The epidemiology of genital human papillomavirus infection

Helen Trotter*a, Eduardo L. Francoa,b,*

a Department of Oncology, McGill University, Montreal, Canada
b Department of Epidemiology & Biostatistics, Division of Cancer Epidemiology, McGill University, 546 Pine Avenue West, Montreal, QC, Canada H2W 1S6

Available online 13 December 2005

Abstract

Clinical and subclinical human papillomavirus (HPV) infections are the most common sexually transmitted infections in the world, and most sexually-active individuals are likely to be exposed to HPV infection during their lifetimes. More than 40 genotypes of HPV infect the epithelial lining of the anogenital tract and other mucosal areas of the body; of these, 13–18 types are considered to be high-oncogenic risk HPV types (HR-HPV). Persistent infection with HR-HPVs is now unambiguously established as a necessary cause of cervical cancer and is likely to be responsible for a substantial proportion of other anogenital neoplasms and upper aero-digestive tract cancers. Low oncogenic risk HPV types (LR-HPV) are also responsible for considerable morbidity as the cause of genital warts. Youth and certain sexual characteristics are key risk factors for HPV acquisition and persistence of HPV infection, but other mediating factors include smoking, oral contraceptive (OC) use, other STIs (e.g. chlamydia, herpes simplex virus), chronic inflammation, immunosuppressive conditions including HIV infection, parity, dietary factors, and polymorphisms in the human leukocyte antigen system. Not surprisingly, these factors are also established or candidate cofactors identified in epidemiologic studies of cervical cancer. HPV transmissibility and molecular events in HPV-induced carcinogenesis have been the focus of recent multidisciplinary epidemiologic studies. This shift in research focus coincides with a shift in cancer prevention techniques towards immunization with HPV vaccines and HPV testing of precancerous lesions.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Genital human papillomavirus; Epidemiology; Cervical cancer

1. Introduction

More than 120 different human papillomavirus (HPV) types have been catalogued so far [1,2], of which more than 40 infect the epithelial lining of the anogenital tract and other mucosal areas of the body. Some 13–18 types have been identified as probable or definite high-oncogenic risk HPV types (HR-HPV) (Table 1). It is now widely accepted that HR-HPV infections are a necessary, but not sufficient, cause of virtually all cases of cervical cancer worldwide and are a likely cause of a substantial proportion of other anogenital neoplasms and oral squamous cell carcinomas. Infection with low oncogenic risk HPV types (LR-HPV), such as HPV 6 and 11, can cause benign lesions of the anogenital areas known as condylomata acuminata (genital warts), as well as a large proportion of low-grade squamous intraepithelial lesions of the cervix. LR-HPV clinical infections are responsible for substantial morbidity and invoke high costs associated with the treatment of clinically relevant lesions. In this report, we review the essential aspects of the epidemiology of genital HPV infections and outline the risk factors for becoming infected with an emphasis on the role of HPV as cancer causing agents, particularly cervical carcinomas.

2. Descriptive epidemiology of genital HPV infection

2.1. How common is genital HPV infection?

HPV infections are the most common diagnosed sexually transmitted diseases today. Studies utilizing HPV DNA testing of asymptomatic women in the general population estimate the prevalence of HPV infection to be in the range of
Table 1
HPV types designated as of high oncogenic risk in representative studies and reviews

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>Probable</td>
<td>X</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>33</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>35</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>39</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>45</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>51</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>52</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>53</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>55</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>56</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>58</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>59</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>66</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>68</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>73</td>
<td>Pap238A, MM9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>Probable</td>
<td>X</td>
</tr>
<tr>
<td>82</td>
<td>W13B, MM4 C139 (subtype)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>Probable</td>
<td>X</td>
</tr>
<tr>
<td>83</td>
<td>Pap291, MM7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Probable</td>
<td>X</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No.</th>
<th>9</th>
<th>11</th>
<th>13</th>
<th>14</th>
<th>18</th>
<th>15–18</th>
<th>13</th>
</tr>
</thead>
</table>

* The HR-HPVs listed have become widely used as probes in diagnostic assays used in epidemiologic and clinical studies. For instance, the HR-HPVs under Nindl et al., 1998, are part of the probe B set in the commercially available Hybrid Capture 2 assay (Digene Co.).

b As above, for the GP5+/6+ general primer polymerase chain reaction (PCR) widely used in many international studies of cervical cancer etiology and screening.

c As above, for the PGMY line blot PCR protocol produced by Roche Diagnostics that is currently under evaluation as a diagnostic tool.

d Based on a consensus review panel of all published studies by the International Agency for Research on Cancer as part of its Carcinogenicity Evaluation Monograph, vol. 90.

