A Long-term Prospective Study of Type-Specific Human Papillomavirus Infection and Risk of Cervical Neoplasia Among 20,000 Women in the Portland Kaiser Cohort Study

Mark Schiffman1, Andrew G. Glass2, Nicolas Wentzensen1, Brenda B. Rush2, Philip E. Castle1, David R. Scott3, Julie Buckland3, Mark E. Sherman1, Greg Rydzak3, Peter Kirk1, Attila T. Lorincz4, Sholom Wacholder1, and Robert D. Burk5

Abstract

Background: Human papillomavirus (HPV) DNA testing is more sensitive than cytology for detection of cervical intraepithelial neoplasia grade 3 and cancer (≥CIN3). Adding HPV testing to cytology is recommended for women ≥30 but long-term prospective studies of HPV testing are rare.

Methods: Beginning in 1989–1990, ~20,000 women in a prepaid health maintenance organization (median age = 34) were followed passively by recommended annual cytology. We tested archived cervicovaginal lavage specimens collected at enrollment, primarily by MY09-MY11 PCR-based methods, for carcinogenic HPV types. We calculated positive and negative predictive values for the entire study period, and Kaplan–Meier estimates of cumulative probability for ≥CIN3, up to 18 years of follow-up.

Results: We observed 15 cases of invasive cervical cancer during the study period, and 189 cases of CIN3. Predictive values and Kaplan–Meier analyses yielded the same conclusions. In women 30 and older, the reanalysis against ≥CIN3 following a single negative HPV test was long-lasting (cumulative probability = 0.7% during follow-up). In this age group, a single HPV test (positive vs. negative, hazard ratio of 8.5, 95% CI = 4.8–15.1) provided greater long-term risk stratification than a single cytologic result (abnormal vs. normal, HR = 2.9, 95% CI = 1.2–6.6). The risk for ≥CIN3 was higher for HPV16 than for the average of the other carcinogenic types (hazard ratio = 2.7).

Conclusion and Impact: The data from this cohort study show the long-term predictive value of HPV testing, particularly in women ≥30, and a possible role for distinguishing particularly carcinogenic types like HPV16. Cancer Epidemiol Biomarkers Prev; 20(7); 1398–409. ©2011 AACR.

Introduction

Cross-sectional investigations including cancer case series and case-control studies (1, 2) provided the first epidemiologic evidence to support the causal link between human papillomavirus (HPV) infection and cervical cancer. Large prospective cohort studies initiated in the late 1980s and early 1990s (3–6) showed in their first years of follow-up that HPV DNA detectability, particularly of the most carcinogenic HPV types like HPV16, precedes and predicts the subsequent diagnosis of the obligate surrogate endpoint for risk of invasive cancer, i.e., cervical intraepithelial neoplasia grade 3 (CIN3) including carcinoma in situ. Conversely, HPV negativity predicts a low risk of subsequent CIN3. It is now certain that persistent infections with a group of approximately a dozen carcinogenic HPV genotypes cause virtually all cases of cervical cancer worldwide (7, 8). HPV16 is known to cause approximately half of cases; HPV18 causes another fifth.

Accordingly, we have moved beyond observation to successful trials and implementation of vaccines and HPV testing. Randomized controlled trials of prophylactic HPV vaccines have shown that prevention of HPV infection prevents subsequent lesions including CIN3 (9). Multiple trials have proven that HPV DNA testing is more sensitive (but less specific) than cytology (Pap tests) in screening for CIN3 [summarized in (10)].

Thus, it is timely to evaluate how HPV testing and cytology should each be used for primary screening. A single cytologic screen is valuable mainly because of good positive predictive value. An abnormal cytologic result indicates increased underlying risk of having...
CIN3 and developing cancer (invasive cancer is uncommonly found in screening), but a single normal result is not sensitive enough to provide strong long-term reassurance. For that reason, the extremely risk-averse strategy in the United States relies on Pap tests every 1–3 years (11, 12).

Co-testing for HPV DNA and cervical cytology is already recommended by several influential organizations, for women 30 and older who are past the peak of acute, typically clearing infections (13). Given the years required to progress from HPV infection to cancer, more sensitive HPV testing will likely provide useful longer-term reassurance than cytology, but it is not clear how to adjust screening intervals optimally to maximize safety while avoiding overtreatment and overspending for unnecessary procedures.

As translational work continues based on the newer trials, the value of the longest-running cohorts like the Portland Kaiser cohort described here is to “complete the picture” of HPV natural history, i.e., to estimate and compare long-term prognostic absolute risks, i.e., predictive values for risk for CIN3 of testing positive or negative for HPV. This is a concluding report of the Portland Cohort, which enrolled the first of more than 23,000 participants in 1989. This analysis incorporates an additional 5+ years of follow-up and individual genotyping for all carcinogenic types. We show that the powerful cervical cancer risk stratification following a single HPV test, particularly HPV16, lasts for greater than 10 years, and that reassurance following a negative test is particularly strong among women 30 and older.

