Invited Commentary

Invited Commentary: Is Monitoring of Human Papillomavirus Infection for Viral Persistence Ready for Use in Cervical Cancer Screening?

Philip E. Castle

From the Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.

Received for publication November 28, 2007; accepted for publication January 17, 2008.

Persistent cervical infections by approximately 15 carcinogenic genotypes of human papillomavirus (HPV) cause virtually all cases of cervical cancer and its immediate precancerous precursor, cervical intraepithelial neoplasia grade 3 or carcinoma in situ. As is shown in a meta-analysis by Koshiol et al. (Am J Epidemiol 2008;168:123–137), detection of carcinogenic HPV viral persistence could be used to identify women at the greatest risk of cervical precancer. Specifically, women who have carcinogenic HPV infection that persists for at least 1 year versus those whose infections clear are at significantly elevated risk of having or developing cervical precancer. However, before detection of HPV persistence can be used in cervical cancer screening, several considerations need to be addressed: 1) validation and Food and Drug Administration approval of a reliable HPV genotyping test, 2) rational clinical algorithms based on risk of precancer and cancer for the clinical management of HPV persistence, 3) clinician and patient acceptability of monitoring of HPV infections (including not responding excessively to the first positive HPV test and waiting 1–2 years for infections to either persist or resolve), and 4) patient compliance with recommended follow-up. Investigators will need to address these and other key issues in order to realize the potential utility of HPV viral monitoring for improving the accuracy of cervical cancer screening.

human papillomavirus 16; human papillomavirus 18; longitudinal studies; papillomavirus infections; uterine cervical neoplasms

Abbreviations: CIN, cervical intraepithelial neoplasia; HPV, human papillomavirus; HSIL, high-grade squamous intraepithelial lesions.

Based on the central role of persistent carcinogenic human papillomavirus (HPV) in cervical carcinogenesis, testing for carcinogenic HPV was recently introduced into cervical cancer screening (1). (Note that in this context, “HPV testing” always refers to the detection of carcinogenic HPV genotypes only; testing for noncarcinogenic HPV genotypes has no clinical utility.) HPV testing has proven greater sensitivity than cytologic screening (Papanicolaou smears) for detection of cervical precancer (Papanicolaou smears) or detection of cervical precancer (cervical intraepithelial neoplasia grade 3 (CIN3)) and cervical cancer (≥CIN3) (2–6) and greater reliability (7, 8). HPV testing is now commonly used in the United States to triage equivocal cytologic findings for colposcopic referral. Co-testing with HPV and cytology is also approved for primary (routine) cervical cancer screening of women aged 30 years or more (9). Women aged 30 years or older who test HPV- and cytology-negative are at extremely low risk of incipient precancer and cancer (≥CIN3) over periods of 10 years or more (10, 11). Therefore, screening intervals in these women can be extended to 3 years in the United States to make co-testing cost-effective (12). Use of HPV testing in primary screening is only recommended for women aged 30 years or older, because these women are typically past the peak age of self-limited infections, which are very common in young women (13, 14). Thus, the positive predictive value of ≥CIN3 is higher in women aged 30 years or more than in younger women (15,
In general, concurrently performed cytologic screening adds little to the sensitivity and negative predictive value of HPV testing (e.g., sensitivity of 97.4 percent for HPV testing alone vs. 100 percent for HPV testing and Pap smears using a threshold of atypical squamous cells of undetermined significance or worse (3)). In a publication based on a meeting of experts, the International Agency for Research on Cancer recently concluded that HPV testing is an acceptable alternative to Pap smears/cervical cytology for cervical cancer screening (17). However, despite its greater sensitivity and overall accuracy, the enthusiasm for using HPV testing in primary screening has been tempered by its somewhat poorer positive predictive value in comparison with cytologic analysis (e.g., 7.0 percent for HPV testing vs. 8.7 percent for Pap smears, using a threshold of atypical squamous cells of undetermined significance or worse (3)). Even at older ages, the prevalence of self-limited infection can reach 10 percent (vs. approximately 5 percent for cervical cytology), with only a minority of these women being at risk of CIN3. A viable strategy for managing the cases of HPV-positive women, specifically identifying the subset at risk of CIN3, would accelerate the adoption of HPV testing into primary cervical cancer screening.