2–44% [3–9]. This wide variation in prevalence estimates is largely explained by age differences among population samples studied, and by differences in the molecular sensitivity of the various HPV DNA assays used to detect viral DNA [7]. It is well appreciated that sexually active young adults are most at risk for acquiring HPV, as epidemiologic studies have shown that the prevalence of HPV infection is highest among young sexually active women [10,11]. Many epidemiological studies have found an age-related decline in HPV prevalence (Fig. 1). Even in a study of prostitutes, a group which is highly exposed to HPV, Kjaer et al. observed a substantial decrease in HPV prevalence with age despite continuously high levels of sexual activity [12]. One possible explanation for this finding is that infected individuals develop adaptive immune responses against HPV that prevent future infection. Also common in many studies, however, is a second pattern with a peak in HPV prevalence among women younger than 25 years of age, followed by the expected decline in prevalence until around age 45–50 and then a second peak in the peri- or post-menopausal years [13–16]. Although the reason for this second, menopausal peak is not clear, it could be plausibly attributed to one or more non-mutually exclusive mechanisms, such as reactivation of latent infections acquired earlier in life due to a gradual loss of type-specific immunity, or to acquisition of new infections due to sexual contacts with new partners later in life. Also, plausible is a cohort effect: age-related variations in prevalence may reflect the diverse HPV exposure of successive birth cohorts. Sexual mores have changed during the last several decades, which may have influenced the HPV exposure of different age cohorts. Unfortunately, the lack of data...
about historical HPV prevalence or cohort studies with long-term follow-up (>15 years) makes it impossible to empirically refute or confirm this hypothesis. Nonetheless, it is clear that the two patterns of age-specific risk are observed for incident HR-HPV infections [17,18].

2.2. Burden to the health care system

HPV infections and their clinically relevant sequelae pose a substantial burden to the health care system. There are an estimated 6.2 million new cases of HR-HPV infection occurring in the US each year, and approximately 20 million Americans are infected with HPV at any one time [19]. HR-HPV infections can cause precancerous cervical lesions that are detected by routine cytological screening with the Papanicolaou (Pap) test. If these lesions are left undiagnosed, they may progress to invasive cervical cancer within a few months to several years (depending on the precancerous lesion grade). Invasive cervical cancer is the second most common cancer in women worldwide [20]. In the US, even though rates have declined during the past 50 years from levels that were comparable with those found in developing countries today, there are still more than 12,000 cases of cervical cancer and more than 4000 deaths from the disease annually [21].

For each new case of invasive cancer found by Pap cytology, estimates suggest that there are approximately 50–100 squamous intraepithelial lesions (SILs). Women with these precancerous lesions need close monitoring by cytology and, if results persist, also by colposcopy and biopsy. Additionally, many women are diagnosed with atypical squamous cells of undetermined significance (ASC-US). Current screening guidelines suggest that these women may undergo repeat cytology or HPV testing, or be sent for colposcopic examination. Altogether, ASC-US and SIL findings account for more than 10% of all Pap smears that are processed in cytopathology laboratories. The diagnosis, management, and follow-up of patients with HPV-induced cervical abnormalities place an important burden on the health care system of nations with opportunistically or organized cervical cancer screening programs. Moreover, in the US, false-negative Pap smears from women with precancerous lesions caused by HPV infection are among the most frequent reasons for medical malpractice litigation [22]. HR-HPV infections are also thought to cause many other anogenital and upper aero-digestive tract cancers. An estimated 85% of anal cancers; 50% of the cancers of the vulva, vagina, and penis; 20% of oropharyngeal cancers; and 10% of laryngeal and esophageal cancers are attributable to HPV [23–25].

LR-HPV infections cause anogenital warts (condylomata acuminata) and low-grade cervical lesions. Unlike malignant neoplasms whose incidence can be measured based on statistics compiled by population-based tumor registries, or syphilis and gonorrhea, sexually-transmitted diseases (STD) of compulsory notification, the occurrence of anogenital condylomata can only be measured indirectly, via hospital series or physician consultation statistics. Although we cannot measure the incidence of genital warts directly, there is strong indication that it has increased substantially in most Western countries during the last few decades. It is estimated that such lesions affect about 1% of all sexually active adults in the United States [26].

The true incidence of genital warts was measured in a study conducted by Mayo Clinic investigators in Rochester, Minnesota, during the period 1950–1978 [27]. This study found that rates had increased substantially during the 29-year period, peaking in 1975 with an average annual incidence of 107 new cases per 100,000. Although a population-based study in a small community in a midwestern state may not truly reflect national trends, the rates seen in the Rochester study probably represent a lower bound for present-day rates in most urban centers in North America. At an approximate annual rate of 100 per 100,000, genital warts are as common a disease as two common malignant neoplasms of grave concern for Americans, namely, breast and prostate cancers. Such a level of incidence can be translated into a lifetime cumulative risk approaching 10% [28].

