

## Performance of p16/Ki-67 Immunostaining to Detect Cervical Cancer Precursors in a Colposcopy Referral Population

Nicolas Wentzensen<sup>1</sup>, Lauren Schwartz<sup>1</sup>, Rosemary E. Zuna<sup>2</sup>, Katie Smith<sup>2</sup>, Cara Mathews<sup>2</sup>, Michael A. Gold<sup>3</sup>, R. Andy Allen<sup>2</sup>, Roy Zhang<sup>2</sup>, S. Terence Dunn<sup>2</sup>, Joan L. Walker<sup>2</sup>, and Mark Schiffman<sup>1</sup>

### Abstract

**Purpose:** Cytology-based screening has limited sensitivity to detect prevalent cervical precancers. Human papilloma virus (HPV) DNA testing is highly sensitive and provides a high, long-term reassurance of low risk of cervical cancer. However, the specificity of HPV DNA testing is limited, requiring additional, more disease-specific markers for efficient screening approaches.

**Experimental Design:** Liquid-based cytology samples were collected from 625 women referred to colposcopy. A slide was stained using the CINtec plus cytology assay. Pap cytology and HPV genotyping were conducted from the same vial. Clinical performance characteristics were calculated for all women, stratified by age, and for women referred with a low-grade squamous intraepithelial lesion (LSIL) Pap.

**Results:** p16/Ki-67 positivity increased with histologic severity, from 26.8% in normal histology, 46.5% in CIN1, 82.8% in CIN2 to 92.8% in CIN3. Among women with CIN3, p16/Ki-67 positivity increased from 77.8% for women younger than 30 years without HPV16 to 100% for women 30 years and older with HPV16. The sensitivity and specificity to detect CIN3+ were 93.2% and 46.1%, respectively, and increased to 97.2% and 60.0% among women 30 years and older. In women with high-risk (HR)-HPV-positive atypical squamous cells of undetermined significance (ASC-US) and LSIL, sensitivity and specificity for detection of CIN3 were 90.6% and 48.6%, respectively.

**Conclusions:** p16/Ki-67 testing could reduce referral to colposcopy by almost half while detecting the most severe cases of CIN3. The high sensitivity of p16/Ki-67 with significantly improved specificity compared with HPV testing makes p16/Ki-67 a viable option for LSIL triage. Further studies are required to evaluate p16/Ki-67 as triage marker in HPV-based screening strategies. *Clin Cancer Res*; 18(15); 4154–62. ©2012 AACR.

### Introduction

Cervical Pap smear screening has led to a substantial reduction of cervical cancer incidence in countries with established screening programs. However, Pap cytology has limited reproducibility, and a single Pap test has limited sensitivity to detect cervical precancer. The recognition that virtually all cervical cancers are caused by infections with carcinogenic human papillomavirus types has led to the development of new molecular assays to improve cervical cancer screening strategies, most importantly testing for human papilloma virus (HPV) DNA (1).

HPV DNA testing is highly sensitive and provides a high, long-term reassurance of low risk of cervical cancer among women testing negative, permitting safe extension of screening intervals (2). However, the specificity of HPV DNA testing for cervical precancer and cancer is limited, as HPV infections are very common in the population, and most infections clear spontaneously after 1 to 2 years. Thus, most HPV infections are not associated with risk for developing invasive carcinoma.

Recently, more disease-specific molecular markers of cervical cancer have been recognized that may provide a combination of high sensitivity and high specificity for detecting cervical precancer (3). Most of these markers have been identified on the basis of our understanding of HPV-related carcinogenesis. The progression from HPV infection to cervical precancer is characterized by a substantial change in the viral gene expression, from a transient infection characterized by expression of structural genes to a transforming infection with strong expression of viral oncogenes that interfere with host cell-cycle control (4, 5).

The expression of several host genes is affected by the oncogene products of HPV, including those involved in cellular proliferation, such as Ki-67, and cell-cycle control, such as p16. Immunostaining for p16 has been determined

**Authors' Affiliations:** <sup>1</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, Maryland; <sup>2</sup>University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma; and <sup>3</sup>Vanderbilt University, Nashville, Tennessee

**Note:** Supplementary data for this article are available at Clinical Cancer Research Online (<http://clincancerres.aacrjournals.org/>).

**Corresponding Author:** Nicolas Wentzensen, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Blvd, Room 5024, Rockville, MD 20852. E-mail: wentzenn@mail.nih.gov

doi: 10.1158/1078-0432.CCR-12-0270

©2012 American Association for Cancer Research.

### Translational Relevance

Cervical cancer screening programs worldwide are moving from primary cytology screening to human papilloma virus (HPV) DNA testing. While HPV DNA testing offers high sensitivity and negative predictive value (NPV), its specificity for cervical precancer is limited. Thus, secondary, more disease-specific markers such as p16/Ki-67 are needed to decide who among the screen-positives require further management and treatment. In a large U.S.-based colposcopy referral population with excellent disease ascertainment, we show that cytologic staining for p16/Ki-67 has comparable sensitivity, but significantly higher specificity than HPV DNA testing, potentially reducing colposcopy referral by half. We show that p16/Ki-67 does best at detecting precancers with the highest risk of progression to cancer, namely, those related to HPV16 among women 30 years and older. These data from a large, independent study suggest that p16/Ki-67 can be an important component of new HPV-based screening strategies.

to be an effective biomarker of cervical disease in histology and cytology specimens (6, 7). It is widely used to improve the reproducibility of cervical biopsy interpretations. In cytology, p16 can improve the accuracy for detecting cervical precancer compared with conventional cytology. However, as some benign cervical epithelial cells can exhibit p16 expression, evaluation of p16 staining requires additional morphologic evaluation (8). Recently, a double-label immunostain for p16 and Ki-67 was developed that allows recognition of abnormal cells simply based on co-staining of the 2 markers in the same cell, potentially obviating the need for morphologic interpretation. The accuracy of combined p16 and Ki-67 cytology staining for detection of cervical precancer has been analyzed in few studies so far (9, 10). Here, we evaluated the clinical performance of p16/Ki-67 in a large study of women referred to colposcopy for abnormal cervical cancer screening results.