Several solutions have been suggested. One is to use cytology as a reflex test for HPV-positive women, because it is more specific for CIN3 (18), but the morphologic criteria for the threshold of positive (abnormal) cytology would need to be adapted (if possible) to avoid losses of sensitivity, as observed by Mayrand et al. (3) using current criteria. Another possibility is HPV genotyping that could target the most carcinogenic of approximately 15 carcinogenic genotypes and could permit the tracking of viral persistence. Already, there is evidence that separate detection of HPV16 and HPV18 in cytologically normal women aged 30 years or more might be a useful risk “stratifier” (9, 19). Those women who test positive for HPV16 or HPV18 might benefit from immediate colposcopy, while those who test negative for both could wait a year before being rescreened (9, 19). Given the fundamental role of persistent HPV in cervical carcinogenesis, its reliable measurement could increase the accuracy of cervical cancer screening, by aiding in further distinguishing HPV infections that pose a risk from those which resolve on their own.

Koshiol et al. (20) conducted a systematic review and meta-analysis of the relation between detection of HPV viral persistence and the risk of precancerous lesions, defined in their analysis as histologic CIN2/3 or cytologic high-grade squamous intraepithelial lesions (HSIL). Their key findings regarding HPV persistence were: 1) it was strongly linked to the detection of precursor; 2) it was more strongly linked with precursor and cancer than with equivocal or low-grade lesions; 3) it was less strongly linked to precursor when women with transient infection were used as a referent group (i.e., women who were HPV-positive at the indicator visit but were negative at a subsequent follow-up visit) than when other referent groups were used (e.g., mixed HPV-positive and -negative or HPV-negative women alone); 4) longer persistence, as detected by absolute duration or the time interval between measurements, was more strongly linked to precursor and cancer; and 5) HPV genotype-specific persistence was no more related to risk of precursor than was repeatedly testing positive for HPV in the aggregate.

With such a heterogeneous set of studies employing different tools, epidemiologic study designs, and limited numbers of outcomes, Koshiol et al.’s analysis naturally had several limitations. Most notable was the use of different endpoints; some studies used histologic endpoints that included CIN2, while other studies used cytologic HSIL to define disease. CIN2 is the standard clinical threshold for treatment, but not all CIN2 is cervical precursor. The CIN2 diagnostic category can include women with nothing more than acute HPV infections, including those caused by noncarcinogenic HPV (21). HSIL cytology is generally a specific but insensitive cytologic indicator of cervical precursor. In addition, we must acknowledge that colposcopy, visualization of the cervix, and diagnostic biopsy are insensitive for the detection of cervical precursor (22–24), especially in well-screened populations in which most large and obvious precancerous lesions have already been detected and treated. Therefore, viral persistence may reflect a concurrent, visually unapparent precancerous lesion as well as indicate those women at risk for developing precursor. All of these limitations probably resulted in underestimation of risk relations in the analysis by Koshiol et al.

While monitoring of HPV infections for persistence versus clearance provides excellent risk stratification in study cohorts, it is fair to ask is whether its use is currently feasible in realistic cervical cancer screening programs.

Is the relation between HPV persistence (vs. transience) and risk of CIN3 sufficiently robust to be clinically useful?

In the paper by Koshiol et al. (20), the summary-estimate relative risk for cervical precursor for any persistent versus transient HPV infection was 14.7 (range, 5.4–119.1). Using the benchmarks of odds ratios/relative risks of approximately 20–25 or greater for a clinically useful biomarker established by Pepe et al. (25), considering any viral persistence may be insufficient for clinical use. However, with data restricted to those studies that examined the risk for persistence of 1 year or more, the summary-estimate relative risk increased to 42.9, suggesting that monitoring of HPV viral persistence for at least 1 year might be useful in patient management.

There is some evidence to support a minimum 1-year threshold for clinically meaningful persistence (26). At least 50 percent of all HPV infections clear within 1 year (27–29). Follow-up studies with repeat measurements taken over a period of 1 year (30) or 2 years (11) begin to distinguish infections and associated lesions that pose greater risk from those posing lower risk. Women with 1-year HPV genotype-specific persistence remain at elevated risk of cervical precursor and cancer (24 percent for CIN2 or worse and 15 percent for CIN3 or worse) for several years after an index colposcopic evaluation has failed to detect disease (4).

What is the appropriate clinical response to detection of HPV persistence or clearance?