In addition to the above population-based statistics, the National Disease and Therapeutic Index (NDTI)/IMS America Ltd.) counted the number of initial consultations for diagnosis and treatment of genital warts in a stratified random sample of private practitioners in the US. Although they do not represent true incidence rates, and fail to account for subclinical warts, the NDTI data have been widely used as a nationwide surveillance indicator for genital warts for inclusion in the STD surveillance reports by the Centers for Disease Control and Prevention [29]. Such clinical visits increased from 56,000 in 1966 to more than 350,000 in 1987. A downward trend occurred until 1997 (145,000 visits), but the numbers have since then increased to 264,000 visits in 2003.

2.3. Incidence of HPV infection measured by molecular detection techniques

A number of cohort studies that began in the early 1990s have provided useful data concerning the natural history of HPV infections. These studies take repeated measurements over time and use DNA hybridization techniques based on target amplification or signal amplification to test for HPV infection. Despite the heterogeneity of study designs and HPV testing systems, some clear similarities in results have emerged. The clearest and most consistent finding is the relatively high incidence of HPV infections among women who were initially HPV negative. Table 2 summarizes the findings from the cohort studies that used polymerase chain reaction (PCR) for HPV detection. The incidence rates in these studies indicate that up to 3% per month of the women who are initially free of HPV DNA may be found on subsequent testing to have acquired an HPV infection.

This translates into high cumulative risks. For instance, among women aged 15–19 years in England, and among two cohorts of college women in the US, the cumulative incidence
of HPV infection exceeded 40% after 3 years [30–32]. In the Brazilian Ludwig–McGill cohort (women with mean age of 33.3), the cumulative incidence was 24% over 18 months [17]. In addition, the incidence of HR types seems to be higher than the incidence of LR types, but this observation is partially dependent on how many HPV types the different studies classified as high risk. In a cohort study of American women aged 18–35 years, Giuliano et al. found the cumulative risk of new HR types at 12 months to be 32% compared with 18% for LR types [33]. Acquisition of infections was highest for HPV 16, HPV 39, HPV 84 and HPV 51: 5.9, 4.6, 4.2 and 3.4 new infections per 1000 women-months, respectively [34]. Incidence rates in a Canadian cohort of female university students were 14.0 cases/1000 women-months for HR-HPVs and 12.4 cases/1000 women-months for LR-HPVs [8]. HPV 16 is usually the most common type, irrespective of study design and geographical area. HPV 18 is also a common type [35], but the prevalence/incidence of HPV 18 varies according to the population studied. For example, Richardson et al. found the 24-month cumulative incidence rates to be highest for HPV 16 (12%), followed by HPV 51 and HPV 84 (8%) [8]. Munoz et al. found that the highest incidences in Colombia were for HPV types 16, 58, 31 and 18, representing 16, 11.2, 10.9 and 10.6% of all infections, respectively [18].

2.4. Persistence of HPV infection

HPV infections are common, but most infections seem to clear spontaneously; cohort studies have consistently found that only a small proportion of women positive for a given HPV type are found to have the same type in subsequent specimens [8,17,30,36–38]. Whether infections clear completely or the virus remains latent in basal cells at undetectable levels is a matter of debate and cannot be verified empirically. What is clear, however, is the fact that risk of subsequent cervical intraepithelial neoplasia (CIN) is proportional to the number of specimens testing positive for HPV [39,40]. This suggests that carcinogenic development results from HPV infections that persist productively (i.e. with sustained viral replication within the squamous epithelium) for prolonged periods of time. Persistence of HR-HPV types (see Table 1 for HR-HPVs) is strongly linked to precancerous cervical lesions and invasive cancer [41–47].

Although the above observations concerning the importance of HPV persistence as a key intermediate in cervical carcinogenesis are not challenged, there is no consensus regarding the definition of what constitutes a persistent or a transient HPV infection. At a minimum, determination of HPV infection status at a baseline visit and at a subsequent follow-up opportunity 6 months or 1 year later allows good discrimination of infections that have higher propensity to develop into cervical lesions. However, a better understanding of the natural history of individual HPV infection episodes can only be obtained from cohort studies that resort to multiple, repeated measurements of viral endpoints over short intervals during several years. Such studies are ongoing in many populations.

