV+}}$  on  $P_{\text{HIV-}}$  using a quadratic regression model:

$$P_{\text{HIV+}} = a * P_{\text{HIV-}} + b * P_{\text{HIV-}}^2,$$

in which a and b are constants:

Because  $P_{\text{HIV-}}$  is not in the denominator of the equation and instead PR is derived as part of the regression model and never directly calculated or modeled as an outcome variable, this quadratic model greatly reduced the possibility of inflated and unstable PR and, moreover, excluded the possibility that the findings might be due to a mathematical tautology. That is, in this quadratic model, the association between the PR and  $P_{\text{HIV-}}$  was estimated by the constant “b”, where a negative value for b would suggest that the PR has a negative association with  $P_{\text{HIV-}}$  (our predicted result), while a positive value for b would suggest the opposite. The fit of the model was then evaluated using R-squared. Both the correlation and the quadratic regression model results were reported in this paper since the correlation analysis had the advantage of being familiar to most readers and was straightforward to interpret, whereas the results of the quadratic model provided the robust findings needed to validate the relationship between PRs and prevalence in HIV[–] women (and demonstrated that the results were not due to a mathematical tautology).

**Incident detection and time to clearance.**—Multivariate Cox models were used to study the incident detection of each HPV type. We use the term “incident detection” rather than “incidence” because the relative contributions of newly acquired HPV infections and reactivation of previously acquired (e.g., latent or quiescent) HPV infections cannot be determined in a population with many years of sexual activity, regardless of the HIV-status of that population. The time of each incident event was defined as the mid-interval (i.e., the midpoint between two consecutive visits) and, as mentioned above, HIV[–] women were used as the reference group. The Wei-Lin-Weissfeld marginal model approach was used to account for the possibility of the incident detection of multiple different HPV types in these Cox models.<sup>19</sup> Participants who had missing data for two consecutive visits were censored at the time of their last visit with complete data. The two-sided Wald test was used to measure the statistical significance of the variables in the models. A similar statistical approach was

used to model HPV clearance, with clearance defined as two sequential negative results for a given HPV type. The assumption of proportionality was assessed using standard methods.<sup>20</sup>

**Additional analyses.**—We assessed the level of heterogeneity in the HPV type-specific PRs by phylogenetic group and oncogenic group using methods similar to those employed in meta-analysis. Specifically, we used the  $I^2$  statistic obtained by adding an interaction between HPV type and HIV/CD4 strata in the above GEE Poisson regression models, and then assessed the test statistics for the significance of this interaction.<sup>21</sup>

All tests were two-sided with significance set at  $p$ -value < 0.05. SAS 9.3 was used for all statistical analyses.

### Data availability

All data are shared with the Multicenter AIDS Cohort Study/Women's Interagency HIV Study (MACS/WIHS) and can be requested through the submission of a formal Concept Sheet available on the MACS/WIHS Combined Cohort Study website <https://mwccs.org/>.

### Results

Table 1 shows baseline characteristics of the study subjects, of whom 68% were enrolled in 1994–1995 and 32% in 2001–2002. A total of 2,470 (88%) of all HIV[+] and 895 (92%) of all HIV[-] WIHS women had adequate HPV and CD4 data available to be included in this analysis. The median follow-up time was nine semi-annual visits; that is, 16 visits for those enrolled in 1994–1995 and seven for those enrolled in 2001–2002. Age at enrollment varied by subcohort (36 years in 1994–1995 vs. 32 in 2001–2002 enrollees;  $p < 0.0001$ ) and by HIV-status (35 years in HIV[+] and 32 in HIV[-] women;  $p < 0.0001$ ), whereas ethnicity did not vary by either variable. Overall, 54% of subjects were Black, 27% Hispanic, 15% White and 3% other. While HIV-status was associated with several behavioral factors, such as the median number of recent sex partners (lower in HIV[+] than HIV[-] women), the differences were modest even when statistically significant (Table 1).

Among HIV[+] women, the median baseline CD4 count (CD4 = 494; interquartile range [IQR] = 337, 698) was higher ( $p < 0.001$ ) in the 2001–2002 compared to 1994–1995 subcohort (CD4 = 332; IQR = 159, 698), consistent with 48% use of highly active antiretroviral therapy (HAART) in the 2001–2002 subcohort at enrolment, but essentially none at baseline in the 1994–1995 subcohort (HAART became widespread in the WIHS during the Spring of 1996; that is, 1½–2 years after enrollment). There were similar but smaller differences in nadir CD4. At their last visit, 50% of HIV[+] women reported using HAART (47% in 1994–1995 and 58% in 2001–2002 enrollees).

Table 2 presents GEE prevalence rates and PRs adjusted for age, smoking, ever IDU, number of male sexual partner in past 6 months, and cervical treatment in past 6 months; variables that were each associated with HIV-status. These models also *a priori* included lifetime number of male sexual partner. For HIV[-] women, the average prevalences ranged from 0.04% for HPV73 and 0.05% for HPV69 to 1.28 for HPV71 and 1.29 for HPV53.