The objective of any test, including monitoring of HPV infections, is to identify women at greater or lesser risk,
leading to rational clinical algorithms based on risk (31). As new risk stratification tools like HPV genotyping and monitoring of HPV infections become available, the clinical response to a positive or negative result (for example, persistence vs. transient infection) should be predicated on and standardized to previous established management guidelines for that risk level. While professional clinical societies and cancer prevention experts will need to establish the thresholds of risk and the appropriate clinical responses, current clinical algorithms based on cytology and colposcopy, such as those recommended by the American Society for Colposcopy and Cervical Pathology (9, 32), can provide some guidance for risk-based algorithms (31). For example, women with a <2 percent risk of cervical pre-cancer and cancer over a typical screening interval (1–3 years) (e.g., negative cytology and HPV-negative equivocal cytology) can be sufficiently reassured that a return to routine screening is acceptable. Risks of 2–<10 percent (e.g., colposcopy and CIN1 biopsy) warrant increased surveillance. A ≥10 percent risk of ≥CIN3, like low-grade squamous intra-epithelial lesion cytology (9), should merit colposcopy. By analogy, women with at least 1-year HPV persistence probably merit colposcopy for as long as the infection persists (4).

New risk-stratification tools like HPV testing and HPV genotyping should be considered in their totality, without “cherry-picking” of particularly palatable parts. For example, the addition of HPV testing for triage of equivocal cytology was accepted quickly in that HPV-positive women are sent immediately to colposcopy. There has been more reluctance to treat HPV-negative women with equivocal cytology as negative and safe (9), despite the proven low risk in these women (33). For monitoring of HPV persistence, even greater potential for misuse exists. If there is an unwillingness to manage less aggressively those women without HPV persistence, strategies for monitoring HPV infections will produce mainly increased costs and excessive treatment, which negatively affects reproductive outcomes (34). Misuse of HPV testing has already been reported (35, 36), and introduction of HPV genotyping for viral monitoring into clinical practice could exacerbate the situation.

How could detection of HPV viral persistence be incorporated into clinical practice?

Based on a risk model, several possible algorithms could be considered, and one based on HPV as the primary screening test is presented in figure 1. Women who tested negative for HPV would be reassured against cervical cancer and would not need to undergo the next round of screening for 3–5 years. Women who tested positive for HPV either would be managed according to the HPV genotype(s) detected, if an HPV genotyping test was used as the primary test, or their residual specimen would be reflex-tested for the HPV genotype(s) present. Women who tested positive for carcinogenic HPV genotypes other than HPV16 and HPV18 would return in 1 year and be retested, with persistence triggering referral to colposcopy. Cytologic analysis might also be conducted as an adjunct for increased safety if a woman’s past history was unknown (e.g., if she had joined a new health plan and her screening history was unavailable) and subsequently phased out as future risk was determined through monitoring of the outcome of new HPV infections.

Because improved accuracy in screening may also be accompanied by greater complexity, with increasing numbers of combinations of test results, simplified reporting of results by grouping of those results that represented similar risks and would be managed identically might be useful. That is, it might be sufficient to report “persistent carcinogenic HPV” for 1-year persistence of carcinogenic genotypes other than HPV16 and HPV18 (and perhaps HPV45) rather than reporting persistent HPV31, HPV35, or HPV58, as their risks are similar. The development and use of a risk calculator, a nomogram, that included important modifiers of risk and management recommendations from professional clinical societies like the American Society for Colposcopy and Cervical Pathology might be helpful.

What are some of the practical considerations involved in monitoring HPV infections for persistence or clearance?

Monitoring of HPV infections for persistence or clearance involves several practical considerations. First, well-validated tests must be available. Currently, there are no tests for HPV genotyping with Food and Drug Administration approval, which is required for their use in cervical cancer screening and serves as a benchmark for validation (37). Second, measurement of viral persistence must be cost-effective. The objective of measuring viral persistence, as we discussed above, is to stratify risks and match the appropriate and established clinical care to the risk. Specifically, unless a new strategy leads to a “sliding scale” for the intensity of clinical management (from extended screening intervals for the lowest risk to colposcopy and treatment for the highest risk) based on population risk, HPV genotyping will only increase the costs of screening without patient benefit. Third, and related to cost, the use of HPV genotyping implies both clinician and patient acceptability. Clinicians must be willing to recommend and follow through with the appropriate management based on the risk linked to the outcome. For example, doctors and patients must be willing to wait for a year or two following a positive HPV test to determine whether the infection will persist or resolve. Fourth, detection of HPV viral persistence not only requires acceptance, it requires compliance with follow-up; that is, viral monitoring may not be applicable to populations of women with poor follow-up.