There is considerable heterogeneity in study design and methods among the various published cohort investigations that have measured the natural history of persistent HPV infections. It is ascertained from all of these studies that there is no clear duration threshold that can be used to distinguish an HPV infection episode that is transient from one that is persistent. Furthermore, for common HPV types, such as HPV 16, it is virtually impossible to distinguish an instance of persistent infection from one that simply represents loss of, and subsequent reacquisition of, the same HPV type from the same or from a different sexual partner. Some experts have proposed testing for molecular variants of HPV, which may provide a finer scrutiny of same-type infections to distinguish true cases of persistent infection from those that represent clearance of one HPV variant and acquisition of another [48]; however, subsequent studies have shown that, at least for HPVs 16 and 18, the vast majority of persistent infections tend to involve the same variant over multiple testing opportunities [49].
2.5. Mean duration and clearance of HPV

Several studies have reported the duration and clearance rates of individual episodes of HPV infections (Table 3). Proper estimation of the mean and median duration and clearance rates requires using actuarial methods such as the Kaplan-Meier estimating technique, which takes into account the censoring of events at the time of data analysis. The importance of this analytical precaution cannot be overemphasized. For instance, if one wanted to measure the mean duration of all individual HPV 16 infections that occur in a cohort of women it would be wrong to average only those episodes in which clearance has already occurred. In any prospective cohort study with ongoing data and specimen collection, there will always be cases where infection has not yet cleared at the time of data analysis. These cases represent censored events in which the duration is likely to be longer than the time elapsed since the infection was first detected. Conditions caused by subject dropout, missing specimens, or pending (or missing) HPV test results are also consistent with the same interpretation. All such cases of incomplete events represent the reality of the HPV natural history and must be taken into account in the calculations. Excluding such cases (in order to opt for the simple arithmetic that includes complete episodes only) will substantially underestimate the true average duration of HPV 16 infection episodes in that cohort. Table 3 only includes studies that used actuarial methods to account for censored observations.

In general, across all studies averages of individual episodes last from 4 to 20 months, clearance rates vary substantially, but between 10 and 60% of all women who develop an infection will still be infected with the same type of HPV 1 year later. Most studies indicate, however, that less than half will continue to be positive at 12 months. HR-HPV infections tend to last longer than those of LR-HPV types. Among the former, HPV 16 infections tended to be among those of longer duration [8, 32], but the difference does not seem to be substantial, particularly in view of the overlap among statistical confidence bands around the estimates. Other findings from some of the studies reviewed in Table 3 provide evidence for a longer duration of HPV 16 infections. For instance, in the cohort study of Bisson et al., HPV 16, 18, 21/33/55 appeared more persistent (odds ratio [OR], 2.5; 95% CI, 1.0–6.2) than other types [50]. Molano et al. found that HPV 16 had a significantly lower clearance rate than infections with LR types (rate ratio [RR] = 0.47; 95% CI, 0.32–0.72), HPV types related phylogenetically to HPV 16 (types 31, 33, 35, 52, 58) had intermediate clearance rates (RR = 0.62; 95% CI, 0.47–0.84), and other HR types did not show evidence of slower clearance compared with LR types [38].

2.6. Co-infection with multiple HPV types

Co-infection with multiple HPV types is a common finding of many epidemiologic studies. In the Brazilian Ludwig-McGill cohort, multiple types were detected at the same visit in one-fifth of all women who tested positive for HPV at any time during follow-up [51]. Moreover, it seems that infection with a given type does not decrease the probability of being infected by phylogenetically related types. Thomas et al. found that acquisition of multiple infections occurred more often than expected by chance and that the risk of acquiring a specific HPV type was not substantially decreased among those with prior infection with a phylogenetically related type (HPV: 16 and 31; 18 and 45; 6 and 11) [52]. Rousseau et al. found that HPV 16 and 18 co-infections occurred less frequently than expected by chance as were other pairwise instances of HR-HPV types, such as HPV: 16 and 52; 16 and 68; and 18 and 6/11 [53].

The role of co-infections with additional HPV types on the duration of an episode with a given HPV type is not clear. In general, infections with single and multiple HPV types have comparable clearance rates. However, it is difficult to make cross-study comparisons because the definition of persistence varies appreciably across investigations. Ho et al. and Perrons et al. found that infection with multiple types of HPV was associated with persistent HPV infection [30, 54]. Woodman et al. found that simultaneous infection with HPV 16 and another type resulted in longer duration of an HPV 16 episode as compared with single infection with HPV 16 [32]. However, not all data supports this. Rousseau et al. observed that persistence of HPV infection was independent of co-infection with other HPV types [55]. Liaw et al. found that the presence of HPV 16 was associated with an excess risk for acquisition of other types without affecting the persistence of the episodes with the additional types [56]. It is important to keep in mind that the PCR assays used in most epidemiologic studies have not been optimized for amplifying multiple types in the same reaction mixture and that the target DNA sequences of some HPV types are amplified less efficiently than others. These methodological issues may result in underestimates of the number of types in a specimen and handicap research into the interplay between multiple types of HPV infection.