### Materials and Methods

#### Study population

We conducted our analysis in the Biopsy Study, a population-based study of women referred to colposcopy for abnormal cervical cancer screening results conducted at the University of Oklahoma Health Sciences Center (OUHSC; Oklahoma City, OK) from February 2009 to August 2011. All women 18 years and older who were referred for colposcopy at the University of Oklahoma colposcopy clinic were eligible for inclusion in the study. Exclusion criteria were previous treatment for cervical disease [including loop electrosurgical excision procedure (LEEP), cold-knife conization, cryotherapy, LASER therapy, or hysterectomy], prior chemotherapy or radiation treatment for cervical neoplasia, pregnancy, HIV infection, and inability to give

informed consent. Referral to colposcopy was based on the American Society for Colposcopy and Cervical Pathology (ASCCP) guidelines and included women with HPV-positive atypical squamous cells of undetermined significance (ASC-US) and low-grade squamous intraepithelial lesion (LSIL) or greater independent of HPV status. A total of 2,270 women with appointments at the colposcopy clinic were approached about participation in the study. Of these, 897 women were ineligible for enrollment or could not be enrolled for other, unknown reasons. A total of 690 of 1,373 (50.3%) eligible women agreed to participate in the study that involved an extended biopsy protocol. Of those, 673 had residual liquid-based cytology material available to produce monolayer slides for conducting p16/Ki-67 immunostaining. Forty-eight samples were excluded because of low cellularity to give a final study population of 625 women.

#### Cytology, colposcopy, and histology

All women enrolled in the study had a colposcopy conducted and at least one biopsy was taken in all women. Before colposcopic biopsy procedures, cervical cytology was collected using a Wallach broom device and transferred to PreservCyt solution (Cytoc Corp.). The cytology specimen was used for ThinPrep liquid-based cytology, for HPV DNA analysis and for p16/Ki-67 immunostaining. The study involved an extended biopsy protocol with digital photographic documentation of biopsy sites. Up to 4 biopsies were taken from distinct acetowhite lesions or large heterogeneous lesions extending over 2 quadrants. If fewer than 4 directed biopsies were taken, a biopsy from a quadrant without any visible CIN (random biopsy) was added. As per standard practice, all histologically confirmed CIN3 and most CIN2 were treated by LEEP of the transformation zone. At least 3 biopsies were taken in more than 80% of the colposcopic procedures conducted. We considered outcomes both based on worst biopsy result during the colposcopy and worst overall outcome, including biopsy results and LEEP outcomes.

#### Liquid-based cytology

Thin-layer cytology slides were prepared using the ThinPrep 2000 slide processor (Hologic). The PreservCyt container was placed in the instrument, and using the gynecology sample procedure, the cell suspension was homogenized and cells were collected on a filter membrane and transferred to a glass slide. Cells were placed in ethanol for fixation and stained using the Papanicolaou method. Cytology slides were evaluated by a cytotechnologist and confirmed by a pathologist using the revised Bethesda nomenclature. The analysis of p16/Ki-67 positivity in cytologic categories is based on the study Pap, whereas the analyses of HPV-positive ASC-US and LSIL triage are based on the outside referral Pap result.

#### HPV testing

HPV detection and genotyping in cytology specimens was done using the Linear Array HPV Genotyping Test (Roche

Molecular Diagnostics). The linear array assay is a type-specific PGMY09/11 L1 primer PCR assay for 37 HPV types (6, 11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 45, 51, 52, 53, 54, 55, 56, 58, 59, 61, 62, 64, 66, 67, 68, 69, 70, 71, 72, 73, 81, 82, 83, 84, and 89). Linear array genotyping was conducted according to the manufacturer's instructions with slight modifications, as previously described (11, 12). In brief, the procedure followed recommendations of the manufacturer with the variation that 10  $\mu$ L of template DNA was amplified, and the amplified products were hybridized and detected using an automated Auto-LiPA staining system using 2.5 mL of each reagent per strip (as compared with 4.0 mL in manual processing). The linear array results were evaluated by unamplified examination of the strips by 2 independent observers. An unambiguous, continuous band was judged to indicate that biotinylated amplicons had hybridized to complementary sequences of probes bound to the strips and was considered a positive result.

