Review

Human Papillomavirus mRNA and p16 Detection as Biomarkers for the Improved Diagnosis of Cervical Neoplasia

Kate Cuschieri1 and Nicolas Wentzensen2

1Specialist Virology Centre, Royal Infirmary of Edinburgh, United Kingdom and 2Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, Maryland

Abstract

Human papillomavirus (HPV) infection of the genital tract is very common and normally follows a benign clinical course; however, in an unfortunate minority of infected individuals, it can cause disease that sometimes leads to cancer. It is accepted that HPV DNA testing has a role in the management of cervical disease both in a prevaccination and postvaccination era; however, to improve the specificity of this approach, there is a requirement to develop and validate tools/assays that can identify women at risk for progressive disease. There is evidence to suggest that detection of viral gene expression both directly and indirectly may constitute a more specific approach for delineating clinically significant infection compared with HPV DNA–based assays. HPV oncogene expression and evidence of its deregulation can be monitored through direct detection of viral mRNA transcripts or through detection of the cellular protein p16. For both approaches, commercial assays have been introduced and numerous studies have been conducted. The present article describes the scientific theory underpinning these approaches, their amenability to routine-diagnostic specimens/settings, and the clinical data that has been garnered through their application thus far. Currently, there is promising data indicating that HPV mRNA and p16 might play an important role in future cervical cancer screening scenarios. Still, large randomized studies are necessary to confirm the preliminary data.


Background

Human Papillomavirus (HPV) DNA testing is already being done in the United States for the management of cervical disease. Moreover, the data generated from the large randomized controlled trials of HPV testing in screening settings—either recently completed or ongoing in Europe—will inform its use in the appropriate countries therein (1-3). Although it is accepted that HPV DNA testing is more sensitive for the detection of cervical disease compared with cytology, its specificity is less so, especially in women under 30 years, largely due to the prevalence of transient, clinically benign infection (4). As a consequence, appropriate contextualization of HPV DNA testing is paramount e.g., age restricting testing and interpreting the result alongside concurrent pathologic diagnosis to avoid unnecessary testing for what is a prevalent viral infection (5).

HPV DNA–Based Tests

The majority of HPV detection tests and certainly the clinically validated ones have incorporated (largely), aggregate detection of a group of HPV types, and can be termed presence/absence tests. These have relied on DNA-based detection of a portion of the gene that encodes the major structural protein of HPV: L1. Presence/absence tests have the benefit of being amenable to high throughput and being easy to interpret and indeed have generated the majority of clinical data, which have informed the use of HPV testing in clinical contexts (6).

However, such tests cannot delineate between transient and potentially transforming infection. To improve the specificity of an HPV test would clearly be of clinical value, and suggested candidates, to this end, have included measurement of viral load and the use of HPV type–specific detection via genotyping (7, 8). With respect to the latter, there is good evidence to suggest that HPV 16 confers a significantly higher absolute risk for the development of significant lesions compared with other high risk types (9). One of the issues of type–specific testing however is the complexity of the data/result it may generate. Multiple infections are common,
and, also the significance of one non–HPV-16 high-risk type over another is poorly understood (10).

Viral load has been shown in some cross-sectional studies to be a useful indicator of clinically significant infection and lesion severity, although there is evidence to suggest that (a) its usefulness may alter according to infecting HPV type and that (b) the wide dynamic range of what constitutes a “significant” viral load currently precludes its use, practically, as a prognostic indicator, especially for types other than HPV 16 (11, 12). Also, accurate measurement requires standardization of sample input to the assay—a significant technical challenge. Testing algorithms are being designed to ameliorate the complexity of testing and result dissemination using typing and load. However, the onus is still on the HPV community to strive for more specific, (bio)markers of significant infection, the detection and interpretation of which can be done in as straightforward a manner as possible.

**Alternatives to HPV DNA Testing**

At the 2006 meeting of the European Research Organization on Genital Infection and Neoplasia, a working group was convened with the remit to compile a consensus report on how to apply (HPV-associated) molecular (bio)markers in practice. The authors concluded that four main lines of research were being followed to identify appropriate markers, namely: (a) detection/measurement of E6 and E7 HPV mRNA transcripts, (b) alterations of the methylation pattern of several genes, (c) alterations of viral (e.g., viral integration) and host (e.g., chromosomal gains and losses) genomes, and (d) detection of cellular proteins that are overexpressed by HPV-infected cells, e.g., the p16ink4A gene product (13). Although biomarkers within (above) categories b and c show demonstrable proof of concepts, experimental work has occurred in academic laboratories on small numbers of samples. Few studies have been reported where they have been applied to sizable clinical sample sets and fewer still where their performance has been compared with validated HPV DNA tests. The focus of this article is to describe the state of the art of two biomarkers, for which standardized, commercial assays are available and which have been the most interrogated in clinical studies, thus far, namely, detection of oncogenic mRNA transcripts and the detection of a protein, p16, which is produced in response to their deregulated expression (Fig. 1). Other biomarkers are currently emerging, among them, a standardized commercial assay for histology and cytology applications, ProExC, which detects two replication associated proteins, MCM2 and Top2A (14, 15). However, limited data precludes thorough assessment of these markers.