Materials and Methods

Cohort

From April 1, 1989 to November 2, 1990, 23,702 non-pregnant women aged 16 and older, receiving routine cytologic screening in a prepaid health plan at Kaiser Permanente were invited to participate in this study (3, 14). The cohort was a demographically representative sample (mainly Caucasian) of Kaiser Permanente, which served about one-quarter of the women residing in Portland during this time. A total of 1,107 women (4.7%) refused to join the study, and 67 had no specimens taken because they were underage (Fig. 1).

Because we were studying the natural history of HPV in the context of routine cervical cytologic-screening, we excluded 1,406 participating women with previous hysterectomies and 780 women with a preenrollment medical history of HSIL or worse cytology and/or CIN2 or worse histology. In addition, 421 women lacked either interpretable cytology (85 inadequate, 208 indeterminate) or HPV testing (n = 128).

The remaining 19,921 women with satisfactory baseline cytology and HPV testing comprised the cohort (84.0% of total approached). The women were followed prospectively by routine cytology for a maximum of 214 months.

Enrollment examination

Informed consent was obtained under the prevailing institutional review board guidelines at Kaiser Permanente and the National Institutes of Health, which permitted “opt out” from “minimal-intervention studies”. Participants underwent a routine pelvic examination, at which time experienced clinicians prepared a standard, single ethanol-fixed Pap smear for each subject using an Ayre spatula and cytobrush. As the only added study component, the cervix was rinsed with 10 mL of sterile saline using a 3–1/4-inch flexible intracatheter extender attached to a syringe (15). The pooled fluid was collected from the posterior vaginal fornix and processed for HPV testing as described below.

Pathology

Pap smears were originally reported using a classification that predated the development of the Bethesda System now used; we converted these interpretations into Bethesda 2001 terminology for this study. We reclassified women with smears reported as “normal” or “benign reactive atypia” as “negative for intraepithelial lesion or malignancy (NILM, meaning normal).” Pap smears with equivocal findings reported as “severe reactive atypia, possibly dysplasia” or “possible koilocytotic or condylomatous atypia” were reclassified as “atypical squamous cells of undetermined significance” (ASC–US). Cytologic interpretations of dysplasia were reclassified as low-grade squamous intraepithelial lesions (LSIL) or high-grade SIL (HSIL), as appropriate. Glandular lesions were extremely rare and grouped with the squamous lesion of the corresponding severity.

Histologic diagnoses were converted into CIN nomenclature. Specifically, severe dysplasia and carcinoma in situ were categorized as CIN3, and moderate dysplasia was categorized as CIN2. We extensively reviewed pathology diagnoses during the first 10 years of follow-up but accepted all CIN3 and cancer thereafter because downgrades during review were very rare.

We concentrated on a stringent disease endpoint of CIN3 or cancer (≥CIN3) in order to exclude questionable precancer; however, we also report more equivocal endpoints of histologic CIN2 for the following reasons. The subtle histopathologic distinction between CIN2 and CIN3 has inadequate reproducibility, even among experts (16). Also, the usual treatment threshold in the United States is CIN2.

Human papillomavirus DNA testing

Cervicovaginal lavage specimens were refrigerated within 1 hour of collection and transported to a laboratory for processing. A 1-mL aliquot was removed and frozen at –70°C. The remaining sample was divided roughly in half, cells were pelleted by centrifugation, the supernatants were discarded, and two half-pellets were frozen at –70°C.

Each HPV test was carried out masked to any other data on that participant. We selected either liquid...
 aliquots or cell pellets for HPV testing, depending on availability and assay requirements.

The HPV typing data presented in this analysis derive primarily from MY09-MY11 PCR using AmpliTaq Gold polymerase (17), with additional testing for HPV16 and HPV18 by a noncommercial hybrid capture assay (18) (Digene Corporation, now Qiagen).

**Previous rounds of testing**

In the first round, we tested the specimens using a now obsolete test called Hybrid Capture Tube test (HCT, Digene), which proved to be inaccurate (19). Therefore, we retested all specimens using Hybrid Capture 2 (HC2), a pooled probe, microplate test at a detection threshold of 1.0 pg/mL (approximately 5000 copies). The Food and Drug Administration–approved HC2 assay is the main HPV test used in the United States; it targets all 13 carcinogenic types but also crossreacts somewhat with additional genetically related HPV types in the same species, reducing its analytic specificity (20).

