The Sensitivity and Specificity of p16\(^{\text{INK4a}}\) Cytology vs HPV Testing for Detecting High-Grade Cervical Disease in the Triage of ASC-US and LSIL Pap Cytology Results

Karin J. Denton, MD,1* Christine Bergeron, MD, PhD,2* Petra Klement,3 Marcus J. Trunk, MD,3† Thomas Keller, PhD,4 and Ruediger Ridder, PhD,3 for the European CINtec Cytology Study Group‡

Abstract

We analyzed the performance of p16\(^{\text{INK4a}}\) immunocytochemistry on a series of 810 retrospectively collected atypical squamous cells of undetermined significance (ASC-US) and low-grade squamous intraepithelial lesion (LSIL) cases with available biopsy follow-up data, including 94 cases of cervical intraepithelial neoplasia (CIN) 2 and 128 cases of CIN 3. Human papillomavirus (HPV) testing was performed from the same residual liquid-based cytologic specimen, and results for both tests were correlated with histologic follow-up data. Sensitivity values for high-grade CIN (HGCIN) confirmed on biopsy within 6 months were 92.6% (ASC-US) and 92.2% (LSIL) for cytotechnologists’ reviews of p16 cytology and 90.1% (ASC-US) and 95.7% (LSIL) for HPV testing. Sensitivity rates of initial pathologists’ reviews were slightly lower, 76.4% to 80.1%, with levels comparable to cytotechnologists’ results after adjudication. The specificity of p16 cytology for HGCIN detection was significantly higher than for HPV testing for cytotechnologists and pathologists: 63.2% to 71.1% (p16 cytology) vs 37.8% for HPV in ASC-US (P < .001) and 37.3% to 53.3% (p16 cytology) vs 18.5% for HPV in LSIL (P < .001). This evaluation of the diagnostic performance of p16 cytology confirms the potential of this stain for the efficient triage of ASC-US and LSIL cytologic results.

Owing to the low prevalence of high-grade precancerous cervical lesions in the routine screening population, there is a substantial number of Papanicolaou (Pap) cytologic cases with equivocal (atypical squamous cells of undetermined significance [ASC-US]) or mildly abnormal (low-grade squamous intraepithelial lesion [LSIL]) morphologic findings for which no high-grade cervical intraepithelial neoplasia (HGCIN) can be confirmed on histologic specimens collected during subsequent colposcopy-guided biopsy sampling (reviewed by Schiffman et al1 and Wright et al2). Although to a substantial extent colposcopic error is known to lead to failure in finding HGCIN on first referral to colposcopy, there is a large proportion of women within these 2 Pap cytologic categories who neither have nor may develop HGCIN over time. Thus, it is important to implement an efficient triage approach for ASC-US and LSIL Pap cytologic results to identify women at highest risk for underlying precancerous disease and who require immediate further diagnostic follow-up.

Because most cervical cancers are induced by persistent infection with high-risk human papillomavirus (HR-HPV) types, testing for the presence of HPV has been incorporated into the management of Pap cytologic results categorized as ASC-US in various countries. The ASC-US–LSIL Triage Study (ALTS) has shown that women with an ASC-US result and being negative for HR-HPV have a very low risk of having an HGCIN.3 In that study, the triage of ASC-US cases with HPV testing has been found to be similarly effective as repeated cytology or direct colposcopy, with a slightly favorable trade-off of sensitivity and referral rates of HPV triage over repeated Pap cytology.3,4

However, despite a high sensitivity for the detection of underlying HGCIN within the ASC-US group, testing for...
HPV has limitations, especially in terms of specificity. The vast majority of ASC-US cases testing positive for HR-HPV have no underlying high-grade abnormality. Furthermore, given the high prevalence of HPV infections, especially in younger age groups, the effectiveness of triaging ASC-US cases with HPV testing is variable, depending on the patient’s age and other socioeconomic factors.