---

**Figure 1.** Progression scheme depicting the stages leading to the development cervical cancer from HPV infection to invasive disease. 
**A.** Functional progression model. **B.** Morphology based description of progression; arrows, HPV oncogene expression. **C.** Representative p16 staining patterns associated with morphologic and functional disease stages.
Table 1. Published in-house E6 and/or E7 mRNA detection assays

<table>
<thead>
<tr>
<th>Author</th>
<th>Detection method</th>
<th>HPV type detection range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stoler, 1992</td>
<td>In situ hybridization with Riboprobes</td>
<td>16 and 18</td>
</tr>
<tr>
<td>Smits, 1995</td>
<td>NASBA with E6/E7-specific primers</td>
<td>16</td>
</tr>
<tr>
<td>Rose, 1995</td>
<td>Qualitative E6/E7 RT-PCR</td>
<td>16</td>
</tr>
<tr>
<td>Nakagawa, 2000</td>
<td>Quantitative consensus RT-PCR</td>
<td>16, 18, 31, 33, 35, 51, 52, 56, 58, and 59</td>
</tr>
<tr>
<td>Wang-Johanning,</td>
<td>Quantitative E6/E7 RT-PCR with FRET probes</td>
<td>16</td>
</tr>
<tr>
<td>Lamarcq, 2002</td>
<td>Quantitative E7 RT-PCR with TaqMan probes</td>
<td>16 and 18</td>
</tr>
<tr>
<td>Sotlar, 2004</td>
<td>Qualitative nested E6/E7 RT-PCR with type-specific primers</td>
<td>16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68</td>
</tr>
<tr>
<td>Scheurer, 2005</td>
<td>Quantitative E7 RT-PCR with SYBR Green</td>
<td>16 and 18</td>
</tr>
</tbody>
</table>

HPV Oncogene Expression

It is widely accepted that for HPV to cause cancer, persistent infection and a cellular environment, which allow for high level expression of viral oncogenes E6 and E7 (initially in the basal cell layer and then throughout the epithelium), are necessary. The E6 and E7 proteins are necessary for the virus to replicate itself and are expressed during the productive “normal” life cycle, where their regulation is tightly controlled. When this regulation is disrupted and E6 and E7 are overexpressed, they can abrogate normal tumor-suppressive function and cell cycling, and this process is considered necessary for incurring a malignant phenotype. This deregulation is evidenced particularly by the presence of E7 throughout the epithelium and in the surface layers (16). It is therefore logical to hypothesize that detection of E6/E7 mRNA may identify more clinically significant infection than a DNA approach where a structural gene is targeted, particularly if the surface layers of the epithelium are sampled for analysis as is the case when a cervical cytology specimen is taken.

Technical Issues Relating to RNA Detection

RNA, as a single stranded moiety, is less stable than DNA. Thus, although conceptually, E6 and E7 mRNA may be suitable targets, is their detection practically feasible in routinely collected clinical samples? This is an especially important point considering that certain types of liquid-based cytology (LBC) medium seem to preserve HPV DNA in exfoliated cervical cells very well (17).

Accumulated, indirect evidence suggests that E6 and E7 mRNA are detectable in stored clinical samples. In a key paper by Stoler et al. (18), 3H-labeled riboprobes specific to individual mRNA families were generated, in vitro, to examine serial sections of archived biopsy specimens. In the early 90s, Rose et al. (19) successfully amplified HPV 16 E6/E7 transcripts from paraffin-embedded tissue.

More direct evidence on the stability of HPV mRNA came from the work of Habis et al. (20) who showed that a widely used cytologic collection fluid, PreservCyt (Cytyc Corp.), produced a high yield of RNA and fared better when compared with a high-salt solution designed for RNA preservation. Tarkowski et al. (21) did a prospective study by seeding PreservCyt medium with an HPV-16–containing cell line and found that HPV E6/E7 transcripts were still detectable after 1 year of storage, with 1 HPV-positive cell in 30,000 being consistently detected. A later, prospective study that involved the longitudinal detection of E6/E7 transcripts in cytology samples (PreservCyt) collected from women attending colposcopy showed no loss of HPV transcript detection over a 2-week time course (22).