### p16/Ki-67 testing

A second cytology slide was prepared from the residual PreservCyt material using a T2000 slide processor (Hologic). Immunostaining of cervical cytology slides for p16/Ki-67 was conducted using the CINtec Plus Kit (Roche mtm laboratories AG), according to the manufacturer's instructions. The kit contains a ready-to-use primary antibody cocktail comprising a mouse monoclonal antibody (clone E6H4) directed at human p16<sup>INK4a</sup> (p16) protein and a rabbit monoclonal antibody (clone 274-11 AC3) directed against human Ki-67 protein. Secondary detection reagents included horseradish peroxidase (HRP) and goat anti-mouse fragment antigen-binding Fab' antibody fragments for detection of p16 antibodies and a polymer reagent conjugated to alkaline phosphatase and goat anti-rabbit Fab' antibody fragments for detection of Ki-67. HRP-mediated conversion of 3,3'-diaminobenzidine (DAB) chromogen and alkaline phosphatase-mediated conversion of Fast Red chromogen generated the brown and red staining of p16 and Ki-67, respectively. Alcohol-free hematoxylin was used as a counterstain, with a 2-step mounting procedure including an aqueous mounting medium followed by a permanent mounting step. Slides that did not meet the squamous cellularity criteria as specified in the Bethesda 2001 Cervical Cytology Classification system for reporting cervical cytology were excluded from evaluation. A trained cytotechnologist reviewed all cases for the presence of cells staining positively with both markers. A case was considered positive if one or more cervical epithelial cell(s) stained both with a brown cytoplasmic stain (p16) and a red nuclear (Ki-67) irrespective of the interpretation of morphologic abnormalities. Slides without any double-stained cells were called negative for p16/Ki-67 dual-stain cytology.

### Statistical analysis

Of the 625 women with valid p16/Ki-67 staining results, 612 (97.8%) had cytology results, 619 (99%) had histology results, and 622 (99.5%) had linear array genotyping data. We considered women with presence of any of 13 carcino-

genic genotypes (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68) as high-risk (HR)-HPV DNA positive. We used contingency tables to evaluate p16/Ki-67 expression by study cytology results and by worst study biopsy results.

McNemar test was used to evaluate differences in positivity for p16/Ki-67 and carcinogenic HPV in cytologic categories. Sensitivity, specificity, positive predictive value (PPV), NPV, with 95% confidence intervals were calculated for 2 different endpoints, CIN2+ and CIN3+ for all women with histology endpoints (using worst endpoints from either biopsy or LEEP).

We sought to evaluate triage of women with HPV-positive ASC-US and LSIL referral Pap results. Because we did not have access to referral cytology specimens (and some referrals were based on conventional Pap smears), we simulated a triage approach using a test interposed between screening cytology and colposcopy. We restricted this analysis to 404 women with an outside referral Pap of HPV-positive ASC-US or LSIL. In addition, we evaluated the performance of p16/Ki-67 triage separately for HPV-positive ASC-US ( $n = 140$ ) and LSIL ( $n = 264$ ). We used McNemar test to compare differences in sensitivity and specificity for p16/Ki-67, HR-HPV testing, and HPV16/18 testing.

Sensitivity and specificity with confidence intervals were plotted on a receiver operator curve (ROC) characteristics graph. Statistical analyses were conducted using Stata (StataCorp), and summary ROC graphs were created using SigmaPlot.  $P < 0.05$  was considered statistically significant.

## Results

### Cytology and histology results and carcinogenic HPV prevalence

A total of 625 women from a colposcopy referral population were included in the analysis. The median age was 26 years (interquartile range, 23–31; complete range, 18–67). Of 622 women with HPV results available, 488 (78.5%) were positive for HR-HPV. Among 612 women with cytology results available, 143 (23.4%) had negative for intraepithelial lesion or malignancy (NILM), 117 (19.1%) had ASC-US, 138 (22.5%) had LSIL, 63 (10.3%) had ASC-H, and 148 (24.2%) had high-grade squamous intraepithelial lesion (HSIL) as study cytology result (Table 1). There was an increasing proportion of HR-HPV with increasing cytology severity, from 56% in NILM, 72% in ASC-US, 86% in LSIL to 93% in HSIL (Table 2). The high proportion of HR-HPV-positive women in NILM and ASC-US categories reflects that women included in the study were referred for either HPV-positive ASC-US HPV+ or LSIL and higher, according to the ASCCP guidelines.

### p16/Ki-67 positivity in cytology categories

p16/Ki-67 double staining positivity (p16/Ki-67+) increased from the lowest to highest cytology categories (based on study cytology result); 19.6% in NILM, 40.2% in ASC-US, 68.8% in LSIL, 82.5% in ASC-H, and 95.3% in HSIL (Table 2). The p16/Ki-67 prevalence was much lower