For the management of cytologic results categorized as LSIL, there is currently no triage tool available that could assist in determining which women with such low-grade cytologic abnormalities should undergo colposcopy, and, therefore, colposcopy is standard for most patients. The ASC-US–LSIL Triage Study and other studies have shown that cytologic interpretations of LSIL are so widely associated with infections with HR-HPV types that there is no clinical usefulness for triaging LSIL cytologic results with testing for the presence of HPV infections.

The relatively low threshold of an ASC-US result positively triaged with HPV testing or an LSIL Pap cytologic result triggering referral to colposcopy greatly enhances sensitivity for the detection of HGCIN, but at the cost of referring most of the women to colposcopy and biopsy follow-up who do not have prevalent CIN 2 or CIN 3 disease or invasive cervical cancer. Therefore, the availability of a marker that provides a sensitivity similar to that of HPV testing but with a significantly higher specificity would be desirable to improve current triage testing of ASC-US cytologic results and to allow for reducing the colposcopy referral rates after LSIL cytologic results.

In recent years, various molecular markers have been proposed as potential candidates for the efficient triage of equivocal and mildly abnormal Pap cytologic results (reviewed by Cuschieri and Wentzensen and von Knebel Doeberitz). One of the most intensely studied markers is the cellular p16INK4a protein (p16), a cyclin-dependent kinase inhibitor that has been demonstrated in a large number of studies to be strongly overexpressed in almost all high-grade precursor lesions and invasive cancers of the cervix uteri (reviewed by Cuschieri and Wentzensen, Wentzensen and von Knebel Doeberitz, and Tsoumpou et al). The expression of this cell cycle regulatory protein has been shown to be directly linked to the transforming activity of the viral E7 oncoprotein at the molecular level. As the continuous inactivation of the retinoblastoma protein (pRb) by the viral E7 oncoprotein is necessary to maintain the malignant phenotype of HPV-associated cancer cells, detection of overexpression of p16 may be used as a highly sensitive and specific surrogate marker for the inactivation of the tumor suppressor pRb in HPV-transformed epithelial cells (reviewed by Wentzensen and von Knebel Doeberitz).

The usefulness of conjunctive testing for p16 overexpression in cervical tissue specimens has been shown in many studies and is widely accepted (reviewed by Cuschieri and Wentzensen and Tsoumpou et al). In cervical cytology, various research studies have been performed to evaluate the potential usefulness of applying p16 immunocytochemical staining protocols, especially for the triage of equivocal or mildly abnormal Pap cytologic results. In most studies, a sensitivity similar to that of HPV testing, but at a substantially higher specificity rate, has been reported for p16 cytology when used for the triage of Pap cytologic results categorized as ASC-US or LSIL or containing atypical glandular cells (reviewed by Cuschieri and Wentzensen and Tsoumpou et al). Although this indicates the potential of p16 cytologic testing for the triage of abnormal Pap cytologic results, many of the research studies show limitations owing to the use of nonvalidated and nonstandardized immunocytochemical reagents and protocols, the lack of standardized interpretation protocols, and a low number of biopsy follow-up results available.

We therefore set out to perform a retrospective, multicenter, multinational, diagnostic case control study to validate the performance characteristics of the CINtec p16 Cytology test (mtm laboratories, Heidelberg, Germany) on a large number (n = 810) of residual liquid-based cervical cytologic specimens with follow-up biopsy specimens available.

Materials and Methods

Case Selection and Sample Preparation

Cervical cytology cases were identified in the archives of 5 anatomic pathology laboratories in Switzerland and Italy by using string search to identify cases with corresponding Pap cytologic results of ASC-US or LSIL. Consecutive such cases were selected for which both residual material was available in the ThinPrep Pap Test liquid-based cytology vial (Hologic, Marlborough, MA) and corresponding tissue blocks of cervical punch biopsies, cone biopsies, and/or endocervical curettage procedures collected within a period of up to 6 months after the date of the index Pap result were available at the respective laboratory. The liquid-based cytology vial had to contain sufficient remaining material for an additional ThinPrep slide preparation and for HPV testing. All patient-related data were delinked from the study samples to ensure the anonymity of patients, and samples were identified by a study sample identification number. The study protocol was approved by the Freiburg Ethics Commission International, Freiburg, Germany (approval number FECI-08/1315).