CD4 was treated as a time-dependent covariate in the multivariate models. The prevalence of virtually every HPV type among HIV[+] women significantly increased with decreasing CD4 (e.g.,  $p_{\text{trend}} = 0.001$ ). The lowest PRs contrasting HIV[+] women with CD4 < 200 to HIV[-] women, denoted henceforth as PRs<sub><200</sub>, were for HPV71 (1.78), HPV16 (3.85), HPV6 (4.78), HPV31 (4.86) and HPV69 (5.03) and the highest PRs<sub><200</sub> were for HPV85 (30.03), HPV89 (23.12), HPV84 (21.78), HPV11 (19.24) and HPV66 (18.03).

While there were no significant differences in ethnicity by HIV-status, previous studies have found that ethnicity may affect the HPV type-specific distribution.<sup>22,23</sup> Therefore, in additional PR analyses, we included ethnicity in the multivariate models. However, ethnicity was not a significant covariate ( $p > 0.4$ ). Likewise, inclusion of enrollment date had no discernible impact on any findings after control for other covariates (data not shown) and, when measured separately by enrollment period, the Spearman correlations did not meaningfully vary nor was there statistical interaction between enrollment date and type-specific HPV prevalence among HIV- women ( $p_{\text{interaction}} = 0.68$ ).

The relationship of PRs<sub><200</sub> with phylogenetic group and oncogenicity was then examined. The heterogeneity of PRs within several individual phylogenetic groups was found to be substantial (i.e.,  $\alpha 3$ ,  $\alpha 9$  and  $\alpha 10$  species) and moderate for the  $\alpha 7$  species, indicating that the individual HPV types in these groups had dissimilar PRs and their data could not be meaningfully statistically combined (Table 3). The exclusion of HPV16 from the  $\alpha 9$  species only partially reduced heterogeneity from substantial to moderate in this phylogenetic group (data not shown). Likewise, the heterogeneity of PRs<sub><200</sub> among HPV grouped as oncogenic and nononcogenic was also substantial (Table 3) and, as with phylogenetic group, exclusion of HPV16 from the oncogenic types only modestly reduced heterogeneity.

In contrast, there was a high inverse Spearman correlation ( $\rho = -0.59$ ;  $p = 0.0001$ ) found between PRs<sub><200</sub> and HPV type-specific prevalence in HIV[-] women (Fig. 1). For example, HPV16 and HPV71 both had a high prevalence in HIV[-] women and a weak association with host immune status as measured by type-specific PRs<sub><200</sub> (Table 2), whereas HPV types 11, 40, 84 and 85 had a low prevalence and strong associations with type-specific PRs<sub><200</sub>.

The impact on the results of changing the PR reference group was then examined. That is, the PRs for these secondary analyses were calculated contrasting using HIV[+] CD4 > 500 (rather than HIV[-] women) as the referent. These HPV type-specific PR<sub>200 vs. 500</sub> were then correlated with the prevalence in HIV[-] women as above. However, these changes made no discernible difference in the correlation estimate ( $\rho = -0.57$ ;  $p = 0.0001$ ).

Moreover, we assessed the association of HPV type-specific PRs in women with HIV[+]/CD4 < 200 vs. HIV[-] women using a quadratic model ( $P_{\text{HIV}+} = a^* P_{\text{HIV}-} + b^* P_{\text{HIV}-}^2$ ). As mentioned, this quadratic model greatly reduced the possibility of inflated and unstable PR and, moreover, excluded the possibility that the findings might be due to a mathematical tautology; i.e., because  $P_{\text{HIV}-}$  is not in the denominator of the equation and instead PR is derived as part of the regression model and never directly calculated or modeled as an outcome variable. Overall, the quadratic model fit these data well (as

measured by a large R-squared;  $R^2 = 0.894$ ), and the estimate of  $\mathbf{b}$  in this model was negative (as reflected in Fig. 2 by a downward curve) with high statistical significance ( $p = 0.0005$ ); that is, indicating a negative association between HPV type-specific  $PR_{<200}$  and  $P_{HIV-}$  (across all HPV types). Thus, both the correlation and quadratic model statistical approaches provided similar results.

To better understand the impact of CD4+ on these relationships, we then applied this quadratic model now using HIV[+] women with (i) CD4 200–500 and (ii) >500 (instead of those with CD4 < 200) as the comparators to HIV[-] women. These two analyses found a significant negative quadratic relationship ( $p = 0.02$  and  $p = 0.04$ , respectively) and, as expected, the coefficient significantly decreased in magnitude with increasing CD4 count ( $p_{\text{trend}} = 0.001$ ;  $R^2 = 0.90$  for the quadratic model across HIV/CD4 strata). Unlike CD4, neither phylogenetic nor oncogenic grouping was statistically significant in these models ( $p > 0.22$ ); that is, using interaction terms to obtain and summarize results across the individual CD4 strata, further demonstrating the consistency of the quadratic equation and GEE PR analysis results.