**Table 1.** Cross-tabulation of referral cytology results and study cytology results

Referral cytology	Statistics	Study cytology						Total
		NILM	ASC-US	LSIL	ASC-H	HSIL+	Missing	
ASC-US	<i>n</i>	54	30	22	19	12	3	140
	p16/Ki-67+	14 (25.93%)	8 (26.67%)	16 (72.73%)	16 (84.21%)	12 (100%)	1 (33.33%)	67 (47.86%)
	CIN2+	4 (7.41%)	2 (6.67%)	8 (36.36%)	9 (47.37%)	10 (83.33%)	0 (0%)	33 (23.57%)
	CIN3+	1 (1.85%)	1 (3.33%)	2 (9.09%)	5 (26.32%)	3 (25%)	0 (0%)	12 (8.57%)
LSIL	<i>n</i>	49	53	88	21	45	8	264
	p16/Ki-67+	10 (20.41%)	23 (43.40%)	58 (65.91%)	17 (80.95%)	42 (93.33%)	2 (25%)	152 (57.58%)
	CIN2+	4 (8.16%)	6 (11.32%)	31 (35.23%)	12 (57.14%)	36 (80%)	2 (25%)	91 (34.47%)
	CIN3+	0 (0%)	2 (3.77%)	5 (5.68%)	1 (4.76%)	11 (24.44%)	1 (12.50%)	20 (7.58%)
ASC-H	<i>n</i>	18	21	14	10	15	1	79
	p16/Ki-67	1 (5.56%)	10 (47.62%)	12 (85.71%)	9 (90%)	15 (100%)	1 (100%)	48 (60.76%)
	CIN2+	4 (22.2%)	4 (19.05%)	10 (71.43%)	5 (50%)	15 (100%)	1 (100%)	39 (49.37%)
	CIN3+	1 (5.56%)	1 (4.76%)	1 (7.14%)	2 (20%)	5 (33.33%)	1 (100%)	10 (12.66%)
HSIL+	<i>n</i>	13	8	11	11	70	1	114
	p16/Ki-67+	3 (23.08%)	3 (37.50%)	8 (72.73%)	9 (81.82%)	68 (87.14%)	1 (100%)	92 (80.70)
	CIN2+	3 (23.08%)	3 (37.50%)	7 (63.64%)	7 (63.64%)	63 (90.00%)	1 (100%)	84 (73.68%)
	CIN3+	0 (0%)	0 (0%)	3 (27.27%)	4 (36.36%)	34 (48.57%)	0 (0%)	41 (35.96%)
Missing	<i>n</i>	9	5	3	2	9	0	28
	p16/Ki-67+	0 (0%)	3 (60%)	1 (33.33%)	1 (50%)	7 (77.78%)	0 (0%)	12 (42.86%)
	CIN2+	0 (0%)	1 (20%)	0 (0%)	1 (50%)	9 (100%)	0 (0%)	11 (39.29%)
	CIN3+	0 (0%)	0 (0%)	0 (0%)	0 (0%)	5 (55.56%)	0 (0%)	5 (17.86%)
Total	<i>N</i>	143	117	138	63	151	13	625
	p16/Ki-67+	28 (19.58%)	47 (40.17%)	95 (68.84%)	52 (82.54%)	144 (95.36%)	5 (38.46%)	371 (59.36%)
	CIN2+	15 (10.49%)	16 (13.68%)	56 (40.58%)	34 (53.97%)	133 (88.08%)	4 (30.77%)	258 (41.28%)
	CIN3+	2 (1.40%)	4 (3.42%)	11 (7.97%)	12 (19.05%)	58 (38.41%)	1 (7.69%)	88 (14.08%)

in HR-HPV-negative NILM (3.9%), ASC-US (16.7%), and ASC-H (0%). In the whole population, the HR-HPV prevalence was higher than the p16/Ki-67 prevalence ( $P < 0.001$ ), but this was differential by cytology category. The percentage of p16/Ki-67-positive subjects was lower for NILM ( $P < 0.001$ ), ASC-US ( $P < 0.001$ ), and LSIL ( $P < 0.001$ ) but did not significantly differ from HR-HPV in ASC-H ( $P = 0.06$ ) and HSIL ( $P = 0.78$ ), indicating higher specificity of p16/Ki-67 than of HR-HPV testing.

#### p16/Ki-67 positivity by histology results

We observed an increasing proportion of p16/Ki-67-positive subjects with increasing histologic severity, from 26.8% in women with normal histology, 48.0% in women with atypical squamous metaplasia, 46.5% in women with CIN1, 82.8% in women with CIN2, and 92.8% in women with CIN3 (Table 3). Only a small subset of CIN2 and about 30% of CIN3 is expected to progress eventually to invasive cancer. HPV16 and older age at diagnosis of precancer are

**Table 2.** p16/Ki-67 positivity in cytology categories

	NILM	ASC-US	LSIL	ASC-H	HSIL	Cancer	Total
All							
Total	143	117	138	63	148	3	612
% DS+	19.6	40.2	68.8	82.5	95.3	100.0	59.8
HR-HPV+							
Total	79	84	117	57	140	3	480
% DS+	30.4	47.6	71.8	89.5	95.7	100.0	70.0
HR-HPV-							
Total	63	33	19	6	8	0	129
% DS+	3.9	16.7	50.0	0	87.5	N/A	20.35
McNemar ( <i>P</i> )	<0.001	<0.001	<0.001	0.06	0.78	N/A	<0.001

Abbreviation: DS+, p16/Ki-67 double-stain positive.

**Table 3.** p16/Ki-67 positivity in histology categories stratified by age and HPV16 status

	Normal	Atypical metaplasia	CIN1	CIN2	CIN3	Cancer	Total
Total	112	25	228	169	83	6	623
% DS+	26.8	48.0	46.5	82.8	92.8	100.0	59.6
HPV16- (<30)	52	14	131	85	18	0	300
% DS+	25.0	64.3	46.6	81.2	77.8	N/A	55.3
HPV16- (30+)	41	3	64	22	11	3	144
% DS+	24.4	0	31.3	68.2	90.9	100	40.3
HPV16+ (<30)	14	5	25	52	34	0	130
% DS+	28.6	40.0	72.0	92.3	97.1	N/A	80.8
HPV16+ (30+)	4	1	6	8	19	3	47
% DS+	50.0	100.0	83.3	75.0	100.0	100.0	87.8

Abbreviations: DS+, p16/Ki-67 double-stain positive; N/A, not applicable.

associated with higher risk of invasion (13, 14). Thus, we stratified women by presence of HPV16 and into 2 age categories, <30 and 30 and older, reflecting the current age cutoff for HPV-based primary screening ( $\geq 30$ ). In all histologic categories except for atypical squamous metaplasia, HPV16-positive women had a higher percentage of p16/Ki-67 positivity than HPV16-negative women, overall and stratified by age (Table 3). Among women with CIN3, the proportion that was p16/Ki-67-positive increased from 77.8% for those <30 years and without HPV16 to 100% among women 30 years and older who were HPV16-positive. Among women with CIN2, those who were HPV16-positive had a higher proportion of p16/Ki-67-positive results, but interestingly, the p16/Ki-67 positivity was higher among women <30 years.