These data are encouraging; however, we cannot conclude that all alcohol-based cytology medium will constitute suitable buffers for HPV RNA preservation. In a recent article by Powell et al. (23), yields of RNA from an HPV cell line spiked into SurePath medium diminished substantially over a 600-hour time course. Poor recovery of RNA from SurePath has been documented elsewhere also (24). Clearly, technical validation of any viral assay is necessary with respect to different specimen types; nevertheless, the evidence would indicate that HPV mRNA seems to be detectable and stable in certain routinely taken cervical specimens.

HPV E6 and E7 mRNA Detection Methods—In House

The in-house RNA detection methods described to date have tended to be based on reverse transcription-PCR (RT-PCR), and the minority has incorporated a quantitative element. Table 1 depicts the remit of a selection of these assays (18, 25-31) and real-time PCR approaches continue to be used for detection of HPV RNA in the anogenital region and in other sites (32, 33).

Published in-house assays have differed from each other with respect to the nature of the transcripts that have been targeted. E6 and E7 proteins arise from polycistrionic transcripts. The full-length E6/E7 transcript is required to encode the complete E6 protein, whereas through splicing in the E6 open reading frame, a shorter E6/E7 transcript, often called E6*, encodes the E7 protein. HPV 16 has two splice acceptor sites in the E6 open reading frame and can generate two E7 transcripts E7a and E7b. Most of the other HR-HPV types have just the one E7 transcript (34).

The point in relation to diagnostic HPV mRNA assays is that their performance may be related to the nature of the transcript they detect, e.g., in a study by Sotlar et al. (35), the authors found that spliced transcripts were detected less frequently than full-length transcripts in pathologic samples ranging from no CIN to CIN3.

In addition, few of the published in-house HPV mRNA assays described have been internally controlled. The need for internal control is justified so inhibitory and/or inadequate specimens are identified. Moreover,
quantitative measurement of expression requires normalization through measurement of a reference gene with a predictable and stable expression to control for the input into the assay; again, it is expedient to evaluate these with respect to specimen type.

HPV E6 and E7 mRNA Detection Methods—Commercial

Two commercial HPV mRNA assays are available—PreTect HPV Proofer (Norchip) and the APTIMA HPV Assay (GenProbe). The former is a qualitative assay based on a NASBA RNA amplification (of full-length E6/E7 transcript) before detection of amplified RNA with molecular beacons corresponding to HPV types 16, 18, 31, 33, and 45. The HPV Proofer is CE marked, resolves type, is designed primarily to be compatible with LBC samples, and is internally controlled. The latter assay was launched in Europe in May 2008 as a CE-marked test, is also qualitative, and relies on aggregate detection of 14 HR types. It too is compatible with LBC samples and is internally controlled. The chemistry is based on transcription-mediated amplification of full-length E6/E7 transcript preempted by target capture.

Overview of E6 and E7 mRNA Detection in Clinical Studies

Few clinical studies in which HPV RNA testing has been done have been published. Undoubtedly, this will change as more commercial assays become available, recently developed assays (i.e., the APTIMA test) are evaluated, and detection of HPV mRNA expression is not seen exclusively as a means to facilitate basic research. Currently, however, several factors hamper establishing consensus findings across published clinical studies—they include the following: different types of clinical specimens tested, lack of demographic data, different type detection remits of mRNA assays and comparator DNA assays (if used), and the different species of mRNA transcript targeted (i.e., full-length versus spliced, etc.). Moreover, the few studies where DNA and RNA testing has been done on the same sample set have (largely) not constituted direct comparisons of HPV DNA and RNA for the detection of clinical end points—rather, they have been correlative—i.e., samples/cases are already selected due to their HPV DNA positivity so that the corresponding transcript can be looked for.

Cross-sectional Clinical Studies. Few would doubt the sensitivity of HPV DNA testing for the detection of infection associated with significant lesions and cervical cancer; thus, an important initial issue to address is how does RNA testing compare? Evidence collated thus far would indicate that E6/E7 detection in cancers is high as depicted in Table 2.