Finally, we explored several aspects of type-specific HPV natural history. Specifically, we examined the relationships of  $PR_{<200}$  with the incidence hazard ratio ( $I_{HR}$ ), clearance hazard ratio ( $C_{HR}$ ) and  $I_{HR}/C_{HR}$  ratio. Because the reciprocal of clearance is duration of infection, these analyses essentially address the relationship  $PR_{<200} = \text{Incidence } HR_{<200} \times \text{Duration } HR_{<200}$ . As predicted, the  $PR_{<200}$  and  $I_{HR} < 200 / C_{HR} < 200$  were strongly correlated ( $\rho = 0.66$ ,  $p < 0.0001$ ). Moreover, we also assessed which aspect of HPV natural history might be the major driving force behind the relationship between HPV type-specific prevalence and host immune status. That is, we assessed the correlations of  $PR_{<200}$  with  $I_{HR} < 200$  and then with  $C_{HR} < 200$ , and found that the relation of  $PR_{<200}$  with  $I_{HR} < 200$  ( $\rho = 0.75$ ,  $p < 0.0001$ ) was much stronger than with  $C_{HR} < 200$  ( $\rho = -0.35$ ,  $p = 0.04$ ).

## Discussion

This article presents novel findings regarding the natural history of individual HPV types and their associations with host immune status (as measured by HIV-status and CD4 count). In particular, the data showed that the HPV types most prevalent in HIV[-] women were also the types whose prevalence was least affected by HIV infection and CD4 count. This relative independence from host immune status is consistent with prior results for HPV16, the most common and carcinogenic HPV type, and has been interpreted as possible evidence of an innate ability to avoid the effects of immunosurveillance.<sup>3–7</sup> If correct, it could help explain why HPV16, and other common HPV types, are in fact common.

The current data also showed that HPV type-specific PRs related to HIV and CD4 are more strongly correlated with  $I_{HR}$  than  $C_{HR}$ . That is, the relative HPV type-specific prevalence was on average more greatly associated with the impact of HIV/CD4 on incident detection than the duration of HPV infection. We posit that this may partly reflect HPV reactivation, as prior studies in WIHS and other cohorts have reported evidence of HPV reactivation and its strong relation with HIV-status and low CD4 count (e.g., based on longitudinal studies with serial HPV testing in long celibate women).<sup>11</sup>

In addition, we addressed certain statistical limitations related to the use of correlations and linear models in the analysis of these data. Specifically, it was possible that a small  $P_{HIV-}$  could lead to an inflated or unstable PR and affect the findings. We therefore reanalyzed the data using a quadratic equation model that did not raise this concern; that is, this approach not only reduced the possibility of an inflated and unstable PRs as a result of a small  $P_{HIV-}$  estimate, it also excluded the possibility that this might influence the correlation between these two variables (a tautology), since  $P_{HIV-}$  was not a denominator in the equation and PR was derived as part of the regression model and was never directly calculated or modeled as an outcome variable.

Based on the collective findings, we therefore propose that the relative independence from host immune status of common HPV types may be an important (albeit, previously unrecognized) characteristic of HPV types with high prevalence in HIV[-] women. Future studies should investigate the genetic factors—distinct from phylogenetic species designation (which was unrelated to HPV PR)—and other biological correlates related to high HPV prevalence and independence from host immune status. Such studies could help elucidate biologic pathways that play a role in HPV's ability to avoid the effects of immune surveillance (immunoavoidance), and provide targets for the development of novel strategies to reduce HPV infection.

In addition to the clinically relevant role immunoavoidance may play in HPV16 oncogenicity, it is also possible that immunoavoidance may play an important role in the high prevalence and predominant role of HPV6, a nononcogenic HPV, in genital warts.<sup>24</sup> HPV6 was among the types of HPV other than HPV16 that had a weak association with host immune status as measured by HIV-status and CD4 count. Thus, there is potential clinical importance to the current findings related to both oncogenic and nononcogenic HPV types.

We note several limitations to our study that should be considered in the interpretation of the findings. First, the current HPV results (based on high-risk HIV[+] and HIV[-] women) may not be generalizable to the overall US population. Nonetheless, the existence of a large HIV[-] comparison group in the WIHS, enrolled through similar clinical and out-reach sources as the HIV[+] women, is a strength of the WIHS and helps make some extrapolation to HIV[-] women possible. In fact, the current data show a strong biologic gradient of increasing HPV infection from HIV[-] to HIV infection with diminishing CD4 count, suggesting an ordinal relationship if not a continuum. We also found similar type-specific HPV findings using HIV[+] women with CD4 > 500 as the reference group. Second, for certain phylogenetic groups we detected too few HPV types to adequately characterize them as a group. This was, for example, true with  $\alpha 15$ , for which we detected only HPV71; albeit, HPV71 was common in our population. Other members of this phylogenetic group, for example, HPV90<sup>25</sup> and HPV106,<sup>26</sup> were rare or, in the case of HPV106, were only recently discovered and therefore not previously tested for by our assay.<sup>25,26</sup> It is additionally possible that differences in sensitivity of our degenerate primer HPV PCR assay for certain HPV types could have affected the results. However, MY09/MY11 PCR is well established and shown to have broad sensitivity and high specificity,<sup>9</sup> and the distribution of HPV types in HIV[-] women in our study was similar to the type-specific distribution found in one of the few truly population-based studies of a general population, the well known NCI/NIH



Guanacaste Project.<sup>27,28</sup> Furthermore, variation in assay sensitivity would not explain the PRs<sub><200</sub> with HPV prevalence in HIV[-] women; for example, the associations we detected did not correlate with phylogenetic group, and random variation would be a bias toward the null. While the current investigation cannot directly differentiate between “incident detection” due to new infections, reinfection or reactivation, this is a limitation that applies equally to any study of women with many years of sexual activity, regardless of HIV-status.