#### Clinical performance of p16/Ki-67 testing in the study population

We analyzed clinical performance characteristics of p16/Ki-67 to detect CIN2 or greater and CIN3 or greater among the whole study population and stratified by age (Table 4). Overall, p16/Ki-67 positivity was 59.6%, indicating that referral to colposcopy would have been reduced by almost half if p16/Ki-67 testing was used as an additional triage test in this population. The sensitivity and specificity to detect CIN2+ was 86.4% and 59.5%, respectively, and for CIN3+, it was 93.2% and 46.1%, respectively. Among women 30 years and older, p16/Ki-67 positivity dropped to 51.1%,

further reducing colposcopy referral in this group. Among women 30 years and older, the sensitivity and specificity to detect CIN2+ was 84.8% and 67.5%, respectively, and for CIN3+, it was 97.2% and 60.0%, respectively (Fig. 1). p16/Ki-67 provided good risk stratification among women 30 years and older, with an absolute risk of CIN3+ of 36.8% among p16/Ki-67-positive women (PPV) and a low absolute risk of CIN3+ of 1.1% among p16/Ki-67-negative women ( $1 - NPV$ ).

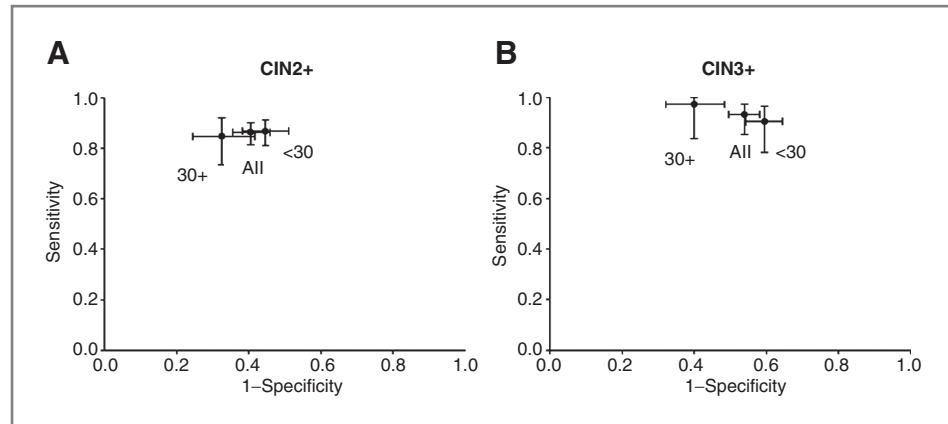
#### Clinical performance of p16/Ki-67 testing for triage of LSIL and HPV-positive ASC-US

Women with LSIL and HPV-positive ASC-US are uniformly referred to colposcopy. We estimated the performance of p16/Ki-67 testing to triage women with these screening test results by restricting the analysis to the 404 women referred for colposcopy with an outside Pap result of HR-HPV-positive ASC-US or LSIL. The p16/Ki-67 positivity was 54.5%, indicating that colposcopy referral would be reduced by almost half if p16/Ki-67 was used as triage test (Table 5). The sensitivity and specificity for detection of CIN2+ were 85.5% and 59.4%, respectively. The sensitivity and specificity for detection of CIN3 were 90.6% and 48.6%, respectively. The absolute risk of CIN3 among p16/Ki-67-positive women with LSIL referral was 13.2%, whereas the absolute risk of CIN3 among p16/Ki-67-negative women was 1.6% ( $1 - NPV$ ). For comparison with p16/Ki-67, we analyzed the results of HR-HPV testing and of

**Table 4.** Clinical performance of p16/Ki-67 in a colposcopy population

	Sensitivity	Specificity	PPV	NPV	Referral
All women CIN2+ ( $n = 258$ )	86.4% (81.5–90.2)	59.5% (54.2–64.5)	60.1% (54.9–65.1)	86.1% (81.1–90.0)	59.55% (55.7–63.3)
All women CIN3+ ( $n = 89$ )	93.2% (85.3–97.2)	46.1% (41.8–50.4)	22.3% (18.3–27.0)	97.6% (94.6–99.0)	
<30 y and CIN2+ ( $n = 190$ )	86.8% (81.0–91.1)	55.4% (48.9–61.7)	60.4% (54.3–66.2)	84.3% (77.5–89.4)	63.19% (58.4–67.6)
<30 y and CIN3+ ( $n = 52$ )	90.4% (78.2–96.4)	40.5% (35.6–45.7)	17.2% (13.0–22.3)	96.9% (92.4–98.8)	
$\geq 30$ y and CIN2+ ( $n = 66$ )	84.8% (73.4–92.1)	67.5% (58.3–75.6)	58.9% (48.3–68.8)	89.0% (80.3–94.3)	
$\geq 30$ y and CIN3+ ( $n = 36$ )	97.2% (83.8–99.9)	60.0% (51.7–67.8)	36.8% (27.4–47.4)	98.9% (93.2–99.9)	51.08% (44.0–58.2)

Figure 1. Sensitivity and specificity of p16/Ki-67 immunostaining to detect CIN2+ (A) and CIN3+ (B) in the complete referral population and stratified by age.