A cytologic slide for subsequent immunocytochemical staining for p16 was prepared for each case using the T2000 slide processor (Hologic). Cases were excluded from the study for the following reasons: (1) Cytologic slides were damaged during the staining run or subsequent procedures. (2) Control slides from the same staining run were not stained properly. (3) Cytologic slides were obscured by blood, mucus, inflammatory
Denton et al / p16 Cytology for ASC-US and LSIL Triage

exudates, or lubricant, as specified in the Bethesda System for reporting cervical cytology.15 (4) Slides did not meet the minimum squamous cellularity criteria as specified in the Bethesda System for reporting cervical cytology.15

Formalin-fixed, paraffin-embedded tissue specimens matching each case were retrieved from the archive and used for the preparation of 4 additional tissue sections. For cases with more than 1 tissue block, the block with the highest histologic grade lesion as specified in the histology report for the respective case was used. The first and the fourth slide of the tissue sections were stained for H&E at the local study center, and tissue compatibility with the tissue slide preparation used for the initial diagnosis was confirmed by the local histopathologist. The remaining slides for each tissue block were used for p16 staining at a central laboratory using the CINtec Histology Kit (REF 9511, mtm laboratories).

Immunocytochemical Studies

Immunostaining of cervical cytologic specimens for p16 was performed using the CINtec Cytology Kit (REF 9521, mtm laboratories) according to the instructions of the manufacturer. For each staining run, a control slide (mtm laboratories) containing positive and negative control cells with respect to immunoreactivity for p16 was used to validate the staining procedure.

p16 Cytologic Slide Interpretation

For interpretation of p16 immunostaining, 2 different classification methods were used in the 2 clinical study centers reading the p16 cytologic slides. The complete set of slides was read at each center (ie, all slides were read independently by both methods). Irrespective of the method used, a positive case was defined by the presence of at least 1 p16 immunoreactive cell that showed morphologic abnormality.

In 1 of the 2 study centers (Southmead Hospital, North Bristol NHS Trust, Bristol, England), all slides were screened for the presence of p16-immunoreactive cells by a cytotechnologist recruited from a group of 4 cytotechnologists. In all cases in which p16-immunoreactive cells were identified, the respective cytotechnologist made an assessment whether the individual p16+ cell(s) showed abnormal morphologic features. In addition, all slides were evaluated by a pathologist (K.J.D.) who screened the slides for the presence of p16-immunoreactive cells and subsequently categorized the cases as follows: Cases with no immunoreactive cells or with p16-immunoreactive cells but showing no morphologic abnormalities were categorized as negative test results. Cases with p16+ cells and showing morphologic abnormalities were categorized as positive test results and subgrouped into p16+ cells showing borderline dyskaryosis, mild dyskaryosis, moderate dyskaryosis, or severe dyskaryosis,16 depending on the degree of morphologic abnormalities of the respective immunoreactive cells. Image 11 shows various examples for p16 cytologic results and the classification for p16-immunoreactive cells based on conventional criteria for morphologic abnormalities. The cytotechnologists and the pathologist had no previous experience with the interpretation of p16 cytology and received a single structured training session before the start of the study.

In the other study center (Laboratoire Cerba, Cergy-Pontoise, France), all slides were reviewed by 1 cytotechnologist and by a pathologist (C.B.) using a nuclear scoring system that has recently been described as a potential alternative method for the interpretation of p16-stained cervical cytologic slides.12,15 In brief, slides were screened for p16+ cells, and a score of 0 was given to cases without any p16+ cells. The presence of p16-immunoreactive cells without signs of nuclear abnormality was categorized as score 1. Score 2 was given to p16+ cells showing as the most profound alteration a single alteration out of the group of morphologic abnormalities such as increased nuclear/cytoplasmic ratio, altered chromatin, altered nuclear shape, or anisokaryosis. Cytologic cases comprising p16-immunoreactive cells with at least 2 such morphologic alterations within the same cell were rated as score 3. For examples of cases showing the scoring of nuclear abnormalities in p16+ cervical cells, refer to Wentzensen et al.17 For the final data analysis, a score of 2 was used as the threshold defining a positive p16 cytologic test result when using the nuclear scoring approach.