Our study also had several strengths. WIHS is a large cohort with semi-annual gynecologic examinations during which CVL specimens are collected and tested in a masked fashion at two expert HPV laboratories. The WIHS collects detailed clinical data including CD4 counts at each visit, and the large size and long follow-up of the cohort allowed us to examine the natural history of individual HPV genotypes, stratified by HIV and host immune status.

In conclusion, our study provides evidence that common HPV types may be common in part because of an innate ability to avoid the effects of host immune surveillance. Furthermore, relative independence from host immune status may define a previously unrecognized aspect of HPV biology, which can now be studied to understand the factors that allow common HPV of different phylogeny and oncogenicity to avoid the effects of immune surveillance, including clinically important HPV types, HPV16 and HPV6.

## Acknowledgements

HPV testing and the current analyses were supported by the National Cancer Institute (NCI) (grant numbers R01-CA-230331, R01-CA-085178 and R01-CA-174634 to H. D. S. and P30-CA-013330, P30-AI24414). Data in this manuscript were collected by the Women’s Interagency HIV Study (WIHS). The contents of this publication are solely the responsibility of the authors and do not represent the official views of the National Institutes of Health (NIH). WIHS (Principal Investigators): UAB-MS WIHS (Michael Saag, Mirjam-Colette Kempf and Deborah Konkle-Parker), U01-AI-103401; Atlanta WIHS (Ighovwerha Oforokun and Gina Wingood), U01-AI-103408; Bronx WIHS (Kathryn Anastos), U01-AI-035004; Brooklyn WIHS (Howard Minkoff and Deborah Gustafson), U01-AI-031834; Chicago WIHS (Mardge Cohen and Audrey French), U01-AI-034993; Metropolitan Washington WIHS (Seble Kassaye), U01-AI-034994; Miami WIHS (Margaret Fischl and Lisa Metsch), U01-AI-103397; UNC WIHS (Adaora Adimora), U01-AI-103390; Connie Wofsy Women’s HIV Study, Northern California (Ruth Greenblatt, Bradley Aouizerat and Phyllis Tien), U01-AI-034989; WIHS Data Management and Analysis Center (Stephen Gange and Elizabeth Golub), U01-AI-042590; Southern California WIHS (Joel Milam), U01-HD-032632 (WIHS I-WIHS IV). The WIHS is funded primarily by the National Institute of Allergy and Infectious Diseases (NIAID), with additional cofunding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), the National Cancer Institute (NCI), the National Institute on Drug Abuse (NIDA) and the National Institute of Mental Health (NIMH). Targeted supplemental funding for specific projects is also provided by the National Institute of Dental and Craniofacial Research (NIDCR), the National Institute on Alcohol Abuse and Alcoholism (NIAAA), the National Institute on Deafness and other Communication Disorders (NIDCD) and the NIH Office of Research on Women’s Health. WIHS data collection is also supported by UL1-TR000004 (UCSF CTSA) and UL1-TR000454 (Atlanta CTSA).

## Abbreviations:

<b>C<sub>HR</sub></b>	clearance hazard ratio
<b>CVL</b>	cervicovaginal lavage
<b>GEE</b>	generalized estimating equation
<b>HPV</b>	human papillomavirus
<b>I<sub>HR</sub></b>	incidence hazard ratio

<b>P<sub>HIV-</sub></b>	prevalence in HIV-negative women
<b>P<sub>HIV+</sub></b>	prevalence in HIV-positive women
<b>PR</b>	prevalence ratio
<b>WIHS</b>	Women's Interagency HIV Study