With respect to high-grade lesions, frequently CIN 3 and CIN 2 or worse are used as a surrogate marker for cancer and as an end point for invasive clinical intervention, respectively. Consequently, sensitivity and specificity for the detection of such are used to assess the performance of HPV detection tools; data relating to E6 E7 detection in CIN2+ can be seen in Table 2. More specifically, Lie et al. (36) did an evaluation of samples collected from 383 women attending colposcopy. HPV DNA and RNA testing was done by the hybrid capture 2 test (13-type range) and the Proofer, respectively. Histology ranged from no disease (61 cases) to invasive carcinoma (20 cases). Overall, agreement between DNA and mRNA positivity was reached in 72%. HPV DNA was detected in a significantly higher number of benign and low-grade lesions lending a higher specificity to the mRNA test. However, 95% and 77% of CIN 2 + lesions were HPV DNA and RNA positive, respectively—a significant difference. It is difficult to attribute whether this difference is solely as a result of mRNA expression being absent/undetectable or whether it was due to the larger type detection range of the hc2. The study of Nakagawa et al. (26) allies with the latter explanation as they found E6/E7 expression in 19 of 19 high-grade lesions when a broad spectrum, consensus approach for E6/E7 detection was used. Lie et al. (36) also suggested that the RNA-negative, DNA-positive high-grade cases would be those more likely to regress without intervention, but prospective data would need to be captured to show this conclusively.

Sotlar et al. (35) tested 779 HPV DNA—positive cervical cytologic samples and looked for E6/E7 oncogene (spliced and unspliced) transcripts via a nested type-specific RT-PCR for 14 high-risk HPV types. Transcript detection was positively correlated with severity of the abnormality, with 18% and 84% of cases of no CIN and CIN 3 testing mRNA positive, respectively. It could be argued that these figures may not represent prevalence of mRNA in cervical samples wholly accurately as HPV DNA—positive samples had been specifically selected preemptively. However, the design of this study does allow for the correlation of DNA and RNA testing approaches that have the same detection range—and the two correlated closely at the higher disease grades, with 103 and 95 cases testing positive for HR-HPV DNA (via multiplex PCR) and mRNA (spliced and/or unspliced transcript), respectively, out of a total of 109 cases with CIN2 or worse. The correlation between RNA and DNA was less so in normal and CIN I lesions with 27.7% and 82% of such testing DNA positive, respectively, compared with 9.1% and 62.5% testing positive for mRNA. Again, these data show that the specificity of an mRNA approach for the detection of significant disease may be higher than a DNA approach (30).

Earlier, we discussed the relevance of a quantitative element to an E6/E7 test, in terms of clinical studies that have involved their application. Wang-Johanning et al. (31) found that HPV 16 E6 and E7 mRNA levels...
positively correlated with increasing severity of cytologic diagnosis using a type-specific RT-PCR. However, Scheurer et al. (28), by using a RT-PCR for HPV 16 and HPV 18 E7 transcripts, found such a correlation existed with HPV 18 only. The difference in the findings of these two studies could be attributable to the length of storage of the RNA extract and sampling method. In the former study, the residual volume of cytologic samples was used directly for testing, whereas the latter took separate samples for cytologic analysis, HPV DNA testing, and HPV RNA testing (with the RNA sample being taken last).

Molden et al. (37) did the largest, clinical, cross-sectional study where HPV mRNA and DNA detection was done, evaluated, and compared. A total of 4,136 women over ages 30 years attending a gynecology clinic were tested using the Proofer for HPV mRNA, and type-specific and consensus (GP5+6+) PCR for HPV DNA. When HPV positivity was related to cytologic status, 2.4% and 9.3% of cytologically normal women were RNA and DNA positive, respectively. When the authors confined the RNA versus DNA comparison to types within the range of the Proofer by using type-specific DNA PCR, overall, 3.1% and 4.4% of samples tested positive for HPV RNA and DNA, respectively, compared with the 10.4% that tested positive by consensus PCR. The authors found that HPV DNA (via GP5+) was significantly higher than detection of RNA in all grades of cytology except for high-grade squamous intraepithelial lesion (HSIL), and of the HSIL cases that were confirmed to be CIN 2 + through histology, 13 of 14 tested mRNA positive compared with 14 of 14 that were DNA positive. These results could imply that the mRNA approach was specific for the detection ‘significant’ disease when compared with a DNA approach.

In a recent study by Castle et al. (38), the authors evaluated the prototype Aptima assay with a DNA line blot assay using 531 cytology samples (with histology confirmation). Two cutoff values for positivity were assessed and related to clinical findings. Significantly fewer women tested positive using the RNA compared with the DNA approach (at both cutoffs) for ≤CIN 1 and at the higher cut off for CIN 1. At the lower cutoff, the same number of CIN2+ cases (95 of 103) were detected by the RNA approach as the DNA approach, and at the higher cutoff, 95 of 103 and 93 of 103 cases were detected by DNA and RNA detection, respectively. These data led to the authors preliminary optimism about the improved specificity of the RNA approach with the caveat that more longitudinal clinical data were needed.