3. HPV infection as the causal intermediate in cervical carcinogenesis

Fig. 2 shows the various components of the etiological model for the natural history of HPV infection and cervical carcinogenesis. The distal risk factor is sexual activity, which has been known for at least five decades to be the most important correlate of cervical cancer risk. As discussed above, the role of persistent HPV infection is central to this etiologic model.

A role for HPV infection in the sexually transmitted disease model of cervical cancer was proposed in the mid-1970s [57]. Strong evidence for an etiologic role of HPV was slow in coming, largely due to inconsistency in epidemiologic findings (reviewed in [58]). The issue was further complicated because prevalence of HPV DNA in cervical tumor speci-
Table 3
Mean and median duration and clearance rates for genital (cervical) HPV infections in different cohort studies of women

<table>
<thead>
<tr>
<th>Study (subjects' mean age)</th>
<th>Mean duration of HPV infection in months</th>
<th>Proportion of women who cleared HPV</th>
<th>Mean time between visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hildesheim et al., 1994 (26 years) [36]</td>
<td>∼1 year</td>
<td>57%</td>
<td>14.6 month (2 visits only)</td>
</tr>
<tr>
<td>Hinchliffe et al., 1995 [95]</td>
<td>∼4 months</td>
<td>93%</td>
<td>4 month (2 visits only)</td>
</tr>
<tr>
<td>Evander et al., 1995 [96]</td>
<td>2 years</td>
<td>80%</td>
<td>NA (2 visits only)</td>
</tr>
<tr>
<td>Brisson et al., 1996 (&lt;29 years) [50]</td>
<td>∼6 months</td>
<td>55%</td>
<td>3.5 month (2 visits only)</td>
</tr>
<tr>
<td>Ho et al., 1998 (20 years) [30]</td>
<td>Any HPV: 8 (7–10)</td>
<td>70%</td>
<td>6 month (multiple visits)</td>
</tr>
<tr>
<td>Moscicki et al., 1998 (∼20 years) (after 3 HPV-visits) [37]</td>
<td>Any HPV: ∼15.0</td>
<td>4 month (multiple visits)</td>
<td></td>
</tr>
<tr>
<td>Franco et al., 1999 (33 years) [17]</td>
<td>HR: 8.9 (7.6–10.2)</td>
<td>5 year (5 visits)</td>
<td></td>
</tr>
<tr>
<td>Elfgren et al., 2000 (35 years) [97]</td>
<td>HR: 8.1 (7.6–9.3)</td>
<td>4 year (5 visits)</td>
<td></td>
</tr>
<tr>
<td>Woodman et al., 2001 (18 years) [32]</td>
<td>HR: 4.8 (3.9–5.6)</td>
<td>4 year (5 visits)</td>
<td></td>
</tr>
<tr>
<td>Abdol et al., 2001 (∼30 years) [83]</td>
<td>Any HPV: 13.7 (8.0–25.4)</td>
<td>5 year (5 visits)</td>
<td></td>
</tr>
<tr>
<td>Giuliano et al., 2002 (∼28 years) [34]</td>
<td>HPV-16: 10.3 (6.8–17.3)</td>
<td>5 year (5 visits)</td>
<td></td>
</tr>
<tr>
<td>Elfgren et al., 2002 (35 years) [98]</td>
<td>HPV-18: 7.8 (6.0–12.6)</td>
<td>5 year (5 visits)</td>
<td></td>
</tr>
<tr>
<td>Xi et al., 2002 (18.7 years) [99]</td>
<td>Any HPV: HR: 9.8</td>
<td>3 month (multiple visits)</td>
<td></td>
</tr>
<tr>
<td>Richardson et al., 2003 (23 years) [8]</td>
<td>LR: 4.3</td>
<td>4 month (multiple visits)</td>
<td></td>
</tr>
<tr>
<td>Dalstein et al., 2003 (∼35.7 years) [100]</td>
<td>HPV-16: 8.5</td>
<td>3 month (multiple visits)</td>
<td></td>
</tr>
<tr>
<td>Sellors et al., 2003 (32.7 years) [69]</td>
<td>European variants: 17.2</td>
<td>4 month (multiple visits)</td>
<td></td>
</tr>
<tr>
<td>Molano et al., 2003 (∼29 years) [38]</td>
<td>HR: 7.5</td>
<td>6 month (multiple repeat visits)</td>
<td></td>
</tr>
<tr>
<td>de Sanjose et al., 2003 (∼51 years) [101]</td>
<td>HR: 4.0 (1.1–14.3)</td>
<td>1 year (2 visits)</td>
<td></td>
</tr>
<tr>
<td>Munoz et al., 2004 (32.3 years) [18]</td>
<td>HR: 14.8 (3.1–17.0)</td>
<td>1 year (2 visits)</td>
<td></td>
</tr>
</tbody>
</table>