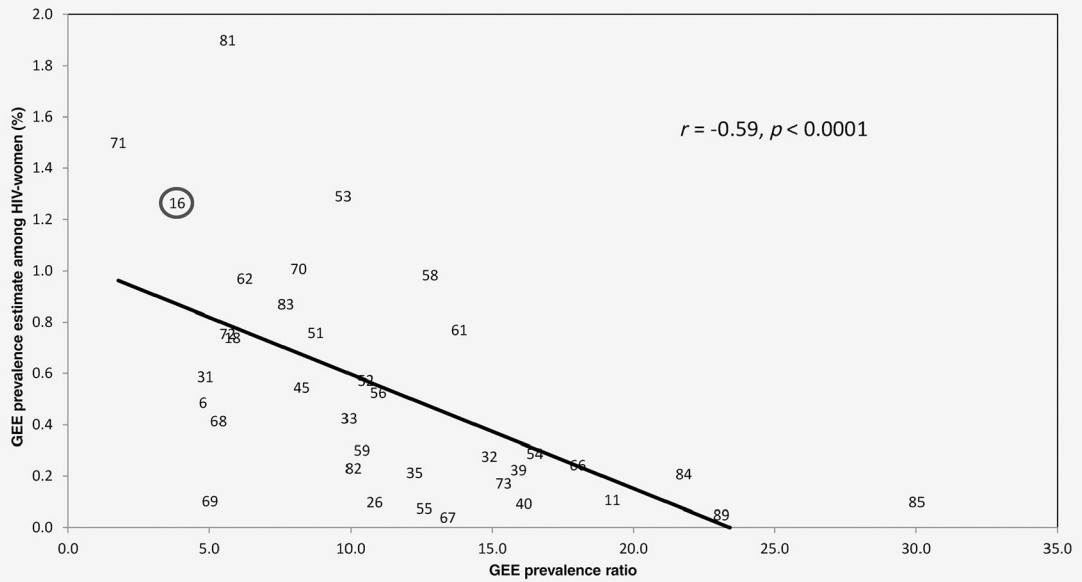
## References

- Schiffman M, Castle PE, Jeronimo J, et al. Human papillomavirus and cervical cancer. *Lancet* 2007; 370:890–907. [PubMed: 17826171]
- Grulich AE, van Leeuwen MT, Falster MO, et al. Incidence of cancers in people with HIV/AIDS compared with immunosuppressed transplant recipients: a meta-analysis. *Lancet* 2007;370: 59–67. [PubMed: 17617273]
- Strickler HD, Palefsky JM, Shah KV, et al. Human papillomavirus type 16 and immune status in human immunodeficiency virus-seropositive women. *J Natl Cancer Inst* 2003;95: 1062–71. [PubMed: 12865452]
- Massad LS, Xie X, Burk RD, et al. Association of cervical precancer with human papillomavirus types other than 16 among HIV co-infected women. *Am J Obstet Gynecol* 2016;214:354–6. [PubMed: 26433170]
- Clifford GM, Goncalves MA, Franceschi S. Human papillomavirus types among women infected with HIV: a meta-analysis. *AIDS* 2006;20:2337–44. [PubMed: 17117020]
- Clifford GM, De VH, Tenet V, et al. Effect of HIV infection on human papillomavirus types causing invasive cervical cancer in Africa. *J Acquir Immune Defic Syndr* 2016;73:332–9. [PubMed: 27331659]
- Lin C, Franceschi S, Clifford GM. HPV types from infection to cancer in the anus, according to sex and HIV status: a systematic review and meta-analysis. *Lancet Infect Dis* 2018;18:198–206. [PubMed: 29158102]
- Bacon MC, von Wyl V, Alden C, et al. The Women's interagency HIV study: an observational cohort brings clinical sciences to the bench. *Clin Diagn Lab Immunol* 2005;12:1013–9. [PubMed: 16148165]
- Castle PE, Schiffman M, Gravitt PE, et al. Comparisons of HPV DNA detection by MY09/11 PCR methods. *J Med Virol* 2002;68:417–23. [PubMed: 12226831]
- Keller MJ, Burk RD, Massad LS, et al. Cervical precancer risk in HIV-infected women who test positive for oncogenic human papillomavirus despite a Normal pap test. *Clin Infect Dis* 2015;61:1573–81. [PubMed: 26187020]
- Strickler HD, Burk RD, Fazzari M, et al. Natural history and possible reactivation of human papillomavirus in human immunodeficiency virus-positive women. *J Natl Cancer Inst* 2005;97: 577–86. [PubMed: 15840880]
- Burk RD, Ho GY, Beardsley L, et al. Sexual behavior and partner characteristics are the predominant risk factors for genital human papillomavirus infection in young women. *J Infect Dis* 1996;174:679–89. [PubMed: 8843203]
- Palefsky JM, Minkoff H, Kalish LA, et al. Cervicovaginal human papillomavirus infection in human immunodeficiency virus-1 (HIV)-positive and high-risk HIV-negative women. *J Natl Cancer Inst* 1999;91:226–36. [PubMed: 10037100]
- Qu W, Jiang G, Cruz Y, et al. PCR detection of human papillomavirus: comparison between MY09/MY11 and GP5+/GP6+ primer systems. *J Clin Microbiol* 1997;35:1304–10. [PubMed: 9163434]
- Jiang G, Qu W, Ruan H, et al. Elimination of false-positive signals in enhanced chemiluminescence (ECL) detection of amplified HPV DNA from clinical samples. *Biotechniques* 1995;19:566–8. [PubMed: 8777045]
- IARC. IARC Monographs on the Evaluation of Oncogenic Risks to Humans Human Papillomaviruses [100B]. Lyon, France: IARC, 2012.