**Prospective Clinical Studies.** Few studies have investigated the prospective sensitivity and specificity of diagnostic mRNA detection. One published study examined a prospective cohort of HPV DNA-positive women with normal cytology. HPV mRNA detection was done on the same specimen that tested DNA positive, and follow-up HPV DNA and RNA testing was done 2 to 3 years after the normal diagnosis (39). The main finding of this study was that women who were HPV mRNA positive at baseline were significantly more likely to have a persistent infection than women who were HPV DNA positive only. Another study looked to assess the predictive value of HPV mRNA and DNA detection for the detection of CIN 2 or worse in 77 women with cytologic low-grade disease, at ~2 years. RNA and DNA detection was done by the Proofer and consensus (GP5+6+) PCR, respectively. The main finding was that HPV mRNA detection was more specific for the detection of CIN 2 at 2 years than the DNA approach (84.9% compared with 50%), and that the two approaches were equally sensitive in detecting the 7 of 77 women who did develop a minimum of CIN2 (40).

Lastly, an RNA assay has been used to test samples collected as part of the Dutch randomized controlled trial POBASCAM—these data are awaited and will be welcome.

**p16 Expression in HPV-Related Disease**

As discussed, alterations in the viral gene expression pattern mark the progression of a productive to a transforming infection, the latter conferring a higher risk for cervical precancer and cancer. p16 is a cellular correlate of the increased expression of oncogenic E6/E7 mRNA: The main actions of the HPV oncoproteins are the degradation of p53 by E6 and thereby the abrogation of apoptosis as well as the release of E2F from pRb that leads to continuous activation of the cell cycle (41, 42). Physiologically, E2F activation is mediated by phosphorylation of the Rb protein. This pathway is strictly regulated by a set of cyclin-dependent kinase inhibitors, among them, p16, which block enzymes phosphorylating pRB (cyclin-dependent kinases). In cells with transforming HPV infections, the regulation of the Rb-E2F pathway is disturbed by E7 and the activation of p16 has no downstream effect (43). As a result, p16 is strongly overexpressed and accumulates in the cells (44). p16 overexpression has been shown in the vast majority of cervical precancers and cancers, whereas in normal tissue, p16 expression is found only rarely (45).

Several properties of p16 make this protein a promising biomarker for HPV-related cancers: The expression is directly linked to the HPV oncogene action because continuous expression of E7 is necessary to maintain a malignant phenotype in HPV-associated cancer (46). The expression of p16 seems to be independent of the HPV type causing the oncogenic infection, obviating the need to detect different HPV types in DNA and RNA assays. Also, in contrast to many classic tumor markers such as ki67 or MYC, p16 is not associated with proliferation, but rather with senescence and cell cycle arrest (47), and is not found expressed in normal basal cells or in other cells with proliferative capacity.

**Role of p16 in the Histologic Diagnosis of Neoplasia**

**Methodology.** The majority of the p16 histology and cytology studies discussed here are based on CE-labeled kits for p16 immunostaining (CINtec; mtm Laboratories), although studies where in-house protocols were used have also been described. In normal cervical histology specimens, barely any p16 expression is observed. Occasionally, metaplastic and endocervical cells may display p16 staining. This is contrasted by strong expression of the protein in cervical precancer and cancer (45). The typical staining pattern of HPV-related
transformation originates in the basal epithelial cells and extends to the upper layers. Three categories have been used to describe p16 staining: no staining, focal staining of single metaplastic or endocervical cells (independent of HPV), and diffuse staining indicative of an HPV-associated lesion. Specifically, focal staining is defined as noncontinuous staining of isolated cells or small cell clusters, usually not located in the basal and parabasal layers.

Diffuse staining is defined as a continuous staining of cells in the basal and parabasal layers (with or without staining of superficial squamous cell layers). The focal staining pattern is considered negative for HPV-associated disease, resulting in a dichotomous evaluation system of positive (diffuse) and negative (negative and focal) staining (48).

**Improvement of Interobserver Variation in Cervical Histology.** An initial study on interobserver variation of cervical histology showed a higher consistency in the assessment of p16-stained specimens compared with conventional H&E slides (48). In a recent study, 247 punch biopsies and 249 cone biopsies were analyzed by 6 pathologists whose initial assessment was based on H&E slides only. Inclusion of a p16-stained slide in the evaluation after a washout phase led to a significant increase in interobserver agreement for both punch and cone biopsies (from 0.49-0.63 and 0.63-0.70, respectively; ref. 49).