To calculate risk estimates separately for HPV16 and HPV18, we retested specimens that were HC2 positive \( (n = 2,853) \) by another hybrid capture method (HC3, Digene) (18) that provided pooled-type data and also genotyped for HPV16 and HPV18. HC3 employed individual type-specific gapped RNA probes coupled with type-specific capture of DNA:RNA hybrids using immobilized DNA oligonucleotides directed to DNA in the gaps.

**Current round of testing**

Starting in 2008, we retested 6,353 specimens from the following groups: (i) Women with abnormal histology or cytology at enrollment or during follow-up, defined as histology \( \geq \) Condyloma/CIN1 or cytology \( \geq \) ASC-US; (ii) women already thought to be HPV positive by any other test including HCT, HC2, HC3, or PCR; and (iii) a random sample of 1,000 women enrolled in the cohort regardless of previous HPV testing, cytology, or colposcopy (i.e., we used a case-cohort approach). Of note, testing of the random sample (previously negative specimens that were not linked to disease) confirmed the initial results, i.e., the retesting yielded few additional new positives, none of which led to CIN2, CIN3, or cancer during follow-up.

In several previous nested case-control studies within the Portland cohort, we used MY09-MY11 PCR-based HPV typing (17, 21, 22), the first PCR-based HPV amplification and typing method. This is a well-validated research test; and we wished to have HPV typing data for the full group of carcinogenic and related HPV types. Therefore, for this final analysis we retested, using MY09-MY11 PCR-based methods, for all HPV types in the alpha-5, -6, -7, -9, and -11 species. We categorized the carcinogenic types (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68) separately from the other, genetically related types that are rarely found in invasive cancers [(HPV26, 30, 34, 35, 40, 50, 53, 55, 57, 66, 67, 69, 70, 73, 82, and 85) 23].

**Adjudicating HPV type-specific data**

Typing disagreements were frequent among multiple tests at varying times over many years on each specimen. We faced trade-offs in choosing data from various testing data typing methods (e.g., old tests of then-fresh specimens vs. retesting with optimized state-of-the-art assays of long-stored specimens). The HPV typing data presented in this analysis derive primarily from MY09-MY11 PCR carried out in 2008, but with equal consideration of earlier rounds of PCR when available, and of the noncommercial HC3 method for HPV16 and HPV18. We
developed a set of rules to adjudicate typing results for all women. We counted the number of positive results and negative results at baseline by PCR for each type (for types 16 and 18, HC3 results were treated like PCR results) and the “majority ruled”. If the number of positive and negative tests for a type were equal, we categorized the specimen as positive for that type. If there were no valid tests by PCR and/or HC3 for 16 or 18, then we checked to see if the specimen was tested by HCT or HC2. If it was tested and the result was not positive, then the official result was negative; if the specimen was positive by HCT or HC2, then the official result was “unknown type” but this last combination was very rare (n = 25).

Follow-up
Subjects in follow-up were 16 years of age or older at enrollment (median age = 34.0 years, mean = 35.8, SD = 12.7 years). During the study period, annual cytologic screening of women at Kaiser continued as part of standard clinical practice. The screening was passive, in that women were screened if they returned for cytology or other reasons; nonattendees were not aggressively pursued. The then-current standard practice guidelines for management of abnormal cytology mandated treatment of patients with CIN2 or worse (> CIN2), but health plan physicians had discretion to treat some patients with CIN1, which is a more aggressive treatment threshold than current guidelines recommend. HPV test results were not known by clinicians and were not used to direct patient management.

Statistical analysis
Separate analyses were carried out on the subgroups of 7,186 women aged <30 (median 24) and 12,735 women aged 30 years or older (median 40) at enrollment, to address age-specific screening recommendations. HPV testing is currently recommended in the United States in combination with cytology only for the older age group; therefore, a priori, we concentrated on the older group.

We were mainly interested in (i) the long-term reassurance against ≥CIN3 provided by a single negative HPV test and (ii) the absolute risks (positive predictive values or PPV) associated with a single positive test for each of the HPV types, and for HPV positivity overall compared with cytologic abnormality. We created subcohorts for women with each HPV type, for women younger than 30 and for those 30+ at enrollment. For each subcohort, we analyzed the associations of HPV result first with enrollment cytology and then with PPVs, meaning worst outcome throughout the cohort study including enrollment and follow-up (cancer, else histologic CIN3, else histologic CIN2, else cytologic HSIL, else minor cytologic or histologic abnormality, else always negative). We censored women after a cytologic result of HSIL or worse (because treatment was very likely even without histologic confirmation) or a histologic endpoint of CIN2 or worse. Because a woman could have multiple, concurrent infections due to shared or sequential sexual exposures, we further subdivided each subcohort into subsets including women with only the defining type and those who also had additional infections.