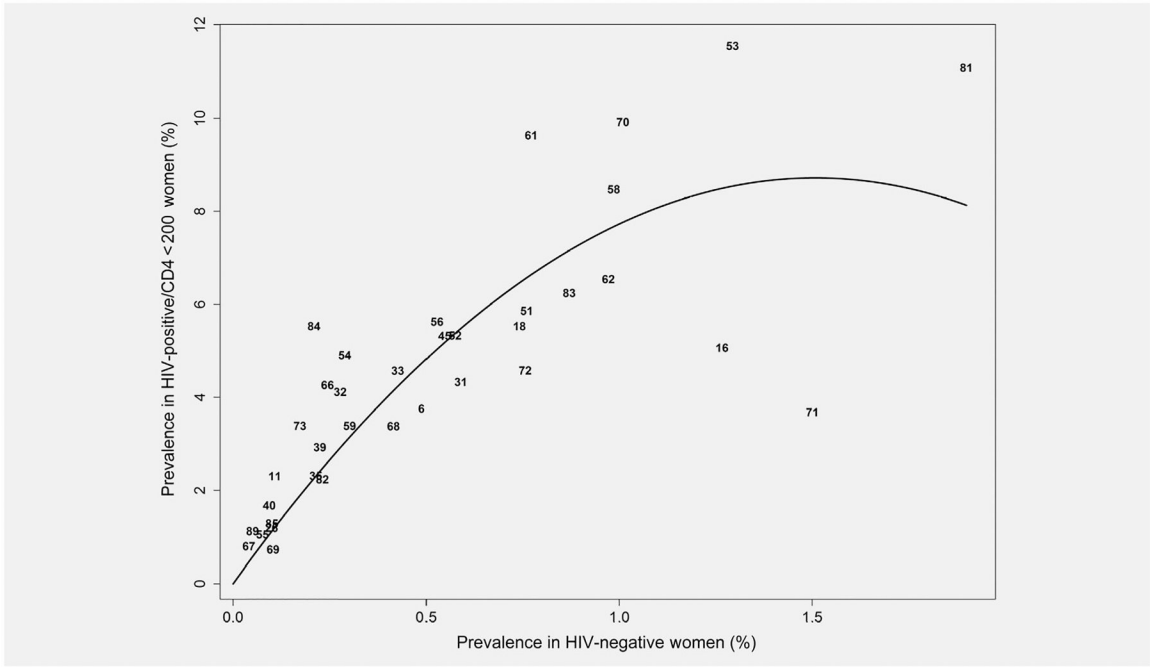
17. Bernard HU, Burk RD, Chen Z, et al. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 2010;401:70–9. [PubMed: 20206957]
18. Xue X, Gange SJ, Zhong Y, et al. Marginal and mixed-effects models in the analysis of human papillomavirus natural history data. *Cancer Epidemiol Biomarkers Prev* 2010;19:159–69. [PubMed: 20056635]
19. Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modelling marginal distributions. *J Am Stat Assoc* 1989;84:1065–73.
20. Altman DG, De Stavola BL. Practical problems in fitting a proportional hazards model to data with updated measurements of the covariates. *Stat Med* 1994;13:301–41. [PubMed: 8177984]
21. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327:557–60. [PubMed: 12958120]
22. Keller MJ, Burk RD, Massad LS, et al. Racial differences in human papilloma virus types amongst United States women with HIV and cervical precancer. *AIDS* 2018;32:2821–6. [PubMed: 30234608]
23. Niccolai LM, Russ C, Julian PJ, et al. Individual and geographic disparities in human papillomavirus types 16/18 in high-grade cervical lesions: associations with race, ethnicity, and poverty. *Cancer* 2013;119:3052–8. [PubMed: 23661284]
24. Hernandez-Suarez G, Pineros M, Vargas JC, et al. Human papillomavirus genotypes in genital warts in Latin America: a cross-sectional study in Bogota. *Colombia Int J STD AIDS* 2013;24: 567–72. [PubMed: 23970773]
25. Terai M, Burk RD. Identification and characterization of 3 novel genital human papillomaviruses by overlapping polymerase chain reaction: candHPV89, candHPV90, and candHPV91. *J Infect Dis* 2002;185:1794–7. [PubMed: 12085327]
26. Chen Z, Schiffman M, Herrero R, et al. Identification and characterization of two novel human papillomaviruses (HPVs) by overlapping PCR: HPV102 and HPV106. *J Gen Virol* 2007; 88(Pt 11):2952–5. [PubMed: 17947516]
27. Castle PE, Schiffman M, Herrero R, et al. A prospective study of age trends in cervical human papillomavirus acquisition and persistence in Guanacaste, Costa Rica. *J Infect Dis* 2005;191: 1808–16. [PubMed: 15871112]
28. Herrero R, Castle PE, Schiffman M, et al. Epidemiologic profile of type-specific human papillomavirus infection and cervical neoplasia in Guanacaste, Costa Rica. *J Infect Dis* 2005;191: 1796–807. [PubMed: 15871111]

**What's new?**

What makes common types of human papillomavirus (HPV) common? To answer this question, the authors examined HPV types in HIV-infected and uninfected women as the immune status in this population is quantifiable by CD4 T cell count. They found that HPV types that were most prevalent in HIV-negative women were also the types whose prevalence was least affected by HIV infection. This relative independence from immune status supports a model that the innate ability to avoid the effects of immune surveillance makes common HPV types, in fact, common.