Zhang et al. (50) used p16 to refine the histologic gold standard of a large cervical cancer screening study comparing cytology, HPV testing, and VIA. Only CIN2/3 histology results that exhibited diffuse p16 staining were considered "real" cases. Using this refined case definition, the sensitivity of HPV DNA testing and cytology was found to be increased, whereas the sensitivity of VIA was decreased. The p16-negative CIN2/3 cases showed characteristics of immature squamous metaplasia accompanied by unusual cellular atypia, mostly probably leading to false positive histology results.

**Cross-sectional Studies of p16 Histology.** Numerous studies have been conducted analyzing p16 as a biomarker for cervical precancer and cancer in immunohistochemistry, most of them based on convenience samples. A limitation in comparing the studies is related to the different evaluation/scoring strategies applied. The studies using the criteria described earlier (48) with a minimum sample size of 100 are summarized in Table 3, but also within this set, there is considerable heterogeneity in the populations analyzed and the slide evaluation process (44, 45, 48, 51-56).

Murphy et al. for consistency (57-59) used a system to analyze p16 histology, assigning a score of 1 for <10% positive cells, a score of 2 for 10% to 50% positive cells, and a score of 3 for >50% positive cells. In a study of 149 cases with varying histologic diagnosis, the authors found that p16 histology delineated dysplastic squamous and glandular cells with good sensitivity and specificity and that p16 outdid the concomitantly analyzed proliferation markers MCM5 and CDC6 (59). However, because no qualitative distinction between focal and diffuse staining patterns was made, the authors found a higher percentage of positive CIN1 and nondysplastic cases.

Based on a summary of 9 studies including 2,178 cases and applying the criteria from a study by Klaes et al. (or slight variations thereof; ref. 48), 7% of the nondysplastic cases, 54% of the CIN1 cases, 86% of the CIN2 cases, and 96% of both the CIN3 and cancer cases showed diffuse p16 expression (Table 3).

**Prospective Studies of p16 Histology.** In an extension of the cross-sectional study by Wang et al., (56) the prospective value of p16 IHC was determined for a subset of women for whom follow up data were available. In these women, p16 staining showed a positive predictive value of 44% and a negative predictive value of 85% for CIN progression or HPV persistence.

Negri et al. (60) did a study analyzing CIN1 lesions categorized into three groups according to follow-up data, (a) cases with spontaneous regression, (b) cases with progression to CIN3, and (c) cases selected irrespective of progression status. All biopsies were stained with p16 and evaluated according to Klaes et al. (48). In the regressive group, 44% were diffusely p16 positive, whereas 74% of the progressing cases showed diffuse p16 expression. In the unselected group, 58% showed diffuse p16 staining, a percentage in keeping with most studies where p16 expression has been assessed in CIN1.

---

### Table 3. p16 staining in immunohistochemistry

<table>
<thead>
<tr>
<th>Author</th>
<th>Nondysplastic* (%)</th>
<th>CIN1  (%)</th>
<th>CIN2  (%)</th>
<th>CIN3  (%)</th>
<th>Invasive CA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sano, 1998</td>
<td>0/15 (0%)</td>
<td>8/15 (53%)</td>
<td>16/17 (94%)</td>
<td>27/27 (100%)</td>
<td>38/39 (97%)</td>
</tr>
<tr>
<td>Klaes, 2001</td>
<td>1/111 (1%)</td>
<td>29/47 (61%)</td>
<td>32/32 (100%)</td>
<td>60/60 (100%)</td>
<td>58/60 (97%)</td>
</tr>
<tr>
<td>Klaes, 2002</td>
<td>7/58 (12%)</td>
<td>15/17 (88%)</td>
<td>10/10 (100%)</td>
<td>43/43 (100%)</td>
<td>46/46 (100%)</td>
</tr>
<tr>
<td>Agoff, 2003</td>
<td>30/247 (12%)</td>
<td>43/76 (57%)</td>
<td>60/80 (75%)</td>
<td>103/113 (91%)</td>
<td>47/53 (89%)</td>
</tr>
<tr>
<td>Wang, 2004</td>
<td>12/179 (7%)</td>
<td>27/75 (36%)</td>
<td>12/19 (63%)</td>
<td>19/19 (100%)</td>
<td></td>
</tr>
<tr>
<td>Hu, 2005</td>
<td>0/75 (0%)</td>
<td>20/45 (44%)</td>
<td>43/46 (93%)</td>
<td>51/51 (100%)</td>
<td></td>
</tr>
<tr>
<td>Benevolo, 2005</td>
<td>2/75 (2%)</td>
<td>17/54 (51%)</td>
<td>9/10 (90%)</td>
<td>11/11 (100%)</td>
<td>8/8 (100%)</td>
</tr>
<tr>
<td>Ishikawa, 2006</td>
<td>0/140 (0%)</td>
<td>13/53 (25%)</td>
<td>32/40 (80%)</td>
<td>45/48 (94%)</td>
<td>16/16 (100%)</td>
</tr>
<tr>
<td>Focchi, 2007</td>
<td>0/75 (0%)</td>
<td>80/88 (91%)</td>
<td>33/33 (100%)</td>
<td>32/32 (100%)</td>
<td>44/47 (94%)</td>
</tr>
<tr>
<td>Total</td>
<td>50/748 (7%)</td>
<td>252/470 (54%)</td>
<td>247/287 (86%)</td>
<td>391/404 (96%)</td>
<td>257/269 (96%)</td>
</tr>
</tbody>
</table>