The PPVs did not take into account loss-to-follow-up, which was substantial by the end of 18 years. Therefore, observed PPVs undoubtedly underestimated total rates. Consequently, as a complementary analytic approach, we estimated cumulative probability of disease outcomes by time after enrollment for selected subcohorts defined by HPV tests at enrollment (HPV16, else other HPV types, else negative) and constructed Kaplan–Meier curves. To compare the cumulative probability, we used proportional hazard analyses to generate proportional hazards (PH) and standard log-rank tests. This method also had its limitations. Because the log-rank test assumes proportional hazards, it was only strictly appropriate when the many endpoints triggered by abnormal cytology at enrollment (which were sometimes HPV DNA negative) were excluded. Therefore, for proportional hazards analyses, we focused on prognostic values for incident cases diagnosed following the enrollment period.

Results

Enrollment HPV and cytology results
In this well-screened, low-risk population sampled in the 1980s, HPV prevalence was relatively low, as was LSIL and especially HSIL (of which 31 of 51 were HSIL, favor CIN2). The details of HPV and cytology and enrollment are shown in Supplementary Table S1 for women <30 and S2 for women 30+. For each cytologic interpretation (especially the cytologic interpretations we reclassified as ASC–US), the percentage of HPV positivity (percentages based on column data in Supplementary Table S1 and S2 with HPV types combined) was higher among the younger women than the older: 17% versus 5% for NILM, 62% versus 17% for ASC–US, 86% versus 56% for LSIL, and 96% versus 78% for HSIL, respectively. Thus, equivocal cytologic terms that we reclassified as ASC–US were used nonspecifically by pathologists at that time in Portland Kaiser, and HPV detection was generally lower in older women.

More than half (51.4%) of all infections were found concurrently with other types; 57.2% for women <30 and 37.0% for women 30+. Consistent with the independent effects of individual types, multiple infections did produce more cytologic abnormalities than single infections. However, only the minority of infections, single or multiple regardless of HPV type, resulted in concurrent cytologic abnormalities. This was particularly true in the older age group.

Follow-up
The number of screening visits gradually decreased, in large part as women left the Kaiser health plan, often due to change in family employment status and geographic mobility (B. Rush, personal communication).
Table 1. Average person-years and screening visits for each baseline HPV/cytology category

<table>
<thead>
<tr>
<th>BL group</th>
<th># Women</th>
<th>Median (IQR)* person-years</th>
<th>Median (IQR)* screening visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>All women &lt; 30 at baseline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV– NILM</td>
<td>5581</td>
<td>4.6 (0.8–13.1)</td>
<td>4 (2–8)</td>
</tr>
<tr>
<td>HPV– ASC/LSIL</td>
<td>155</td>
<td>2.5 (0.6–9.9)</td>
<td>4 (2–7)</td>
</tr>
<tr>
<td>HPV+ NILM</td>
<td>948</td>
<td>3.8 (1.0–11.4)</td>
<td>4 (2–7)</td>
</tr>
<tr>
<td>HPV+ ASC/LSIL</td>
<td>343</td>
<td>2.5 (0.3–10.7)</td>
<td>3 (2–7)</td>
</tr>
<tr>
<td>Women 30+ at baseline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPV– NILM</td>
<td>11462</td>
<td>10.7 (3.3–15.2)</td>
<td>6 (3–10)</td>
</tr>
<tr>
<td>HPV– ASC/LSIL</td>
<td>393</td>
<td>10.3 (2.6–14.9)</td>
<td>7 (4–12)</td>
</tr>
<tr>
<td>HPV+ NILM</td>
<td>459</td>
<td>9.1 (2.0–14.6)</td>
<td>6 (2–10)</td>
</tr>
<tr>
<td>HPV+ ASC/LSIL</td>
<td>120</td>
<td>4.8 (0.5–13.5)</td>
<td>4.5 (2–9)</td>
</tr>
</tbody>
</table>

NOTE: Person years defined as the time until HSIL+ or the end of follow-up.
*aInter-quartile range.

The average length and intensity of follow-up on a per woman basis are outlined in Table 1. Women <30 years old had a median length of follow-up of 4.3 years, inter-quartile range (IQR) of 0.7–12.8 years, and a median number of screening visits of 4, IQR of 2–7. Our research focus was on women 30+ years old, who tended to stay in the Kaiser health plan. They had a considerably longer median length of follow-up of 10.5 years, IQR of 3.1–15.1 years, and a median number of screening visits of 6, IQR of 3–10. The probability that at least one tissue specimen was taken during follow-up varied by enrollment cytology status: 52.9% for ASC-US, 87.9% for LSIL, and 98.0% for HSIL. Within each age group, women without screening abnormalities were followed at least as long and had as many screening visits as those with abnormalities, ruling out any spurious positive association due to differential follow-up.