**Figure 1.** Scatter plot and linear association between [X-axis] HPV type-specific GEE summary prevalence ratios (PRs) of HIV(+) and CD4 <200 related to HIV(-) women and the [Y-axis] average GEE prevalence of each individual HPV type across all study visits in HIV(-) women. The linear correlation (indicated by the line) and the Spearman correlation ( $r$ ) are both shown. Each number in the figure specifies the HPV type related to that data point on the graph. HPV16 is circled as a reference.



**Figure 2.** Quadratic model of GEE HPV type-specific prevalence for HIV[-] women ( $P_{HIV-}$ ; X-axis) vs. GEE HPV type-specific prevalence for HIV[+] women with CD4 < 200 ( $P_{HIV+}$ ; Y-axis) across all HPV types. A quadratic model of  $P_{HIV+} = a * P_{HIV-} + b * P_{HIV-}^2$  fit the data well as indicated by  $R^2 = 0.894$ . Compared to simple correlation or a linear model (e.g.,  $P_{HIV+} / P_{HIV-} = c + d * P_{HIV-}$ ) in Figure 1 this quadratic model greatly reduces the possibility of an inflated and unstable PR as a result of a small  $P_{HIV-}$  estimate and excludes the possibility that this might influence their observed correlation (a tautology). That is,  $P_{HIV-}$  is not a denominator in the quadratic equation and instead PR is derived as part of the regression model and is never directly calculated or modeled as an outcome variable. Overall, the estimate of b in this model was negative (as reflected by a downward curve) with high statistical significance ( $p < 0.0001$ ), indicating a negative association between HPV type-specific PR and  $P_{HIV-}$  (across all HPV types).

**Table 1.**

Selected baseline characteristics

Characteristic	1994–1995		2001–2002		p-value
	HIV-negative (n = 514)	HIV-positive (n = 1859)	HIV-negative (n = 387)	HIV-positive (n = 707)	
Age, Median (IQR)	34 years (28–40)	36 years (31–41)	29 years (23–35)	33 years (28–38)	<0.0001
Ethnic					0.04
Black	52%	54%	55%	56%	
Hispanic	29%	25%	29%	32%	
White	15%	19%	12%	7%	
Other	4%	3%	4%	5%	
Median number of sex partners, Lifetime (IQR)	10 (5–30)	11 (5–50)	10 (5–30)	7 (4–20)	<0.0001
Median number of sex partners, Last 6 Mo (IQR)	1 (1–1)	1 (0–1)	1 (1–3)	1 (1–1)	<0.0001
Ever used condom	53%	62%	67%	68%	0.79
Ever smoked cigarettes	62%	55%	49%	38%	0.0004
Ever used injection drugs	36%	41%	14%	10%	0.054
Underwent cervical treatment in past 6 months	0.0%	0.3%	0.0%	0.1%	1.00
HIV+ Subjects					
HAART		0.4%		48%	
CD4+, Median (IQR)		332 (159–518)		494 (337–698)	
% women CD4 < 200		30%		10%	

T-tests or Wilcoxon tests were used for continuous variables as appropriate and chi-square for categorical data.

Abbreviations: HAART, highly active antiretroviral therapy; IQR, interquartile range.

**Table 2.**

Average type-specific HPV prevalence (percentage) across all study visits by HIV-status/CD4 count, and GEE prevalence ratios (PRs) that contrast type-specific prevalence rates in HIV[+] women (with CD4 < 200) vs. HIV[-] women

HPV type	Phylogenetic Group	Rank of Lowest PRs <sup>2</sup>	HIV[+]			CD4+ >500 n = 816	CD4+ 200-500 n = 1,066	PR: HIV[+] vs. HIV[-]		P <sub>trend</sub> <sup>3</sup>
			All n = 2,566	HIV[-] n = 901	HIV[+] n = 1,665			PR: HIV[+] <200 vs. HIV[-]		
6	α10	6	0.49	1.97	0.70	1.86	2.52	4.78	<0.0001	
11	α10		0.11	1.23	0.48	1.03	8.21	19.24	<0.0001	
16 <sup>/</sup>	α9	2	1.27	3.33	1.85	3.35	2.26	3.85	<0.0001	
18 <sup>/</sup>	α7	9	0.74	3.12	1.40	2.82	3.25	5.81	<0.0001	
26	α5		0.10	0.57	0.12	0.42	4.24	10.84	<0.0001	
31 <sup>/</sup>	α9	3	0.59	2.44	1.29	2.31	2.96	4.86	<0.0001	
32	α1		0.28	2.58	1.12	2.83	7.83	14.90	<0.0001	
33 <sup>/</sup>	α9		0.43	2.81	1.32	2.86	5.46	9.94	<0.0001	
35 <sup>/</sup>	α9		0.21	1.15	0.39	1.02	4.85	12.26	<0.0001	
39 <sup>/</sup>	α7		0.22	1.18	0.25	0.95	5.15	15.94	<0.0001	
40	α8		0.09	0.79	0.26	0.65	8.52	16.13	0.01	
45 <sup>/</sup>	α7		0.55	2.83	1.64	2.29	4.58	8.26	0.001	
51 <sup>/</sup>	α5		0.76	3.06	1.28	2.85	4.41	8.76	<0.0001	
52 <sup>/</sup>	α9		0.57	2.93	1.73	2.51	4.92	10.53	<0.0001	
53	α6		1.29	6.40	2.82	6.09	5.17	9.73	<0.0001	
54	α13		0.29	2.71	1.52	2.38	9.44	16.51	0.0002	
55	α13		0.08	0.68	0.39	0.62	7.12	12.60	0.01	
56 <sup>/</sup>	α6		0.53	3.27	1.39	3.28	5.52	10.98	<0.0001	
58 <sup>/</sup>	α9		0.99	5.11	2.13	5.29	6.98	12.81	<0.0001	
59 <sup>/</sup>	α7		0.30	1.85	0.65	1.71	4.95	10.39	<0.0001	
61	α3		0.77	6.51	3.86	6.45	8.53	13.84	<0.0001	
62	α3	10	0.97	5.11	3.47	5.50	4.37	6.25	0.004	