*Includes normal, reactive, inflammation, hyperplasia, atypical, equivocal.

1 Population-based study.

2 Criteria modified from Horn et al. (49).

3 Adolescents.
Hariri et al. (61) collected 100 CIN1, 50 HGCIN, and 50 nondysplastic lesions and did a 5- to 7-year follow-up. All p16-negative CIN1 cases showed regression during follow-up, whereas 45% of the p16-positive CIN1 cases progressed or had persistent CIN1.

In conclusion, the studies with prospective end points suggest that CIN1 lesions exhibiting diffuse p16 expression may have a higher risk of progression compared with p16-negative lesions, but more prospective data are necessary to support this notion.

Role of p16 in the Cytologic Detection of Neoplasia

Methodology. p16 immunocytochemistry has been used with several LBC systems, including Thinprep, SurePath, and SEROA. In cytology, the spatial context of single cells that is preserved in histology is lost. Therefore, the distinction between focal and diffuse p16 staining cannot be transferred to cytology. Different approaches have been used to identify dysplastic lesions by p16 cytology. In several studies, p16-positive cells were counted and a cutoff level was defined to ascribe a specimen as p16 positive. Bibbo et al. (62) used 10 p16-positive cells per total slide as a cutoff for detecting high-grade dysplasias, whereas Sahebali et al. (63) have used 1.87 p16-positive cells per 1,000 cells. A different approach is the qualitative assessment of p16-positive cells. Trunk et al. (64) described characteristics of different p16-positive cell types and showed that p16-stained specimens can be analyzed based on morphologic criteria.

In contrast to conventional Pap staining, in p16 cytology, only a fraction of the cells on a slide (those that are stained by p16) need to be analyzed. To facilitate this approach, a scoring system was developed on which to base the assessment of p16-positive cells (65). Using the score, the specificity of p16 aided detection of HGCIN without sacrificing sensitivity. In direct comparison with counting p16-positive cells, the qualitative assessment had a better overall performance (66). Recently, modified p16 cytology staining protocols have been presented. Negri et al. (67) used p16 immunocytochemistry combined with a Pap counterstain. Baak et al. (68) have proposed a concomitant staining approach using antibodies directed against p16 and ki-67 to specifically highlight HPV-transformed proliferating cells.

Cross-sectional Studies of p16 Cytology. Many studies on p16 cytology have been published, most of them using convenience samples or consecutive specimens of a particular cytologic diagnosis. Table 4 summarizes 10 large studies incorporating 1,610 specimens (62, 65, 66, 69-75). About 5% of the normal cytology specimens, 32% of the atypical squamous cells of undetermined significance (including also ASC-US and ASC-H), 37% of the low-grade squamous intraepithelial lesion (LSIL), 93% of the HSIL, and 99% of the cancer specimens were found to be positive in p16 cytology (Table 4). These data are summarized with the caveat that different cytology platforms and systems for p16 cytology assessment were used. Some p16 cytology studies could not be summarized in Table 4. Sahebali et al. (63) determined p16-positive cell counts in consecutive slides from a routine cytology laboratory.
(76). ASC-H and HSIL cases had significantly higher p16 counts compared with the LSIL, ASC-US, and negative cytology cases. Akpolat et al. (77) prepared and analyzed paraffin-embedded blocks from residual Thinprep specimens and found a good correlation between p16 positivity and Pap cytology. In addition to the identification of squamous cervical lesions, p16 cytology has also been used to identify glandular lesions with promising results (78-80).