Note: CI = Confidence Interval; HR = Hazard Ratio; LR = Likelihood Ratio.
### Table 3 (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>HPV type</th>
<th>Mean duration of HPV infection in months (95%CI)</th>
<th>Mean time between visits</th>
<th>Proportion of women who cleared HPV</th>
<th>Median duration of HPV infection in months (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moscicki et al., 2004 [102]</td>
<td>Any HPV</td>
<td>13.4 (4 years)</td>
<td>4 months (2 visits only)</td>
<td>Any HPV: 13.4</td>
<td>Any HPV: 13.4 (4 years)</td>
</tr>
<tr>
<td>Perrons et al., 2005 [54]</td>
<td>HPV type 16-like</td>
<td>15.6</td>
<td>6 month (HR)</td>
<td>6 month (2 visits only)</td>
<td>HPV type 16-like: 15.6 (6 months)</td>
</tr>
<tr>
<td>Brown et al., 2005 [103]</td>
<td>HPV type 18-like</td>
<td>15.2</td>
<td>3 month (multiple visits)</td>
<td>3 month (multiple visits)</td>
<td>HPV type 18-like: 15.2 (3 months)</td>
</tr>
<tr>
<td>Braun et al., 2005 (5-Year) [101]</td>
<td>HPV type 11-like</td>
<td>14.1</td>
<td>3 year (HR)</td>
<td>3 year (HR)</td>
<td>HPV type 11-like: 14.1 (3 years)</td>
</tr>
<tr>
<td>Ekbom et al., 2005 (15-Year) [32]</td>
<td>HPV type 5-like</td>
<td>15.6</td>
<td>6 month (2 visits only)</td>
<td>6 month (HR)</td>
<td>HPV type 5-like: 15.6 (6 months)</td>
</tr>
</tbody>
</table>

a Except for the Woodman et al. [32] study which used inter-quartile ranges, all other intervals represent 95% CIs.

---

**Fig. 2.** Etiologic model of human papillomavirus (HPV) infection as a necessary cause of cervical cancer incorporating the role of host, reproductive, lifestyle, and viral co-factors. Abbreviations: CIN: cervical intraepithelial neoplasia (a histological classification); LSIL: low-grade squamous intraepithelial lesions; HSIL: high-grade squamous intraepithelial lesions; HLA: human leukocyte antigen.

Source: Franco and Harper, 2005, with permission.
The causative link between HPV infection and cervical cancer opens up new paths of prevention, including screening for infection with high-risk types of HPV and immunization to prevent infection with high-risk types of HPV. Indeed, prophylactic HPV vaccines directed against the most common LR- and HR-HPV types are expected to prevent a substantial number of HPV infections. Further methods of prevention may become available when more is known about the mechanism whereby HPV infection initiates cervical carcinogenesis and the transmissibility of HPV infection; however, prevention through vaccination is predicted to substantially reduce HPV-associated morbidity and mortality.

4. Risk factors for HPV infection

Risk determinants for HPV infection that have been identified in various cross-sectional and prospective cohort studies include number of sexual partners (lifetime and recent), age at first intercourse, smoking, oral contraceptive (OC) use, other STIs (e.g. chlamydia and herpes simplex virus), chronic inflammation, immunosuppressive conditions including HIV infection, and parity [8,31,47,67–70]. Results have been inconsistent partly owing to the fact that different populations have been studied. Furthermore, risk factor profiles have been found to differ depending on whether HR- or LR-HPV types were considered [33,71–73]. Nevertheless, in addition to sexual activity correlates, the most consistent determinant of HPV infection is age, with most studies indicating a sharp decrease in risk after the age of 30 [33,55,60]. The decrease in risk of HPV infection with increasing age seems to be independent of changes in sexual behavior, suggesting a role for immune response. As described above, however, a second peak at older ages (Fig. 1) is also a common feature of epidemiologic findings.

