The Expanded Use of HPV Testing in Gynecologic Practice per ASCCP-Guided Management Requires the Use of Well-Validated Assays

Mark H. Stoler, MD, Philip E. Castle, PhD, MPH, Diane Solomon, MD, and Mark Schiffman, MD, MPH

DOI: 10.1309/RNF3C01JKADQCLKP

In 2001, scientists and clinicians in gynecology and pathology gathered with the common goals of standardizing and improving the quality of patient care for women with abnormal cervical cytology. The American Society for Colposcopy and Cervical Pathology (ASCCP)-sponsored consensus conference established a series of evidence-based guidelines to help guide clinical practice and management of women with cervical abnormalities. At this time, the value of a large public investment was realized: the National Cancer Institute–sponsored clinical trial called ALTS (ASCUS [atypical squamous cells of undetermined significance] and LSIL [low-grade squamous intraepithelial lesion] Triage Study) firmly established the clinical value of human papillomavirus (HPV) testing in the management of patients with equivocal cytologic abnormality. Infection with carcinogenic genotypes of HPV is the necessary cause of cervical cancer[1] and its precursor lesions[2] and, as a corollary, the absence of HPV provides strong reassurance of low cancer risk. To a large extent, the wealth of data generated by ALTS regarding multiple parts of precolposcopic and postcolposcopic management "guided the guidelines" developed by the ASCCP-sponsored consensus conference. These guidelines clarified and simplified management and have been widely adopted. In a commentary accompanying the 2002 *JAMA* publication of the ASCCP guidelines it was noted[3]:

"Human papillomavirus (HPV) testing has matured, appears clinically validated and should become integral to both screening and clinical management…. [The] bar has been raised for bringing forward newer HPV diagnostics. Having well-established positive and negative predictive values for the current [Food and Drug Administration (FDA)-approved] test, which are applicable to most populations, allows for clear probabilistic reporting of the results in direct correlation with those of the source cervical cytology. Any new test must document its performance relative to this standard, because many of the proposed management guidelines are based on the performance data…."

Only 1 FDA-approved HPV test was available at the time, although it was widely assumed that the marketplace would rapidly drive the establishment of several newer clinically validated HPV tests.

In addition to the use of HPV testing in the management of women with abnormal cytology, multiple studies have demonstrated that HPV testing in a primary screening setting is more sensitive, more reproducible, and of better predictive value than cytology alone. When used in combination with cytology, HPV testing compensates for the relative insensitivity of a single Papanicolaou test.[4,5] No detectable high-risk HPV essentially translates to "virtually no risk" of precancer or cancer until the next screening.

More than 5 years have passed, and the ASCCP has once again engaged the community to formulate revised guidelines. These guidelines expand the clinical indications for HPV testing at additional points in management of cervical abnormalities, to identify women at risk for precancer (HPV+) from women (often the majority) who can be safely reassured that they are at virtually no risk for cancer (HPV–). Yet as of this writing, there are still no other FDA-approved HPV tests, although several nonapproved tests are being sold. The lack of multiple, competitive, well-validated tests is a problem, as noted in the new, soon-to-be-published ASCCP guidelines.

"These Guidelines expand clinical indications for HPV testing based on studies using validated HPV assays. One cannot assume that management decisions that are based on results of HPV tests that have not been similarly validated will result in the

outcomes that are intended by these guidelines. Furthermore the application of these guidelines using such [unvalidated] tests may increase the potential for patient harm. The appropriate use of these guidelines requires laboratories utilize only HPV tests that have been analytically and clinically validated with proven acceptable reproducibility, clinical sensitivity, specificity and positive and negative predictive values for cervical cancer and verified precancer (CIN [cervical intraepithelial neoplasia] 2,3), as documented by U.S. Food and Drug Administration (FDA) approval and/or publication in peer-reviewed scientific literature."

What exactly is the meaning of the phrase "analytically and clinically validated with proven acceptable reproducibility, clinical sensitivity, specificity and positive and negative predictive values for cervical cancer and precancer (CIN2,3)"? First, the tests must be reliable and reproducible so that a specimen sent to one laboratory would yield the same result at other laboratories. The tests must also be accurate in judging whether a clinically relevant HPV infection is present.