The data described above indicate that the performance of p16 cytology is dependent on the evaluation strategy applied. Currently, a qualitative analysis of p16-positive cells either by standard cytologic criteria or by using a simplified score based on cytologic abnormalities associated with p16-positive cells can substantially increase the specificity of p16 cytology without sacrificing sensitivity (66). In time, the combination of p16 with other biomarkers might further eliminate subjectivity in cytologic assessment.

**Prospective Studies of p16 Cytology, Triage Studies.**

Bibbo et al. (81) analyzed a series of cytology specimens with a HSIL result, for which follow-up biopsy information was available. In most of the p16 cytology–positive cases, high-grade lesions were detected in the biopsy, whereas the p16-negative cytology cases were confirmed as cervicitis, squamous metaplasia, and CIN1. Nieh et al. (82, 83) did p16 cytology on a series of conventional smears with related follow-up biopsies. Using p16 cytology, 95% sensitivity and 96% specificity was achieved for the detection of HGCIN and greater, compared with 86% and 31% for HPV DNA testing.

In a study by Ziemke (84), p16 cytology was used to triage mild-moderate Pap results. All women with persistent abnormalities (PapIIIID in the Munich classification) were positive for p16, whereas only 23% of those that had subsequent regression were positive for p16 cytology.

Carrozzi et al. (85) used p16 cytology on LBC samples obtained from 283 women with abnormal Pap results. The authors analyzed p16 cytology as an adjunct to HPV testing to prompt colposcopy and found a better PPV for the combined marker set albeit at a decreased detection rate of CIN2+.

In a recent study, p16 cytology was used to triage ASC-US and LSIL cytology specimens in a reflex triage setting (66). Overall, a sensitivity of >90% and a specificity of >80% for the detection of HGCIN were achieved. The triage worked equally well in both the atypical squamous cells of undetermined significance and LSIL cytology groups, making p16 an interesting marker for the triage of LSIL cytology, particularly when compared with HPV DNA testing.

**Role of p16 Protein Quantification for the Detection of Cervical Precancer**

Recently, a biochemical assay to detect p16 levels in solubilized cervical cells has been developed. In contrast to morphologic evaluation, the measurement of protein levels is independent of the observer’s education and training. This assay was done in high-risk populations and was shown to have a good sensitivity for the detection of HGCIN in two independent studies (86, 87).

These preliminary data are promising, but further population-based studies are necessary to evaluate the value of p16 quantitation in cervical cancer screening.

**Conclusions**

Many biomarkers have been proposed for cervical cancer screening, a recent overview is given by Wentzensen et al. (88). Most of the markers have been analyzed in a few small studies only and their value cannot be assessed properly at the moment. Here, we have described two related biomarkers of transforming HPV infections, E6/E7 mRNA expression and detection of the p16 protein. The candidacy of these as relevant biomarkers has been justified through our knowledge of basic HPV virology and in vitro studies. Although both HPV mRNA and p16 have been analyzed in a number of studies, there is clearly some way to go before we can say how they will “best fit” to improve the diagnosis of cervical neoplasia, as either stand alone or as adjunctive tests, in triage or in primary screening contexts. For this, we need more clinical data, particularly sufficiently powered, longitudinal studies where the candidates are assessed alongside concurrent pathology. The preliminary clinical data look promising although more work is needed to show sufficient sensitivity and specificity for the detection of high-grade lesions in real-life clinical situations. An appealing feature of both approaches is that potentially, they would not require highly specialized platforms or a technical sea change for future operators. For example, diagnostic platforms already exist for high-throughput testing of viral RNA that could be adapted to HPV as necessary. p16, if used to inform/refine traditional cytology and histology, builds on existing skills, and if the p16 ELISA approach proves robust, ELISA technology and result interpretation is a mainstay of many diagnostic laboratories. There is currently only one Food and Drug Administration–approved test for HPV infection; this is an unusual situation for what is a stable and, moreover, clinically significant virus. There is therefore a requirement on the research community to seek out and validate new HPV tests or markers of significant infection that are robust and informative but at the same time feasible in routine diagnostic contexts— in this light, those we have discussed in this article show promise.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**References**


84. Ziemke P. [Immunocytochemistry (p16INK4A) as an additional examination for cytological specimens Class III D (Munich Classification II)]. Zentralbl Gynakol 2005;127:140–5.


Human Papillomavirus mRNA and p16 Detection as Biomarkers for the Improved Diagnosis of Cervical Neoplasia

Kate Cuschieri and Nicolas Wentzensen

Cancer Epidemiol Biomarkers Prev 2008;17:2536-2545.