Ho et al. surveyed college-aged women and found that an increased risk of HPV infection was significantly associated with younger age, Hispanic ethnicity, black race, an increased number of vaginal-sex partners, high frequencies of vaginal sex and alcohol consumption, anal sex, and certain characteristics of partners (regular partners having an increased number of lifetime partners and not being in school) [30]. Richardson et al. found lifetime frequency of sexual intercourse and lifetime number of oral sex partners was associated with HR-HPV infections, whereas HPV infection with LR types was invariant with respect to markers of sexual activity [74]. Rousseau et al. found a strong negative association between age and HR-HPVs, but not with LR types, and OC use was strongly and exclusively associated with HR-HPV and HPV 16 infections [55]. Overall, markers of sexual activity were strongly associated with all types of infections.

Age at sexual debut may increase risk of HPV infection either as a marker for other sexual behaviors (e.g. young age at sexual debut may be associated with a greater lifetime number of sexual partners or with propensity to engage new sexual partners more frequently), or as a true causal risk factor due to greater cervical ectopy during adolescence [75,76]. Research on the effect of condom use has found equivocal results [77,78]. A paradoxical effect is occasionally reported, such that condom use appears to increase risk of HPV infection [72,74,77]. This is likely a result of an increased probability of infection among partners with whom condoms are used. For example, people tend to use condoms with partners whom they consider to be more at risk (e.g. new partners, casual partners, sex trade workers), but not with partners whom they consider to be safe (e.g. long-term, regular partners or spouses) [79–81]. HPV transmission may also occur through non-penetrative sexual activity. Indeed, low-risk HPV has been detected on fingers of individuals with genital warts [82]. Furthermore, among virgin female college students, “any type of non-penetrative sexual contact” was associated with a 2.4% 24-month cumulative incidence of HPV infection, although this type of activity did not result in increased risk among women engaging in intercourse [31].

The thrust of epidemiological research in recent years has focused on understanding the role of risk factors that influence acquisition of persistent HPV infection, or of factors that mediate progression in the continuum of lesion grades [32]. Some authors have shown that persistence is associated with older age [13,30,36,83]. A few studies have also found a protective effect on the risk of persistent infection associated with consumption of fruits and vegetables and dietary or circulating levels of vitamin C and E, beta-carotene, lycopene, lutein/zeaxanthin, and cryptoxanthin [84–86]. Some studies have also suggested a role for viral factors (molecular variants and viral load) in persistence of HPV and progression [49,87]. The fact that HPV infection does not always progress to neoplastic disease also suggests that interpersonal variations in the immune system may play a role in the clearance of HPV infections and/or in their acquisition. Such a mechanism may be linked to genetic polymorphisms in the Human Leukocyte Antigen (HLA) system. The HLA genes, whose protein products are involved in antigen presentation to T cells, play a role in the regulation of immune response. Certain HLA alleles or haplotypes seem to be involved with susceptibility to HPV infection and cervical cancer probably by regulating the immune response against HPV infection and ultimately interfering in the establishment of productive persistent infections and lesion development [88]. Maciag et al. found that among the class II HLA genes, the DRB1*0301–DQB1*0201 haplotype was associated with a two-fold reduction in risk of any HPV infections, regardless of duration [89]. The DRB1*1102–DQB1*0301 haplotype was associated with reduced risk of persistent infections, whereas the DRB1*1601–DQB1*0502 and DRB1*0807–DQB1*0402 ones were associated with a seven-fold and a three-fold increase, respectively, in risk of persistent HPV infections. Although these associations seemed moderate to strong, the rarity of these alleles and haplotypes and the costs related to HLA typing have precluded replication of findings in multiple cohort studies.
the knowledge base that has led to the ongoing paradigm shift in understanding the role of HPV in cervical neoplasia. The advent of modern amplified DNA hybridization techniques and the demonstration that HPV infection is the cause of cervical cancer, and as such, they have dealt with the virus mostly as an intermediate endpoint in cervical carcinogenesis. The majority of studies have included women only, and although we are beginning to learn more about the epidemiology of HPV infections in men, much still needs to be done.

A new generation of epidemiologic studies is now emerging with a focus on HPV as a sexually-transmitted infection. These studies are making judicious use of molecular tools and novel statistical approaches to understand HPV transmission and key steps in the natural history of HPV-related malignancies. Issues of transmissibility such as infectivity following a sexual encounter and the controversial effect of condom use can only be properly addressed in studies involving forming couples as the unit of observation. This new generation of studies will provide the answers needed by women and their health care providers as cervical cancer prevention shifts from an oncological to an STD perspective. Prevention of HPV infection through vaccination is preferable to treatment of precancerous and cancerous lesions, and is expected to dramatically reduce the morbidity and mortality associated with HPV infection.