Staff at this study center (Laboratoire Cerba) had previous experience with reading p16 cytologic slides using the nuclear scoring method. In contrast with the first study center, the results of the cytotechnologist’s review were not independently recorded but overwritten by the results of the separate pathologist’s review and, therefore, were not available for separate analysis in this study.

HPV Testing

Testing for HR-HPV was performed by a central independent clinical laboratory (Labor Limbach, Heidelberg, Germany) using residual material from the ThinPrep vials and the Digene High-Risk HPV Hybrid Capture (HC2) DNA Test (Qiagen, Hilden, Germany) according to the instructions of the manufacturer.

“Gold Standard”

The histologic diagnosis established on the recut tissue block specimens was used as the gold standard (ie, diagnostic accuracy criterion) for the study. A majority consensus diagnosis was established for each case by 2 or more pathologists regarded as experts in the interpretation of gynecopathology, including cervical pathology. The reviewers (Frieder Kommoss, MD, Mannheim, Germany; Giovanni Negri, MD, PhD, Bolzano, Italy; Sigrid Regauer, MD, Graz, Austria; and Dietmar Schmidt, MD, Mannheim, Germany) were blinded.
Interpretation of p16 cytologic results using conventional criteria for morphologic abnormalities. A, Atypical squamous cells of undetermined significance (ASC-US) case with p16+ squamous metaplastic cells and no morphologic abnormalities (×40). B, ASC-US case with p16+ squamous cells showing borderline dyskaryosis (×40). C and D, ASC-US (C, ×40) and low-grade squamous intraepithelial lesion (LSIL) (D, ×40) case with p16+ squamous cells showing moderate dyskaryosis. E and F, LSIL cases with p16+ squamous cells showing severe dyskaryosis (E and F, ×100).
to the original local histopathologic result and to all other test results. A combined review of H&E-stained slides and consecutive slides stained with the CINtec Histology Kit was performed by the expert pathologists.

Statistical Analysis

All data were collected on paper case report forms and entered into a database in duplicate, followed by database cleaning and subsequent statistical analysis.

Estimates for sensitivity and specificity were calculated from cross-tabulation of (binary) test results and the gold standard and are presented with their exact 95% confidence intervals (CIs). True-positive fractions (TPF; sensitivity) and false-positive fractions (FPF; 1 – specificity) were compared by investigating ratios of TPFs (rTPF) and ratios of FPFs (rFPF). Significant improvement of sensitivity or specificity was concluded if the rTPF was larger than 1 (superior sensitivity) or the rFPF was smaller than 1, respectively, and if the 95% CIs for the respective ratio did not contain 1. In addition, P values resulting from the McNemar test are reported.

Sample size estimation supposing a sensitivity rate of at least 70% and taking into account a 10% CI revealed that 81 cases of HGCIN should be included for each of the ASC-US and LSIL groups. Assuming 67% vs 50% and 50% vs 20% improvement for specificities of ASC-US and LSIL cases, respectively, and applying formulas given in Pepe, a minimum of 245 ASC-US and 92 LSIL cases without underlying HGCIN should be included. The study was not powered to demonstrate statistical significance for differences in sensitivity levels for the tests analyzed in this assessment.

Results

A total of 945 liquid-based cytology samples categorized as ASC-US or LSIL and the matched corresponding tissue blocks from punch biopsies, cone biopsies, and/or endocervical curettage specimens obtained within a 6-month period after the index Pap cytologic sample and fulfilling all of the inclusion criteria were retrospectively collected from the archives of 5 anatomic pathology laboratories. The age of the cytologic samples was in the range of 1 to 81 months, and median and mean sample ages were 12 and 18 months, respectively.