HPV16/18 genotyping in the same population (Table 5). HR-HPV DNA testing had the highest positivity (76.0%), followed by p16/Ki-67 positivity (54.5%) and HPV16/18 (28.0%). For both endpoints, p16/Ki-67 had a sensitivity close to HR-HPV DNA testing ( $P = 0.01$  and  $0.32$  for CIN2+ and CIN3, respectively) but a higher specificity ( $P < 0.0001$  both for CIN2+ and CIN3, respectively). HPV16/18 genotyping was less sensitive than the other tests ( $P < 0.0001$  for CIN2+ and CIN3, respectively) but substantially more specific than both HR-HPV DNA and p16/Ki-67 testing ( $P < 0.0001$  for all tests and all endpoints; Fig. 2). Similarly, we evaluated the performance of p16/Ki-67 to triage women referred to colposcopy for HPV-positive ASC-US and LSIL separately (Supplementary Tables S1 and S2).

**Discussion**

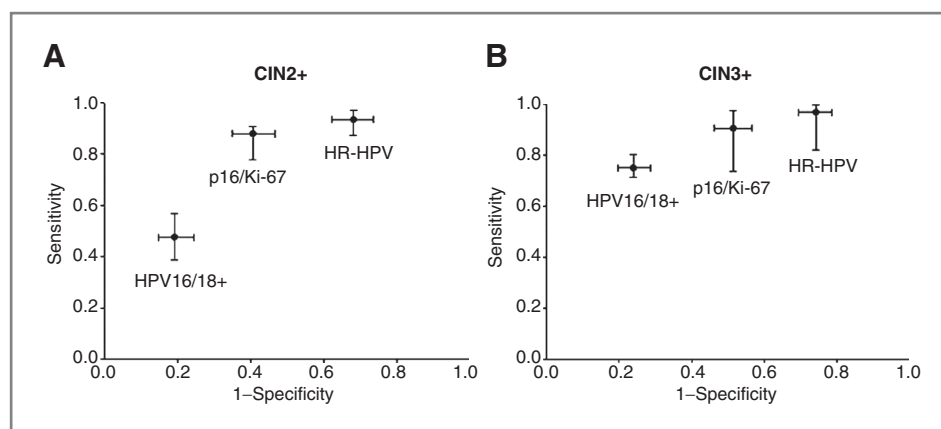
The recognition of HPV as the necessary, but insufficient factor for developing cervical cancer has led to a paradigm shift in cervical cancer prevention programs. HPV DNA testing is increasingly being considered for primary cervical cancer screening. While HPV DNA testing is highly sensitive for detecting cervical precancer, thus providing a high reassurance against developing cervical cancer in the next 5 to 10 years after a negative test, it lacks specificity and

cannot discriminate between common transient infections and rare prevalent precancers (1). Several molecular markers have been developed that are more specific for HPV-related transformation with the goal to improve the detection of cervical precancer. p16 alone has been widely evaluated as a biomarker for cervical precancer (6, 7). However, as some normal cells may express p16, observation of p16-positive cells in cytology preparations requires additional morphologic evaluation to achieve adequate specificity. Recently, an assay has been developed that combines staining for p16 with staining for the proliferation marker Ki-67 on cytologic slides. Theoretically, coexpression of p16 and Ki-67 in the same cell should indicate HPV-related transformation and obviate the need for morphologic interpretation.

We conducted an evaluation of p16/Ki-67 cytology in a large colposcopy referral population in the United States. We found that p16/Ki-67 was positive in 59% of all women enrolled in the study, with a sensitivity of 86% for CIN2+ and 93% for CIN3+. These results indicate that p16/Ki-67 testing could reduce referral to colposcopy by almost half while detecting almost all relevant disease. The sensitivity estimates in our study were very similar to what has been reported in previous industry-sponsored studies using p16/Ki-67 (9, 10).

**Table 5.** Clinical performance of p16/Ki-67 to detect CIN2+ ( $n = 124$ ) or CIN3+ ( $n = 32$ ) in women with HR-HPV-positive ASC-US or LSIL

	Sensitivity	Specificity	PPV	NPV	Referral
p16/Ki-67					
CIN2+	85.5% (77.8–90.9)	59.4% (53.3–65.1)	48.4% (41.6–55.2)	90.2% (84.7–93.9)	54.48% (49.6–59.3)
CIN3+	90.6% (73.8–97.5)	48.6% (43.5–53.9)	13.2% (9.2–18.6)	98.4% (94.9–99.6)	
HPV-HR+					
CIN2+	93.5% (87.3–97.0)	31.9% (26.5–37.8)	38.2% (32.7–43.9)	91.2% (83.8–96.1)	76.0% (71.6–79.9)
CIN3+	96.9% (82.0–99.8)	25.8% (21.5–30.7)	10.2% (7.1–14.3)	99.0% (93.5–99.9)	
HPV16/18+					
CIN2+	47.6% (38.6–56.7)	80.8% (75.5–85.2)	52.7% (43.1–62.1)	77.4% (72.1–82.0)	28.0% (23.8–32.6)
CIN3+	75.0% (71.3–80.3)	76.1% (71.3–80.3)	21.4% (14.5–30.4)	97.2% (94.4–98.7)	



**Figure 2.** Sensitivity and specificity of p16/Ki-67 immunostaining, HR-HPV testing, and HPV16/18 testing to detect CIN2+ (A) and CIN3+ (B) among women with an HR-HPV-positive ASC-US or LSIL referral Pap cytology result.