5. Conclusion

Molecular epidemiologic studies of the natural history of HPV infection and cervical neoplasia have provided much of the knowledge base that has led to the ongoing paradigm changes in cervical cancer prevention via HPV screening and HPV vaccination. The contributions of epidemiology in this process during the last 15 years have been focused on the demonstration that HPV infection is the cause of cervical cancer, and as such, they have dealt with the virus mostly as an intermediate endpoint in cervical carcinogenesis. The majority of studies have included women only, and although we are beginning to learn more about the epidemiology of HPV infections in men, much still needs to be done.

A new generation of epidemiologic studies is now emerging with a focus on HPV as a sexually-transmitted infection. These studies are making judicious use of molecular tools and novel statistical approaches to understand HPV transmission and key steps in the natural history of HPV-related malignancies. Issues of transmissibility such as infectivity following a sexual encounter and the controversial effect of condom use can only be properly addressed in studies involving forming couples as the unit of observation. This new generation of studies will provide the answers needed by women and their health care providers as cervical cancer prevention shifts from an oncological to an STD perspective. Prevention of HPV infection through vaccination is preferable to treatment of precancerous and cancerous lesions, and is expected to dramatically reduce the morbidity and mortality associated with HPV infection.

References

4. Gjorven K, Olsen AO, Magnus P, Grade B, Sauer T, Ostvold I. Prevalence of human papillomavirus in cervical scrapes, as ana-

Table 4 summarizes the above variables with respect to the strength of the association, their consistency, and specificity to different types of HPVs as obtained in multiple epidemiologic studies conducted during the last 15 years, since the advent of modern amplified DNA hybridization techniques (target or signal amplification). It is not surprising that most of these variables have been shown to be risk factors in case-control or cohort studies of pre-invasive and invasive cervical cancer because they are predictors of HPV infection at the distal causal intermediate in cervical carcinogenesis. It is possible also that some of the variables listed in Table 4 may also exert a role on cervical carcinogenesis that is ”downstream” from HPV infection. For instance, smoking may contribute additional genetic damage to epithelial cells already harboring the initiating events triggered by HPV infection (e.g. interference with the cellular p53 and Rb gene pathways via over expression of the E6 and E7 viral oncoproteins). Such additional damage may accelerate progression to malignancy. Likewise, the role of other STIs, such as chlamydia, is likely to occur via inflammatory activity that may result in oxidative damage to HPV infected cells. Chlamydia infection is also likely to lead to cervical inflammation, which may facilitate HPV infection to become persistent. In such a scenario, chlamydia infection would act as a risk factor for HPV infection and should thus be listed in Table 4. However, whether the latter biological mechanism occurs (and if it does, whether it can be distinguished from the inflammation-based mechanism) is very difficult to prove in an epidemiologic study because both microbial agents are strongly linked to sexual activity and adjustment for the latter variable would preclude assessment of an independent effect of chlamydia infection on the risk of HPV.

Table 4

<table>
<thead>
<tr>
<th>Determinant or risk factor</th>
<th>Direction of association</th>
<th>Strength of the association</th>
<th>Consistency</th>
<th>Specificity for HPV types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Negative, U- or J-shaped$^a$</td>
<td>++</td>
<td>+</td>
<td>Yes, country-dependent</td>
</tr>
<tr>
<td>Sexual activity markers</td>
<td>Negative for age at sexual debut</td>
<td>+</td>
<td>+</td>
<td>Yes (HR types)</td>
</tr>
<tr>
<td>Oral contraceptive use</td>
<td>Positive (duration)</td>
<td>+</td>
<td>+</td>
<td>Yes (HR types)</td>
</tr>
<tr>
<td>Condom use</td>
<td>Negative or inconsistent</td>
<td>+</td>
<td>/</td>
<td>Possibly</td>
</tr>
<tr>
<td>Smoking</td>
<td>Positive</td>
<td>+</td>
<td>/</td>
<td>No</td>
</tr>
<tr>
<td>HLA polymorphisms</td>
<td>Variable, allele and haplotype dependent</td>
<td>++</td>
<td>=</td>
<td>Yes, allele and haplotype dependent</td>
</tr>
<tr>
<td>Nutrition</td>
<td>Negative for specific carotenoids and other dietary or circulating micronutrients</td>
<td>+</td>
<td>+</td>
<td>No</td>
</tr>
</tbody>
</table>

$^a$ Related to the average RR or OR across all studies: +: weak to moderate association, ++: moderate to strong association.

$^b$ Related to how consistent the association has been across studies.

$^c$ Whether the association seemed to vary in magnitude for HR and LR HPVs or for individual HPV types, such as HPV 16.

$^d$ See Fig. 1 and text for details on age-specific prevalence patterns in different populations.