Accuracy is not simple to achieve in the realm of HPV diagnostics.[5,6] In laboratories that offer clinical testing, most tests are *analytically* validated but may not be *clinically* validated. The Clinical Laboratory Improvement Amendments of 1988, which govern laboratory processes in the United States, require laboratories to establish analytic validation for most tests. In the realm of infectious disease, it may be clinically important to have maximal analytic sensitivity for very small numbers of infectious particles like HIV or hepatitis C. In contrast, absolute analytic sensitivity for the smallest possible number of molecules of HPV-16 (or other types associated with risk of cervical cancer) is *not* a desirable result. We know that many patients have HPV detectable by molecular analyses in which no clinical evidence of disease can ever be demonstrated. Most infections clear spontaneously. Excessive *analytic* sensitivity of HPV molecular diagnostics can cause *clinically* nonspecific outcomes, ie, referral to colposcopy and possible biopsy in the absence of CIN 2 or CIN 3. Further exacerbating the problem is that colposcopy, the main tool we have for detecting clinical lesions, is fairly insensitive.[7] If viral testing is too sensitive, the clinicians caring for the patient may come to consider the test results as false-positive because there may be no demonstrable evidence of disease cytologically, colposcopically, or histologically.

Furthermore, HPV testing is an especially complex kind of molecular diagnostic assay, which is certainly one of the reasons that there are not more FDA-approved tests. Approximately 15 HPV types can cause CIN 3, adenocarcinoma in situ, and carcinoma, but with varying carcinogenic risk.[8,9] The numbers of HPV genotypes assayed and the positive cutoff for the number of viral copies of each are related analytic and clinical issues. It is the interplay between analytic sensitivity and clinical sensitivity that is critical to the validation question.[5,6,8,10] Hence, the current FDA-approved test cut point has been optimized and clinically validated to trade off sensitivity vs specificity for a clinically important end point,[11] rather than some arbitrary analytic cutoff of numbers of HPV molecules. The critical cut point is sensitivity for predicting the presence of or future detection of CIN 3 within the screening interval. Although the guidelines use the safer and broader target of CIN 2 or CIN 3, the more rigorous and reproducible standard of adjudicated CIN 3 was preferred in ALTS and is recommended herein as a more trustworthy end point.

Parenthetically, it is also important to recognize that no single test will attain perfect clinical sensitivity. That is, all tests have an inherent false-negative rate. Moreover, aside from test failures, the inherent variability of cervical sampling is intrinsic to any screening test and is prone to error leading to false-negative rates. Practically, the upper limit of clinical sensitivity is actually approximately 95% to 97%, rather than 100%, at any analytic sensitivity when testing for carcinogenic HPV, which is the necessary cause of cervical cancer. Even in the theoretical scenario in which an ultra-analytically sensitive test achieved near 100% sensitivity, resulting in extreme overreferral to colposcopy (or our simply sending everyone to colposcopy), we now understand that the diagnostic standard of colposcopy is imperfectly sensitive.[7] Thus, we must recognize the true limits for achieving maximum programmatic sensitivity for detection of precancer and the huge financial and undeniable iatrogenic costs of trying in vain to achieve it.

Most clinicians and their patients have no desire to understand the nuances between clinical and analytic validation, nor should they really have to in an evidence-based practice. Most assume as a matter of course that any test offered by a clinical laboratory has been clinically validated for the indications for which they are using the test. Unfortunately, in the realm of HPV diagnostics, these assumptions are often without foundation. Homebrew testing is common in molecular diagnostics, and many molecular tests are validated against poorly defined and sometimes difficult to establish "gold standards."[6,8-10,12]

Based on current evidence, we propose for discussion the following working definition of an analytically and clinically validated HPV test as a minimal standard for any test before widespread clinical implementation:

1. The test should be capable of detecting at least the 13 high-risk HPV types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68) defined by international consensus as the types responsible for more than 95% of cervical squamous carcinoma and its precursors and in excess of 90% to 95% of true cervical adenocarcinoma and its precursor, adenocarcinoma in situ. The detection of HPV-66, recently categorized as carcinogenic,[13] is also desirable. Inclusion of other borderline carcinogenic HPV types, such as HPV-73 and HPV-82,[9] may provide only minuscule gains at the cost of further reducing the specificity of HPV testing. Tests should not include noncarcinogenic HPV types, which have no proven independent value in cancer-risk

DOI: 10.1309/RNF3C01JKADQCLKP

assessment. The group that should not be assayed includes HPV-53, which is much too common and too rarely found in precancer and cancer to be included, and HPV-6 and HPV-11, which cause genital warts.