Out of this pool of 945 individual cases, cytologic samples and their corresponding histology specimens were included in an ordered way with a ranking dependent on the cytologic sample age, starting with the youngest samples and until at least the required number of disease cases (n = 81) according to the statistical analysis plan was reached for each of the 2 Pap cytologic categories (ASC-US, LSIL). The age of the cytologic samples of the final cohort of 810 cases for which p16 cytologic results could be established was 36 months or less.

Figure 1 shows the distribution of histologic diagnoses for ASC-US and LSIL cases based on the majority consensus diagnoses on cervical biopsies served as the “gold standard.” Results were dichotomized into disease (cervical intraepithelial neoplasia [CIN] 2 or CIN 3) and nondisease (CIN 1 and negative for dysplasia) cases.

Figure 1 Flow chart showing the distribution of study samples for atypical squamous cells of undetermined significance (ASC-US) and low-grade squamous intraepithelial lesion (LSIL) categories based on histologic diagnoses. Expert pathologists’ majority consensus diagnoses on cervical biopsies served as the “gold standard.” Results were dichotomized into disease (cervical intraepithelial neoplasia [CIN] 2 or CIN 3) and nondisease (CIN 1 and negative for dysplasia) cases. LBC, liquid-based cytology.
diagnoses of a panel of pathologists and after dichotomizing the histologic results into disease (CIN 2+) and nondisease (CIN 1 and negative for dysplasia) cases, respectively. A total of 222 CIN 2+ cases were included in this study, 81 of them in the ASC-US Pap cytologic group and 141 disease cases within the LSIL group.

**Sensitivity and Specificity of p16 Cytology**

The estimated sensitivity and specificity rates for CIN 2+ and CIN 3 of the p16 cytologic slide interpretation results for the initial pathologists’ review are shown in Table 1. In the ASC-US category, sensitivity rates for CIN 2+ were 76.5% and 78.8% and for CIN 3 75.9% and 83.0%, respectively. Within the LSIL category, a positive p16 cytologic result would have identified 76.4% and 80.1% of the underlying CIN 2+ and consistently 81.1% of the CIN 3 cases using either of the 2 cytologic interpretation approaches. Specificity rates of the pathologists’ reviews for CIN 2+ were calculated to be 65.5% and 71.1% in the ASC-US group and 47.0% and 53.3% in the LSIL group. There were no substantial differences between the sensitivity and specificity rates for the 2 classification systems used for the interpretation of the p16 cytologic slides, ie, either applying conventional classification criteria of morphologic abnormalities to p16-immunoreactive cells (borderline dyskaryosis or worse morphologic features of p16-immunoreactive cells defining a positive test in this study) or using a nuclear scoring system recently proposed as an alternative method for the interpretation of p16 cytologic testing (a score of 2 or higher of p16-immunoreactive cells defining a positive test in this study).12,17

For the majority of CIN 2+ cases that initially were missed by pathologists on the p16 cytologic slides, 1 of the 2 pathologists had categorized the respective case as positive for p16 cytology. Of the 222 cases with underlying CIN 2+, 145 cases had positive p16 test results as interpreted by both pathologist reviewers and using the 2 aforementioned alternative interpretation methods. In 27 p16+ cases with abnormal morphologic features (conventional method) and 30 cases with a p16 score of 2+ (nuclear scoring), the rating was negative by one reviewer but positive by the other reviewer. Of 222 cases with underlying disease, 20 (9.0%) were consistently called negative by both independent reviewers.