In Fig. 2 (with full details given in Supplementary Tables S3 and S4), we show type-specific HPV results at enrollment compared with worst disease observed in follow-up. As in Supplementary Table S1 and S2, in Supplementary Table S3 and S4, we separate single from multiple infections, and women <30 from those 30+. In total, 189 (1.0%) women were diagnosed with CIN3 and 15 (0.1%) developed invasive carcinoma, despite the availability of screening as part of the prepaid health-maintenance plan.

Over the course of follow-up, HPV16 clearly produced the most cases of ≥CIN3, followed by HPV31 and HPV18. In both age groups, multiple infections were associated with significantly more cytologic-histologic abnormalities (Supplementary Table S3 and S4), which would be expected simply as a sum of individual risks.

Also shown in Supplementary Table S3 and S4 (columns), most cases of ≥CIN3 found in women <30 during follow-up were HPV positive at enrollment (78 of 115 = 68%); for women 30+, fewer than half of cases 41/89=46% were HPV positive. The difference remained even when we restricted to the first five years of follow-up (85.9% vs. 58.9%, respectively), to take into account the shorter follow-up among younger women.

As shown in Figs. 3a and 3b, a single negative test for HPV at enrollment predicted an extremely low cumulative probability of ≥CIN3, especially among women 30+ (0.7%) compared with women <30 (1.8%).

Also shown in Figs. 3a and 3b, we further explored the risks following detection of HPV16, other carcinogenic HPV types (as a group, in the absence of HPV16), and HPV negativity at enrollment. Due to small number of other types, we were able to provide reliable type-specific estimates for HPV16 only. Figures 3a and 3b present the Kaplan–Meier estimates of cumulative probability of ≥CIN3, for <30 and 30+, respectively (excluding prevalent cases found at enrollment). For both age groups, comparing HPV16 with other types, and other types to HPV negativity, the log-rank tests yielded significant P-values (P = 0.04 for HPV16 vs. other carcinogenic types in women 30 and above, P < 0.001 for all other comparisons). The hazard ratios (HR) of HPV16 compared with other carcinogenic types and to HPV-negativity in the important older group were 2.7 (95% CI = 1.0–7.3) and 6.2 (95% CI = 3.0–12.6), respectively. Cases associated with HPV16 tended to occur early. In an ancillary analysis (data not shown) among women 40 and above, no new cases of ≥CIN3 were found for enrollment HPV16-positive women after only 4 years, although there were not many women at risk.

Figs. 4a and 4b show the cumulative probability of CIN2 in women <30 and 30+, respectively. For this less severe outcome that is a mixture of acute infection and true precancer, risks continued to increase longer and higher than for ≥CIN3, for women with HPV16, other carcinogenic types, or no carcinogenic types at enrollment. A negative HPV test at enrollment did not predict as low a cumulative incidence rate for CIN2 for younger women (4.1%) as for older women (1.3%). For women...
<30, HPV16 predicted higher risk than other carcinogenic types (HR = 1.9, 95% CI = 1.3–3.0) but for women 30+, other carcinogenic types predicted nonsignificantly higher risk than HPV16 (HR = 2.1, 95% CI = 0.8–6.1).

To address current interest in cotesting with HPV and cytology, we examined the cumulative probability of histologic ≥CIN2 for combinations of test results in Figs. 5a and 5b, for ages <30 and 30+, respectively. For this clinical question, we chose the histologic disease endpoint of ≥CIN2 to reduce any interpretative problem due to treatment of CIN2 before CIN3 and cancer could develop. In both age groups, HPV status (positive vs. negative) was a more important determinant of risk (HR = 7.1, 95% CI = 4.6–10.9 and 8.5, 95% CI = 4.8–15.1 for ages <30 and 30+, respectively) than cytological abnormality of ASC–US/LSIL versus NILM (HR = 1.5, 95% CI = 0.7–3.3 and 2.9, 1.2–6.6 for ages <30 and 30+, respectively). Risk stratification in the critical 30+ age group was achieved mainly by HPV test result. However, HPV-negative NILM predicted an extremely low cumulative probability of CIN2+ through the end of follow-up for ages 30+ (1.1%).

Discussion

Our data clearly show the extraordinarily low risk of CIN3 or cancer following a negative HPV test result, in follow-up lasting up to 15+ years. These data corroborate and extend those of other long-term cohorts (5, 6, 24) and our own previous publications on the Portland cohort (14). Cervical cancer almost always arises from detectable persistence of HPV infections (10). A negative test provides excellent "negative predictive value", i.e., confidence that the "clock is still set at zero" in the typically decades-long natural history of cervical cancer. The risk following a negative test is increasingly low as women age (5) and become less likely to acquire new infections, which, even if acquired, would tend to clear as in younger women (25).