In addition to irregularities of routine screening or poor compliance with recommendations, women may move or switch medical programs or primary clinicians. In many circumstances, data on clinical history and test results are not linked. Therefore, given an HPV-positive test result, clinicians will be uncertain as to how long the patient has harbored the infection. Again, cytology screening may be a useful addition when the patient’s history is unknown (figure 1).

Do we need to detect all HPV genotypes individually?

One of the more intriguing findings by Koskiol et al. (20) was that measurement of HPV genotype-specific viral
persistence was no more strongly linked to ≥CIN3 than was repeated positive testing for any carcinogenic HPV genotype without discrimination as to which HPV genotype(s) was present. Kjaer et al. (11) demonstrated that testing for any carcinogenic HPV genotype twice over a 2-year interval effectively stratified cytologically negative women into higher-risk (positive/positive), medium-risk (positive/negative or negative/positive), and lower-risk (negative/negative) populations over an 11-year follow-up period (figure 2). It may therefore be sufficient for clinical applications to individually detect the most risky HPV genotypes like HPV16 and HPV18, for which there is some evidence (19) of utility, and detect the remaining HPV genotypes in aggregate rather than individually (i.e., repeatedly testing positive for any other carcinogenic HPV vs. carcinogenic HPV genotype-specific persistence). This would simplify communication of test results from clinical laboratories to clinicians while potentially providing robust risk stratification. More data are needed to examine the best format (partial vs. complete HPV genotyping) for clinical applications.

One caveat is that many of these data were based on earlier, research-oriented HPV genotyping tests that may not be as analytically sensitive as current, commercialized HPV genotyping tests (38). Nevertheless, it might be inferred from these data that HPV genotyping systems may not be sufficiently robust to identify women with persistent HPV infections more accurately than the Food and Drug Administration-approved test (the Hybrid Capture 2 assay (Digene Corporation, Gaithersburg, Maryland)), which detects all carcinogenic genotypes as a pool. Most current HPV genotyping assays rely on polymerase chain reaction amplification of a 60- to 450-base-pair region of the HPV L1 gene using consensus primers (i.e., targeting conserved regions of the gene). However, a transient coinfection of any HPV genotype, carcinogenic or not, can compete for polymerase chain reaction primers (39) and cause a false-negative test for the clinically relevant genotype. Of the current HPV genotyping systems being developed and commercialized, there are no data on their reliability to repeatedly measure individual genotypes over time to establish risk of

FIGURE 1. A proposed clinical algorithm for using carcinogenic human papillomavirus (HPV) testing as the primary screening test and HPV genotyping to triage carcinogenic HPV-positive results and monitor viral persistence as a risk factor for cervical precancer and cancer. If the molecular HPV test used for screening (e.g., Hybrid Capture 2 (Digene Corporation, Gaithersburg, Maryland)) does not provide information on the individual HPV genotype(s) present, only on whether a woman is or is not positive for carcinogenic HPV, a second test would be necessary to determine which HPV genotype(s) are present. Cytologic analysis may be a useful adjunct for women whose screening history is unknown, with women having cytologically detected high-grade squamous intraepithelial lesions (HSIL) being referred to colposcopy immediately. Depending on the exact format of the HPV testing (full or partial HPV genotyping), the most robust use of HPV genotyping, and the associated risks of cervical intraepithelial neoplasia grade 3 or cervical cancer (≥CIN3), women who repeatedly test positive for the other carcinogenic HPV genotypes (individually or in aggregate) may need increased surveillance or colposcopy (31). Following a carcinogenic HPV-negative test, women may be rescreened at an extended interval of 3 or more years, depending on the acceptable risk of ≥CIN3 (and, more specifically, the reassurance against cancer), which may depend on the previous test results.
What additional questions need to be addressed?