To obtain further information about the potential reasons for these discrepant results, we subjected the cases that were initially missed by at least one of the pathologists to a joint adjudication review process. Additional positive and negative cases from the initial review were included to prevent biases owing to expectation and unmasking of disease (workup bias). Bethesda terminology criteria of morphologic abnormalities were used for the adjudication process. In the joint review and discussion of cases, sensitivity of the adjudicated pathologists’ results for CIN 2+ would have been 95.1% in the group of ASC-US cases and 96.4% in the group of LSIL cases.
For the interpretation of p16 cytologic slides using the conventional classification criteria applied to p16-immunoreactive cells, a separate assessment result was available for each slide established by 1 reviewer per slide from a group of 4 cytotechnologists. For this independent review, cytotechnologists were asked to identify p16-immunoreactive cells and then to categorize these cells with respect to the presence or absence of abnormal morphologic features (borderline dyskaryosis or worse). Sensitivities for CIN 2+ (CIN 3) of the cytotechnologists’ interpretation results for the ASC-US and LSIL categories were 92.6% (92.6%) and 92.2% (94.6%), respectively. Specificities for CIN 2+ (CIN 3) were calculated to be 63.2% (58.6%) in the ASC-US group and 37.3% (32.2%) in the LSIL group (Table 1).

**Comparison of p16 Cytologic Results With HPV Testing Results**

HPV testing using the Digene High-Risk HPV HC2 DNA Test was performed independently on residual material from the same liquid-based cytology vial. Sensitivity for CIN 2+ was estimated to be 90.1% within the pool of ASC-US cases with underlying high-grade disease and 95.7% within the LSIL category (Table 1). Specificity rates for the CIN 2+ disease threshold were calculated at 37.8% in the ASC-US category and 18.5% in the LSIL category. The comparison of the p16 cytologic results with HPV testing results revealed a similar sensitivity of the cytotechnologist review results and the adjudicated pathologist review results for the estimated detection of underlying HGCIN in both groups of ASC-US and LSIL cases as obtained for HPV testing. All rTPFs were found to be approximately 1 (95% CIs contain 1, results not shown). However, specificity was found at significantly higher levels for the p16 cytologic rates in the ASC-US category (63.2% vs 37.8%, an increase of specificity rate over HPV testing by 67%; rFPF = 0.59; 95% CI, 0.50-0.69) and the LSIL category (37.3% vs 18.5%, doubling of specificity rate as compared with HPV testing; rFPF = 0.77; 95% CI, 0.70-0.85).

**Discussion**

This study is the first clinical evaluation study for p16 cytology with a sample size that is statistically powered to assess its true performance characteristics in cervical liquid-based cytology cases categorized as ASC-US or LSIL on Pap review and comparing the results with histologic follow-up data for disease verification. The study has shown that p16 cytology provides significantly better specificity than HR-HPV for the triage of ASC-US and LSIL Pap cytologic cases. Furthermore, the specificity associated with p16 cytology on LSIL is comparable to that seen with HR-HPV triage of ASC-US Pap cytologic results. Sensitivity rates of p16 cytology for underlying CIN 2+ or CIN 3 were found in a range similar to that for HPV testing for cytotechnologists’ reviews and after pathologist adjudication. The initial pathologists’ reviews by either method provided slightly lower sensitivity rates.

The findings of the study are in good agreement with the majority of the previous studies (reviewed by Cuschieri and Wentzensen and Tsoumpou et al). A total of 7 studies have been published so far that included ASC-US and/or LSIL cases and that also provided HPV testing data for the study population. In these previous studies, p16 cytologic testing showed somewhat lower or similar sensitivity rates to HR-HPV HC2 DNA testing but substantially higher specificity compared with HC2 (reviewed by Tsoumpou et al). However, most of the previous studies were hampered by relatively low sample numbers, which did not allow for a statistically meaningful analysis of the specificity gain of the p16 cytologic testing approach, or lack of appropriate clinical controls. All of these studies, including the study presented in this analysis, were not designed and powered to demonstrate statistical significance for differences in sensitivity rates between the p16 cytologic testing approach and HPV testing.

The results of our study validate the idea of using a biomarker for the triage of equivocal or mildly abnormal Pap cytology results that acts further downstream of a mere infection with HR-HPV types to increase the specificity of the triage testing result. At the same time, owing to the strong association of the overexpression of p16 at the molecular level with the onset of cellular transformation mediated by the HR-HPV E7 oncoprotein, overall there is also a high level of sensitivity of the p16 immunostaining approach for the detection of underlying HGCIN. This is clinically relevant because 8% to 12% of women with ASC-US on Pap cytology may have underlying high-grade disease, whereas still more than 50% of women with ASC-US need referral to colposcopy when using HPV testing as the triage tool. Furthermore, depending on the age group and the prevalence of HR-HPV infections in the respective population, the referral numbers may be variable and substantially higher, especially in the younger age groups.