Except in low-resource regions, cervical cancer–screening is a repeated process. Still, as recently suggested by the authors of several cohort studies and randomized clinical trials (4, 26, 27), the strong long-term reassurance from a single negative HPV test should inform clinical practice. We are not suggesting that screening in high-resource regions be extended to the length of follow-up of the Portland Cohort. However, rescreening at three years for women 30 and older after negative HPV and cytology cotesting, as commonly recommended (11, 12, 28), is possibly too soon; it is concerning that many clinicians are repeating HPV tests at even shorter intervals (28). Even at older ages, HPV detected within three years after a negative test, like any new infection, will rarely progress to > CIN3 let alone cancer (25).
Accumulated evidence and conclusions reported recently from Europe (4) support longer screening intervals. In a pooled study of approximately 25,000 women enrolled in seven European cohorts, the cumulative incidence rate of ≥CIN3 after a negative HPV test at baseline was 0.3% (95% CI of 0.1% to 0.4%) over six years of follow-up. This is comparable to the cumulative percentage of 0.7% (95% CI of 0.5% to 0.9%) that we observed in HPV-negative women 30 and older in longer follow-up.

Our results might be most important for low-resource regions where frequent screening is not the norm. The low risk following HPV-negative screening frees health planners in low-resource regions from a perceived ethical requirement to emulate the Western model of frequently

Figure 3. Cumulative probability of ≥CIN3 stratified by HPV test results. The curves exclude enrollment cases of cytologic HSIL or histologic ≥CIN2, to permit calculation of hazard ratios using proportional hazards methods. The numbers of women still at risk for each HPV test stratum are listed below the graph. The cumulative risk (cumulative probability) is plotted using Kaplan–Meier methods that take into account censoring due to cytologic HSIL or histologic ≥CIN2 or losses to follow-up. The PPV includes all cases observed combining enrollment and follow-up, but does not take into account censoring. A, women <30: HPV16 infection, cumulative probability = 14.6 (95% CI = 10.0–20.9) and PPV = 11.1 (8.1–14.0); other carcinogenic types, cumulative probability = 7.0 (4.2–11.4); PPV = 3.2 (2.0–4.3); no carcinogenic types, cumulative probability = 1.8 (1.2–2.5) and PPV = 0.7 (0.4–0.9). B, women 30+: HPV16 infection, cumulative probability = 8.6 (95% CI = 4.1–17.2) and PPV = 13.8 (8.2–19.4); other carcinogenic types, cumulative probability = 3.1 (1.6–6.1) and PPV = 3.7 (2.0–5.4); no carcinogenic types, cumulative probability = 0.7 (0.5–0.9) and PPV = 0.4 (0.3–0.5).
repeated cytology screening and justifies the use of HPV testing at extended intervals using new low-cost HPV tests (29).

Compared with other carcinogenic types, higher risk of CIN3 remained for a decade after an HPV16-positive test (5, 6, 30). Despite the large study population, we had less power to study less prevalent carcinogenic HPV although our data did (without statistical significance) corroborate elevated risk following infection with HPV18 or HPV31 (24), and little elevation in risk for other carcinogenic types (5). As shown by other cohorts, the CIN3 lesions detected by HPV testing especially in women 30 and older are often (but not always) persistent and therefore clinically significant (5, 30).

Figure 4. Cumulative probability of CIN2 stratified by HPV test results. The curves exclude enrollment cases of cytologic HSIL or histologic CIN2+, to permit calculation of hazard ratios using proportional hazard methods. The numbers of women still at risk for each HPV test stratum are listed below the graph. The cumulative risk (cumulative probability) is plotted using Kaplan-Meier methods that take into account censoring due to cytologic HSIL or histologic CIN2+ or losses to follow-up. The PPV includes all cases observed combining enrollment and follow-up, but does not take into account censoring. A, women <30: HPV16 infection, cumulative probability = 10.5 (95% CI = 7.0–15.7) and PPV = 9.7 (95% CI = 7.0–12.5); Other carcinogenic types, cumulative probability = 5.9 (95% CI = 3.9–8.9), PPV = 4.9 (95% CI = 3.4–6.3); no carcinogenic types, cumulative probability = 3.1 (95% CI = 2.4–3.9) and PPV = 1.4 (95% CI = 1.1–1.6). B, women 30+: HPV16 infection, cumulative probability = 4.2 (95% CI = 1.0–17.3) and PPV = 11.0 (95% CI = 5.9–16.1); Other carcinogenic types, cumulative probability = 6.1 (95% CI = 3.7–9.8), PPV = 6.7 (95% CI = 4.4–9.0); no carcinogenic types, cumulative probability = 0.6 (95% CI = 0.4–0.9) and PPV = 0.4 (95% CI = 0.3–0.5).
Figure 5. Cumulative probability of ≥CIN2 stratified by HPV test and cytology results. The curves exclude enrollment cases of ≥CIN2, to permit calculation of hazard ratios using proportional hazards methods. The numbers of women still at risk for each HPV test stratum are listed below the graph. The cumulative probability is plotted using Kaplan–Meier methods that take into account censoring due to cytologic HSIL or histologic ≥CIN2 or losses to follow-up. The PPV includes all cases observed combining enrollment and follow-up, but do not take into account censoring.