HPV type	Phylogenetic Group	Rank of Lowest PRs <sup>2</sup>	HIV[-]		HIV[+]		CD4+ >500 n = 816	CD4+ 200-500 n = 1,066	PR: HIV[+] vs. HIV[-]	PR: HIV[+] <200 vs. HIV[-]	P <sub>trend</sub> <sup>3</sup>
			All n = 901	All n = 2,566	All n = 2,566	All n = 816					
66	α.6		0.38	2.37	1.59	3.04	10.15	18.03	0.0001		
67	α.9		0.15	0.44	0.24	0.31	8.15	13.44	0.25		
68 <sup>1</sup>	α.7	5	0.45	1.76	0.84	1.85	3.11	5.32	<0.0001		
69	α.5	4	0.00	0.39	0.22	0.61	2.32	5.03	0.08		
70	α.7		0.67	5.58	3.21	5.10	4.35	8.15	<0.0001		
71	α.15	1	1.28	3.44	1.79	3.33	1.83	1.78	0.12		
72	α.3	7	0.50	3.29	1.03	1.92	3.04	5.64	0.02		
73	α.11		0.00	1.63	1.45	2.22	7.39	15.41	<0.0001		
81	α.3	7	1.23	7.56	4.66	5.72	3.46	5.64	0.0001		
82	α.5		0.28	1.27	0.71	1.06	4.77	10.10	0.0001		
83	α.3		1.05	4.83	3.07	5.99	5.82	7.70	<0.0001		
84	α.3		0.49	3.02	1.74	3.56	11.36	21.78	<0.0001		
85	α.7		0.23	1.09	1.76	0.94	17.10	30.03	0.06		
89	α.3		0.15	0.59	0.24	1.09	8.73	23.12	0.0003		

<sup>1</sup> Carcinogenic.

<sup>2</sup> The GEE prevalence ratios (PR) were adjusted for age, smoking, ever IDU, baseline lifetime number of male sexual partner, number of male sexual partners in past 6 months, and cervical treatment in past 6 months. CD4 was treated as a time-dependent covariate.

<sup>3</sup> P-trend is based on average prevalence by CD4+ stratum among HIV+.

**Table 3.**

The level of heterogeneity ( $I^2$ ) in the HPV type-specific summary prevalence ratios (PRs) within each phylogenetic group

Grouping	List of HPV types	$\chi^2$	DF	p	$I^2$ (%)	Interpretation of heterogeneity
<i>Phylogenetic</i>						
$\alpha$ 3	HPV61, 62, 72, 81, 83, 84, 89	18.48	6	0.005	67.54	Substantial
$\alpha$ 5	HPV26, 51, 69, 82	1.07	3	0.785	<0	Not detectable
$\alpha$ 6	HPV53, 56, 66	3.00	2	0.223	33.32	Small
$\alpha$ 7	HPV18, 39, 45, 59, 68, 70, 85	10.91	6	0.091	45.01	Moderate
$\alpha$ 9	HPV16, 31, 33, 35, 52, 58, 67	17.39	6	0.008	65.49	Substantial
$\alpha$ 10	HPV6, 11	5.71	1	0.017	82.49	Substantial
$\alpha$ 13	HPV54, 55	0.04	1	0.838	<0	Not detectable
<i>Carcinogenicity</i>						
Carcinogenic	HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68	28.39	12	0.005	57.73	Substantial
Noncarcinogenic	HPV6, 11, 26, 32, 40, 53–55, 61, 62, 66, 67, 69–73, 81–85, 89	53.39	22	<0.001	58.79	Substantial

The level of heterogeneity was estimated using methods similar to those employed in meta-analysis. Specifically, the  $I^2$  statistic was obtained by adding an interaction between HPV type and HIV/CD4 strata in the GEE logistic regression models and then assessing the test statistics for the significance of this interaction.