We sought to analyze further the determinants of p16/Ki-67 positivity in CIN2 and CIN3. Stratification by age and HPV16 status showed that the assay had 100% sensitivity for detection of HPV16-positive lesions among women 30 years and older, whereas the lowest percentage was observed in women <30 years without HPV16 infections (78%). On the basis of the well-documented risk of CIN3 associated with HPV16 infections (15–17) and the higher risk of invasion with long-term persistent CIN3s (18), it is assumed that CIN3 related to HPV16 among women older than 30 years has the highest risk of invading to cancer and should be detected by a safe triage assay. In contrast, CIN3 related to types other than HPV16 among younger women may grow slower and may have a much lower risk of invasion, as indicated by their lower proportion in cervical cancers (13). We show that p16/Ki-67 positivity reflects this heterogeneity of CIN3 and does best at detecting the subgroup of CIN3s with highest risk of invasion (HPV16+, 30 years and older). Furthermore, all 5 cancers detected in the study population were positive for p16/Ki-67. Thus, our data indicate that it would be safe to delay colposcopic evaluation of women negative for p16/Ki-67 and await further testing. For example, repeat HPV testing or p16/Ki-67 testing could be conducted after 1 year, similar to currently recommended management guidelines for women positive for HPV but negative for cytology (19).

In our study, all women with an LSIL referral Pap and an HPV-positive ASC-US referral Pap were evaluated by colposcopy and biopsy, allowing us to analyze the performance of p16/Ki-67 for triage of these cytology categories. For these women, p16/Ki-67 achieved a sensitivity equal to HR-HPV testing with significantly improved specificity, and a possible reduction of referral by almost half compared with current practice of referring all women with HR-HPV-positive ASC-US and LSIL to colposcopy. These findings support that p16/Ki-67 can be a viable option for cytology triage, if proven cost-effective. We were not able to measure p16/Ki-67 in the referral cytology specimen to simulate a reflex triage approach. Instead, our study provides data for a delayed evaluation with p16/Ki-67.

Overall, the specificity estimates in our population were comparable, albeit slightly lower than what has been reported in 2 studies sponsored by the manufacturer of the p16/Ki-67 double-stain assay (9, 10). We had excellent disease ascertainment because an extended biopsy protocol was used with up to 4 targeted biopsies. Thus, we were able to distinguish true-positive from false-positive test results better than in studies where disease is missed because of insensitive colposcopic biopsy protocols. Our assays were conducted from the same specimen type (residual PreservCyt material), run in the same laboratory and evaluated according to the same criteria as the previous studies. In contrast to the previous reports, we only used the cytotechnologist's evaluation of p16/Ki-67 positivity, without additional pathologist's review. The major promise of the p16/Ki-67 double-stain assay is that it does not require adjunct morphologic interpretation of positive cells and vastly simplifies the evaluation process compared with p16 staining alone. We sought to evaluate the assay according to this premise, without extensive review and adjudication of the slides. Currently, the test is considered positive when at least one cell exhibits p16/Ki-67 costaining. While this approach obviously has a high sensitivity, increasing the threshold may lead to a better overall tradeoff between sensitivity and specificity.

Our study shows that p16/Ki-67 has a very high cross-sectional sensitivity and that the residual risk of CIN3 in women testing p16/Ki-67-negative is very low. Prospective studies will be required to evaluate how long a negative p16/Ki-67 test provides reassurance against developing CIN3. Currently, we do not have sufficient data to estimate the prospective risk of CIN3 in women testing positive for HPV but negative for p16/Ki-67. Our cross-sectional data, showing 100% sensitivity for CIN3 among older women with HPV16-related lesions, suggests that colposcopic evaluation of p16/Ki-67-negative women can be delayed safely. With the performance characteristics observed in our study, p16/Ki-67 could safely reduce colposcopy referral by almost half in the overall population, with a higher reduction seen in women 30 years and older. The benefits of reduced numbers of

colposcopies have to be weighed against the additional costs related to p16/Ki-67 testing in this population. Further studies are required to develop algorithms for retesting using molecular markers in HPV-positive, p16/Ki-67-negative women.

With primary HPV DNA testing widely adopted in cervical cancer screening programs, more disease-specific markers such as p16/Ki-67 may assume an important role to triage HPV+ women, as has been successfully shown using p16 staining alone (20). Among women 30 years and older, the target population for triage of HPV-positive women, we observed higher specificity with unchanged sensitivity than in women younger than 30 years. If cytologic immunostains such as p16/Ki-67 or other biomarkers are to be widely accepted, automated evaluation strategies will be important. We have previously shown that automated evaluation of p16 staining on cytologic slides is feasible (21) and will continue to adapt the algorithms to p16/Ki-67 staining.

In summary, in a large colposcopy population with excellent disease ascertainment due to an aggressive colposcopic biopsy protocol, we show that p16/Ki-67 cytology has a high sensitivity and specificity for detecting cervical precancer that may warrant use of p16/Ki-67 testing to reduce colposcopy referral. Further studies among HPV+ women are required to evaluate the potential role of p16/Ki-67 as a triage marker in new HPV-based screening strategies.