In the case of LSIL, testing for the presence of HPV infection has been shown in various studies to be not meaningful because the vast majority of cases are HPV+. For this reason, in most countries, women with LSIL Pap cytologic results are directly referred to colposcopy to detect underlying HGCIN. The availability of a test that identifies women having the highest risk for the presence of an underlying HGCIN would be highly desirable to provide a safe alternative to the current management of patients with LSIL Pap cytologic results. The results of the current study indicate that p16 cytology has the potential to be used as a triage test for LSIL Pap cytologic results.
Based on the study design, which required diagnostic follow-up by colposcopy and biopsy within 6 months after the index Pap cytologic result of ASC-US or LSIL, no assessment of the long-term prognostic values of p16 cytologic and HPV test results could be performed in this study. Additional longitudinal studies will be required to evaluate the potential long-term predictive value of p16 cytology and to compare it with that of HPV testing, the latter of which is known to identify women with an increased long-term risk for developing high-grade CIN. Such studies may provide further information about the potential value of a combined p16 cytologic and HPV testing triage approach in which p16 cytology may confirm its potential as a potentially better immediate triage tool, whereas testing for HR-HPV may be superior with respect to the long-term risk stratification for the development of cervical malignancies.

As with HPV testing, for routine Pap cytologic screening, p16 immunocytochemical testing can be performed as a reflex test in conjunction with liquid-based cytology following an index Pap cytologic result of ASC-US or LSIL. As an additional advantage, the p16 immunocytochemical stain may also be used on destained conventional or liquid-based cytology Pap specimens in case this procedure is in compliance with local regulations. This approach would allow “reflexing” ASC-US or LSIL Pap cytologic results directly to p16 cytologic testing, even when conventional Pap smears rather than liquid-based cytology are used, such as in many European settings. This may further contribute to the reduction of costs and improvement of the efficacy of a p16 cytology-based approach of triaging ASC-US and LSIL Pap cytologic results. Although a cost-effectiveness analysis has not yet been performed, it might be reasonable to expect that reducing the number of referrals to colposcopy for ASC-US Pap cytologic results by using a test showing similar sensitivity but significantly higher specificity for the detection of high-grade CIN than HPV testing should have a favorable effect on the cost-effectiveness of p16 cytology-based triaging approaches. Similarly, reducing the numbers of colposcopies for women showing mild Pap cytologic abnormalities should lower the overall costs associated with the management of women with LSIL on Pap cytology.

As discussed in a recent meta-analysis there is substantial variability in the literature regarding the cutoffs for positive p16 cytologic test results defined by individual investigators, such as qualitative parameters (ie, morphologic cytologic features), quantitative parameters (ie, number of immunoreactive cells), and combinations of both. Given the fact that p16 overexpression may be sporadically detected in a fraction of nondysplastic cells, such as in some metaplastic and endocervical cells, in which the p16 protein executes its normal physiologic role of controlling cell cycle progression, the application of morphologic interpretation criteria to p16 immunoreactive cells seems to be required to leverage the full potential for a high specificity of p16 cytologic testing.

In our study, we assessed the performance of 2 slightly different interpretation procedures for p16 cytologic slide reading. Besides the regular interpretation protocol that is based on the application of routine criteria for morphologic abnormalities of cervical epithelial cells to p16-immunoreactive cells, we investigated the performance of a recently proposed alternative interpretation protocol that is based on a scoring algorithm for grading the nuclear abnormalities of p16-immunoreactive cells. The results for both interpretation protocols performed in parallel by 2 pathologists on the same set of 810 p16 cytologic samples in this study revealed that there is no substantial difference in the diagnostic performance of the 2 protocols. The differences in sensitivity and specificity profiles observed for the 2 methods (Table 1) are more likely to reflect reader variability between the 2 pathologists interpreting the slides rather than true differences based on the interpretation protocol used.