A, women <30: HPV+ ASC/LSIL, cumulative probability = 8.9 (95% CI = 5.2–15.0) and PPV = 14.9 (11.1–18.6); HPV− ASC/LSIL, cumulative probability = 6.4 (2.2–17.7), PPV = 5.8 (2.1–9.5); HPV+ NILM, cumulative probability = 15.2 (11.7–19.6) and PPV = 7.2 (5.5–8.8); HPV− NILM, cumulative probability = 3.9 (3.1–4.9) and PPV = 1.6 (1.3–1.9).

B, women 30+: HPV+ ASC/LSIL, cumulative probability = 7.5 (95% CI = 3.4–16) and PPV = 24.2 (16.5–31.8); HPV− ASC/LSIL, cumulative probability = 1.4 (0.5–3.7), PPV = 2.3 (0.8–3.8); HPV+ NILM, cumulative probability = 8.9 (5.9–13.5) and PPV = 6.1 (3.9–8.3); HPV− NILM, cumulative probability = 1.1 (0.9–1.4) and PPV = 0.7 (0.5–0.8).
27). We need health-decision analyses to determine if distinguishing HPV16 or other especially carcinogenic types in HPV test kits is cost-effective, and how we might manage different levels of risk once accurate typing is widely available clinically. In this study, we found no support for separate typing for the majority of the carcinogenic types although a number of cases caused by most types were limited.

The comparison of HPV testing with cytology indicated that HPV status provides longer and stronger risk stratification than cytology, supporting the conclusions of previous short- and long-term comparisons (4–6, 24, 26, 27, 31–34). Finding cytologic ASC–US/LSIL added little long-term risk stratification to that provided by HPV testing. In the European pooled analysis (4) after six years the risk of CIN3 among initially HPV-negative women was lower than after three years among cytology-negative women, suggesting again that switching to HPV testing should result in extended screening intervals. In making this comparison, it is important to note that cytology was not designed for single-time, long-term use but rather as part of a screening routine to be repeated at fixed intervals. We also note that our specific comparison was limited by the nonspecificity, particularly among older women, of equivocal cytologic interpretations that we classified as ASC–US. With more contemporary cytology, many would now be called negative. As another limitation, some of the early-elevated risk of temporary cytology, many would now be called negative. As another limitation, some of the early-elevated risk of diagnosis of ≥CIN3 in our study was due to diagnostic efforts triggered by cytologic abnormalities.

As more minor points, we found that the importance of HPV16 was attenuated for CIN2, which can be caused in larger part by many HPV types and which, even more than early CIN3, frequently regresses spontaneously (5). We also confirmed at enrollment that most HPV infections, especially single infections at older ages, do not cause concurrent cytologic abnormalities (35).

Additional study limitations influenced the results, if not the conclusions. We collected cells with a cervicovaginal lavage, which proved insensitive among older women. The lack of sensitivity of the cervical-lavage collection method among older women might reflect that their cervical transformation zones were more likely to be within the cervical os (29), even though the median age of women with ≥CIN3 in the older age group was only 37.5 and HPV-negative case women were not significantly older than HPV-positive ones. It could also reflect the higher prevalence of HPV infections in the younger age group.

This lowered sensitivity led to decreased negative predictive value among older women. There is a possibility that this decrease could be greater for particular HPV types more prone to infect the cervix as opposed to the vagina (36).

We tested specimens multiple times, until some had no available satisfactory aliquots. Thus, we had to depend on a combination of the test results, which is a major difference from routine clinical practice and a key limitation of this long-term cohort project. Most specimens negative by HCT and HC2, in women never diagnosed with any abnormality, were never tested by PCR. A small percentage of such specimens (as seen in the random subcohort, data not shown) were PCR positive; in absolute numbers, these missed infections could have accounted for a sizable minority of all infections in the cohort. Therefore, because they never led to any cytologic or histologic abnormal diagnosis, the PPV estimates we presented might be somewhat elevated; a more sensitive testing scheme would perhaps lower the risk (i.e., the PPV or the cumulative probability) after a positive test, but strengthen the reassurance from a negative test.