2. To meet currently achievable standards, the test should have a clinical sensitivity to detect at least 92% ± 3% of CIN 3+ such that the negative predictive value of the test is extremely high. The documentation of this level of performance and clinical safety can be provided through an FDA approval documenting the test system as performing at this benchmark in laboratories following the manufacturer's protocol or through independent validation studies that can be directly proven to meet this standard in a statistically valid manner. Data from such studies should be publicly available for peer review. Ideally, independent analysis and documentation should be available for auditing.
3. The test should have clinical specificity of at least 85% such that adequate positive predictive value for CIN 3 or more can be obtained in patients referred for colposcopy and ideally followed up for at least 1 to 2 years to permit another round of diagnosis of incipient lesions.
4. Intrabatch for homebrews and intrabatch and interbatch and interlaboratory reproducibility for manufactured and commercially marketed tests should be documented to ensure reliable performance for use in clinical management. κ values of 0.7 or more for repeated positivity of targeted HPV genotypes (pooled or individual) should be documented to support test robustness.[14-16]
5. One limitation of all current tests is the lack of internal standards to evaluate specimen adequacy. Of course, even a standard that measures squamous cellularity for specimen adequacy does not completely ensure that the correct site has been sampled and the cells of interest are represented. Hence, although such an internal standard could somewhat improve reliability by reducing or eliminating the low percentage of inadequate specimens, its development is problematic and will await future development.

The preceding definition represents a necessary standard or benchmark target because the guidelines for use of HPV testing in the ASCCP guidelines are based on these minimal standards of performance. One may note that there is not a lot of room for improvement in terms of disease sensitivity, which determines negative predictive value, the 2 parameters that govern most of the clinical usefulness of these tests. But use of tests that are less sensitive for the target disease could destroy the usefulness and undermine the performance of the guidelines. Where there is substantial room for improvement is in clinical specificity and, therefore, positive predictive value, which will improve the cost-effectiveness of testing. Well-controlled, statistically powered clinical validation trials of the several candidate assays that are now approaching FDA consideration will soon lead to much needed competition. We hope that the medical community will resolve to *not* accept any new method until appropriate clinical validation studies, submitted to the FDA and/or rigorous peer review, are proven to have similar or better clinical characteristics than the testing we have worked so long to bring into practice.[6,12]

From the Robert E. Fechner Laboratory of Surgical Pathology, University of Virginia Health System, Charlottesville.

References

1. Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol*. 1999;189:12-19.
2. Schiffman MH, Bauer HM, Hoover RN, et al. Epidemiologic evidence showing that human papillomavirus infection causes most cervical intraepithelial neoplasia. *J Natl Cancer Inst*. 1993;85:958-964.
3. Stoler MH. New Bethesda terminology and evidence-based management guidelines for cervical cytology findings [editorial]. *JAMA*. 2002;287:2140-2141.
4. Cuzick J, Clavel C, Petry KU, et al. Overview of the European and North American studies on HPV testing in primary cervical cancer screening. *Int J Cancer*. 2006;119:1095-1101.
5. Titus K. Making a valid point about HPV tests. *CAP Today*. September 2006:1.
6. Stoler MH. HPV testing: if it's not clinically valid, it's dangerous. *HPV Today*. 2006;9:4-5.
7. Guido R, Schiffman M, Solomon D, et al. Postcolposcopy management strategies for women referred with low-grade squamous intraepithelial lesions or human papillomavirus DNA-positive atypical squamous cells of undetermined significance: a two-year prospective study. *Am J Obstet Gynecol*. 2003;188:1401-1405.
8. Schiffman M, Khan MJ, Solomon D, et al. A study of the impact of adding HPV types to cervical cancer screening and triage tests. *J Natl Cancer Inst*. 2005;97:147-150.
9. Munoz N, Bosch FX, de Sanjose S, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med*. 2003;348:518-527.
10. Schiffman M, Wheeler CM, Dasgupta A, et al. A comparison of a prototype PCR assay and Hybrid Capture 2 for detection of carcinogenic human papillomavirus DNA in women with equivocal or mildly abnormal Papanicolaou smears. *Am J Clin Pathol*. 2005;124:1-11.
11. Schiffman M, Herrero R, Hildesheim A, et al. HPV DNA testing in cervical cancer screening: results from women in a high-risk province of Costa Rica. *JAMA*. 2000;283:87-93.
12. Stoler MH. HPV testing in cervical cytology practice: it's all about choice. *Acta Cytol*. 2005;49:117-119.
13. Cogliano V, Baan R, Straif K, et al. Carcinogenicity of human papillomaviruses. *Lancet Oncol*. 2005;6:204.
14. Castle PE, Lorincz AT, Mielzynska-Lohnas I, et al. Results of human papillomavirus DNA testing with the Hybrid Capture 2 assay are reproducible. *J Clin Microbiol*. 2002;40:1088-1090.
15. Castle PE, Wheeler CM, Solomon D, et al. Interlaboratory reliability of Hybrid Capture 2. *Am J Clin Pathol*. 2004;122:238-245.
16. Carozzi FM, Del Mistro A, Confortini M, et al. Reproducibility of HPV DNA testing by Hybrid Capture 2 in a screening setting. *Am J Clin Pathol*. 2005;124:716-721.

DOI: 10.1309/RNF3C01JKADQCLKP