The results from the interpretation using the nuclear scoring system on a large number of cases with histologic follow-up in this study indicates that, different from the original reports, a nuclear score of grade 2 (ie, presence of a p16-immunoreactive cell with a single or more morphologic abnormalities) should be regarded as the clinically relevant cutoff for a positive p16 cytologic test result when using the proposed nuclear scoring approach. Thus, as a conclusion, this study has shown that it is sufficient and efficient as to the clinical outcome to follow an interpretation algorithm for the p16 cytologic slide reading that uses p16 immunoreactivity as a locator tool for cells of diagnostic interest and to then apply routine, conventional interpretation criteria for morphologic abnormality to the p16+ cells. The presence of at least 1 cell fulfilling these criteria should define a positive p16 cytologic test result. Using the alternatively proposed nuclear score provides similar results and, therefore, may be used instead of the conventional interpretation criteria, depending on the individual reviewer’s preference.

In this study, differences were seen for the p16 cytologic results obtained for the initial pathologists’ reviews vs the interpretation results of cytotechnologists independently reading the cases. As can be expected, the cytotechnologists’ interpretations had higher initial sensitivity than the pathologists’ interpretations vs higher specificity levels observed for the pathologists’ reviews. Adjudication between the pathologists indicated a comparable sensitivity for the pathologists’ reviews as seen for the cytotechnologists’ review, which suggests that a discussion of discrepant results may be important. Although a workup bias for the data in the adjudication meeting in our study cannot be excluded owing to case selection, an adjudication process reflects typical clinical practice. In fact, such an adjudication process between a cytotechnologist
and a pathologist is typically used for the routine interpretation of Pap cytologic slides, differing from the adjudication process between the 2 pathologists used in this study, which may not be practical in the routine clinical setting. It is noteworthy that in most of the cases in which p16-immunoreactive dysplastic cells were missed, there were very few positive cells and most showed only very faint staining, which in part could be related to the extended age of the individual leftover cytologic specimens (data not shown).

Our study had the following strengths: (1) The evaluation was performed on a large set of clinical cases (n = 810) with disease verification on histology for all cases included in the study and a large number (n = 222) of HGCIN disease cases. (2) A standardized and quality-controlled reagent set and immunostaining protocol were used that have been validated for use in cervical cytology. (3) The gold standard was built by majority consensus diagnoses on cervical tissues using H&E- and p16-stained slides. (4) Clearly defined and standardized p16 cytologic slide interpretation algorithms were applied. Furthermore, the study was performed according to good clinical practice requirements to ensure maximum data integrity.

Study weaknesses were as follows: (1) The noninterventional design of the study using retrospectively collected specimens may lead to underestimation of true sensitivity and specificity of p16 cytologic testing when used for triaging ASC-US and LSIL cytologic results in real-world clinical settings. (2) The potential negative influence of sample age is a potentially confounding factor that may lower sensitivity and increase specificity rates owing to technical limitations that are not relevant in real clinical practice settings in which the test would be performed within days or weeks after sample collection. (3) There was bias in the selection of cases included in the study based on the need for availability of biopsy follow-up data as one of the study requirements. (4) There was potential bias because the technical component of the preparation of p16-immunostained slides was performed in the laboratory of the study sponsor, whereas the HPV test was done in an independent, qualified routine molecular diagnostic testing laboratory. These are limitations that may interfere with the assessment of the true performance of p16 testing and its comparison with HPV testing.

This study has shown that CINtec p16 cytology may be used as a triage test for ASC-US and LSIL Pap cytologic results. The study demonstrates that p16 cytology has the potential to provide high sensitivity for the presence of underlying HGCIN, with significantly higher specificity rates than HPV testing. Therefore, p16 may further improve the current approach for triaging ASC-US cytology with HPV testing and may provide an efficient immediate triage tool for LSIL cytologic results.