Some additional questions need to be addressed. First, a key clinical question is how long to follow women with HPV infections. Koshiol et al. (20) stratified the risk on measured HPV persistence of less than 12 months versus 12 or more months. It would be useful to explore more finely the relation between duration of persistence and risk of cervical precancer and cancer. Longer-enduring infections will be more strongly linked to the risk of ≥CIN3, but any gain in positive predictive value obtained by monitoring HPV infections longer must be weighed against any decreased reassurance against cancer developing while waiting for transient infections to clear. Second, age is likely to be an important modifier of risk, and at what age might HPV viral monitoring be useful? The cutoff age of 30 years for HPV testing was chosen as a surrogate for viral persistence; that is, a greater proportion of women who are aged 30 years or older are past the age peak of self-limited infections, and when they test HPV-positive they are more likely to have worrisome, persistent HPV infections than are women younger than 30. It is possible that more accurate identification of women with HPV persistence will permit HPV viral monitoring in younger women (ages 25–29 years) in order to find early precancerous lesions, as long as the clinical response upon first detection of HPV is not excessive.

Third, more data are needed on the persistence of individual HPV genotypes and the risk of precancer. In particular, it would be useful to know how the risks for persistent HPV16 and HPV18 infections differ from those for other HPV genotypes, given that HPV16 and HPV18 are the most carcinogenic HPV genotypes. Certainly, there is evidence that persistence of HPV16 is highly linked to the risk of ≥CIN3 (40). There is no epidemiologic evidence that HPV18 persistence is more tightly linked to the development of ≥CIN3 than persistence of other HPV genotypes, but this may be the consequence of underdetection of HPV18-induced cervical precancer (41–43). It would additionally be useful to evaluate the risk for non-HPV16/HPV18 genotypes, since vaccination against HPV16 and HPV18 (44, 45) may eliminate these infections from some populations. Finally, there are important new data on the difference between prevalently detected and incidently detected HPV infections. As already discussed, prevalently detected HPV infections are “left-censored”; that is, there is an unknown duration of infection prior to detection, and there is evidence (27) that the longer an infection endures, the more likely it will continue to do so. This relation naturally leads to an age effect in which there is a greater likelihood of persistence of prevalently detected HPV infections with increasing age at first detection (46). As HPV screening programs mature, infections will increasingly be found as incident infections during longitudinal follow-up. The relation between HPV persistence and risk of ≥CIN3, and the modifiers of that risk, will likely differ substantially between prevalently detected and incidently detected HPV infections because the length of duration preceding detection—that is, “age of infection”—will be substantially different.

In conclusion, the goal of adding new tests to successful cervical cancer screening programs is to improve the accuracy of screening and thereby reduce overtreatment and possibly increase cost-effectiveness. Monitoring of HPV infections for cervical cancer screening is a potentially promising tool, but its use clinically awaits further evaluations to determine the critical parameters and the availability of reliable, validated methods of detection. Koshiol et al. (20) have provided an important analysis that highlights the possible utility of HPV viral monitoring while helping to identify gaps in knowledge. As discussed above, unresolved key issues remain before HPV monitoring can be introduced into clinical applications. More fundamentally, we do not yet understand the biologic determinants of viral clearance versus persistence. Identification of biomarkers that strongly predict viral persistence might overcome the aforementioned limitations of repeat measurements required to identify those women with persistent infections.

It is also unclear when persistent infections become precancerous lesions, that is, when detection of viral persistence is synonymous with the presence of a precancerous lesion. This transition will not be clarified while the diagnostic procedure, colposcopy and directed biopsies, is only moderately sensitive (22, 23). Women with persistent HPV infection remain at high risk for ≥CIN2 following a colposcopic evaluation that is negative for ≥CIN2 (4)—again evidence for the insensitivity of and poor reassurance provided by the diagnostic procedure. As screening tools

![Figure 2](https://academic.oup.com/aje/article-abstract/168/2/138/100646/142)
continue to advance, our ability to identify women who have or will have cervical precancer will approach certainty and will supersede the sensitivity of colposcopy. Without improvements in colposcopy (47), the distinctions made with HPV genotyping and viral monitoring will be restricted by the limitations in colposcopy, and clinicians, faced with unrealistic medico-legal pressures to prevent all cancer, may lose confidence in its use.

ACKNOWLEDGMENTS

The author thanks Dr. Suzanne Krüger-Kjaer (Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark) for providing data for use in this article. The author also thanks Dr. Mark Schiffman (National Cancer Institute, Bethesda, Maryland) for his thoughtful comments and critical review of the manuscript.

Conflict of interest: none declared.

REFERENCES