Women, especially younger ones, tended to leave the Kaiser health plan over the years, leading to a small number and unstable estimates for some analyses. Reassuringly, our results corroborate the recent report of a cohort study among women in Copenhagen younger than 30 years at enrollment (24).

We had too few events in the low risk, highly screened Kaiser cohort to assess the carcinogenicity of HPV types other than HPV16. Despite the large size of the Portland Cohort, reduced number of women infected with any specific type returned after 10 years. With limited precision, we did observe expected elevated cumulative risks of ≥CIN3 for HPV31 and HPV18. The risk of invasive cancer in this cohort was clearly linked to only a few HPV types.

Finally, we did not have repeat, longitudinal HPV testing of the kind that have shown it is HPV persistence that leads to CIN3 and cancer (24, 32). In particular, we did not have genotyping of the lesional tissue at the time of diagnosis preventing exact attribution of each case to a specific HPV type (37, 38). It is important to realize that diagnosis of ≥CIN3 following detection of an HPV infection does not prove causality, especially during long-term follow-up. Even when a single HPV type is present at enrollment, our "single snapshot" study cannot rule out that the detected type cleared, and another subsequently acquired type eventually progressed to precancer. We believe that the continuing accumulation of CIN2 especially among younger women indeed suggests that some endpoints in our study arose, perhaps very quickly, from infections acquired subsequent to specimen collection and are not related to enrollment infections. If so, the long-term low rates of >CIN3 following a negative HPV test are even more remarkable. Although we cannot be absolutely certain of the causal HPV genotype for follow-up >CIN3 based on the baseline measurement of HPV, we note the strong predictive value of HPV16 and that roughly half of all >CIN3 as expected were HPV16 positive at baseline. Therefore, we surmise that for the vast majority of cases found during follow-up, the causal HPV genotype was present at baseline.

In conclusion, long-term data from women spanning all adult ages in the Portland Cohort Study strongly support the negative predictive value of HPV testing.
A double negative result of HPV-negative NILM predicts an extremely low risk of ≥CIN2 in the subsequent decade. The risk decreases even further as women age. Logically, screening intervals should widen further as successive double negative screenings are accrued. A remaining challenge is to determine how to manage HPV positive women, particularly those with HPV16 or other especially high-risk types, if a switch to HPV testing for primary screening includes HPV typing.

Disclosure of Potential Conflicts of Interest
Attila T. Lorincz used to work for Digene Diagnostics. He still consults for Qiagen/Digene.

References


Correction: A Long-term Prospective Study of Type-Specific Human Papillomavirus Infection and Risk of Cervical Neoplasia among 20,000 Women in the Portland Kaiser Cohort Study

In this article (Cancer Epidemiol Biomarkers Prev 2011;20:1398–409), which was published in the July 2011 issue of Cancer Epidemiology, Biomarkers & Prevention (1), the authors misclassified the histology status of 33 cases that should have been categorized as CIN3. Thirty-two cases of carcinoma in situ (CIN) were mistakenly included as cancer rather than CIN3. One other case of CIN3 was missed altogether and coded as <CIN2. As a result, the authors observed 15 cases of invasive cancer (not 47) and 189 cases of CIN3 (not 156). The conclusions of the analysis were based on a combined category of CIN3/cancer and, thus, the results in the main figures are changed minimally, by the addition of 1 case of 189. However, as a substantial modification, there is no longer adequate statistical power to examine invasive cancer as a separate category. A corrected Fig. 2 in which the combined CIN3/cancer category is used is supplied below. A typographical error was also identified in Supplementary Table 1: The count under NILM for type 73 multiple infections should be 23 rather than 2. The authors apologize for the errors. In the interest of clarity, the publisher has made revised versions of the article and the supplementary tables available online.

Figure 2. Type-specific HPV results at enrollment, and positive predictive value of worst disease observed in follow-up. A, women <30. B, women 30+.
Reference


Published OnlineFirst June 20, 2012.
doi: 10.1158/1055-9965.EPI-12-0618
©2012 American Association for Cancer Research.
A Long-term Prospective Study of Type-Specific Human Papillomavirus Infection and Risk of Cervical Neoplasia Among 20,000 Women in the Portland Kaiser Cohort Study

Mark Schiffman, Andrew G. Glass, Nicolas Wentzensen, et al.

Cancer Epidemiol Biomarkers Prev 2011;20:1398-1409. Published OnlineFirst May 20, 2011.

Updated version
Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-11-0206

Supplementary Material
Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2011/05/20/1055-9965.EPI-11-0206.DC1

Cited articles
This article cites 36 articles, 6 of which you can access for free at:
http://cebp.aacrjournals.org/content/20/7/1398.full#ref-list-1

Citing articles
This article has been cited by 8 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/20/7/1398.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.