## References

- Schiffman M, Wentzensen N. From human papillomavirus to cervical cancer. *Obstet Gynecol* 2010;116:177–85.
- Dillner J, Rebolj M, Birembaut P, Petry KU, Szarewski A, Munk C, et al. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. *BMJ* 2008;337:a1754.
- Sahasrabudde VV, Luhn P, Wentzensen N. Human papillomavirus and cervical cancer: biomarkers for improved prevention efforts. *Future Microbiol* 2011;6:1083–98.
- Doorbar J. Papillomavirus life cycle organization and biomarker selection. *Dis Markers* 2007;23:297–313.
- Wentzensen N, von Knebel DM. Biomarkers in cervical cancer screening. *Dis Markers* 2007;23:315–30.
- Cuschieri K, Wentzensen N. Human papillomavirus mRNA and p16 detection as biomarkers for the improved diagnosis of cervical neoplasia. *Cancer Epidemiol Biomarkers Prev* 2008;17:2536–45.
- Tsoumpou I, Arbyn M, Kyrgiou M, Wentzensen N, Koliopoulos G, Martin-Hirsch P, et al. p16(INK4a) immunostaining in cytological and histological specimens from the uterine cervix: a systematic review and meta-analysis. *Cancer Treat Rev* 2009;35:210–20.
- Wentzensen N, Bergeron C, Cas F, Eschenbach D, Vinokurova S, von Knebel DM. Evaluation of a nuclear score for p16INK4a-stained cervical squamous cells in liquid-based cytology samples. *Cancer* 2005;105:461–7.
- Petry KU, Schmidt D, Scherbring S, Luyten A, Reinecke-Luthge A, Bergeron C, et al. Triage of Pap cytology negative, HPV positive cervical cancer screening results with p16/Ki-67 Dual-stained cytology. *Gynecol Oncol* 2011;121:505–9.
- Schmidt D, Bergeron C, Denton KJ, Ridder R. p16/ki-67 dual-stain cytology in the triage of ASCUS and LSIL papanicolaou cytology: results from the European equivocal or mildly abnormal Papanicolaou cytology study. *Cancer Cytopathol* 2011;119:158–66.
- Wang SS, Zuna RE, Wentzensen N, Dunn ST, Sherman ME, Gold MA, et al. Human papillomavirus cofactors by disease progression and human papillomavirus types in the study to understand cervical cancer early endpoints and determinants. *Cancer Epidemiol Biomarkers Prev* 2009;18:113–20.
- Wentzensen N, Schiffman M, Dunn ST, Zuna RE, Walker J, Allen RA, et al. Grading the severity of cervical neoplasia based on combined histopathology, cytopathology, and HPV genotype distribution among 1,700 women referred to colposcopy in Oklahoma. *Int J Cancer* 2009;124:964–9.
- Li N, Franceschi S, Howell-Jones R, Snijders PJ, Clifford GM. Human papillomavirus type distribution in 30,848 invasive cervical cancers worldwide: variation by geographical region, histological type and year of publication. *Int J Cancer* 2011;128:927–35.
- van BM, van den Akker-van Marle ME, Warmerdam PG, Meijer CJ, Walboomers JM, Habbema JD. Present evidence on the value of HPV testing for cervical cancer screening: a model-based exploration of the (cost-)effectiveness. *Br J Cancer* 1997;76:651–7.
- Chen HC, Schiffman M, Lin CY, Pan MH, You SL, Chuang LC, et al. Persistence of type-specific human papillomavirus infection and increased long-term risk of cervical cancer. *J Natl Cancer Inst* 2011;103:1387–96.
- Kjaer S, Hogdall E, Frederiksen K, Munk C, van den Brule A, Svare E, et al. The absolute risk of cervical abnormalities in high-risk human papillomavirus-positive, cytologically normal women over a 10-year period. *Cancer Res* 2006;66:10630–6.
- Schiffman M, Glass AG, Wentzensen N, Rush BB, Castle PE, Scott DR, et al. A long-term prospective study of type-specific human papillomavirus infection and risk of cervical neoplasia among 20,000 women in the Portland Kaiser Cohort Study. *Cancer Epidemiol Biomarkers Prev* 2011;20:1398–409.

## Disclosure of Potential Conflicts of Interest

M. Schiffman has Other Commercial Research Support from Qiagen CareHPV and Roche free specimen testing. No potential conflicts of interest were disclosed by the other authors.

## Authors' Contributions

**Conception and design:** N. Wentzensen, M.A. Gold, R. Zhang, J.L. Walker, M. Schiffman

**Development of methodology:** N. Wentzensen, R. Zhang, S.T. Dunn, J.L. Walker

**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** R.E. Zuna, K. Smith, C. Mathews, M.A. Gold, R. Zhang, S.T. Dunn, J.L. Walker

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** N. Wentzensen, L. Schwartz, R. Zhang, J. L. Walker, M. Schiffman

**Writing, review, and/or revision of the manuscript:** N. Wentzensen, L. Schwartz, R.E. Zuna, K. Smith, C. Mathews, M.A. Gold, R.A. Allen, R. Zhang, S.T. Dunn, J.L. Walker, M. Schiffman

**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** C. Mathews, R. Zhang, J.L. Walker

**Study supervision:** N. Wentzensen, M.A. Gold, J.L. Walker  
**Interpreted Pap tests, biopsies and LEEP specimens for the SUCCEED study (including both the original and the biopsy study):** R.E. Zuna  
**PI enrolling all the patients in the study:** J.L. Walker

## Grant Support

The study was supported by the Intramural Research Program of the National Cancer Institute.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 25, 2012; revised May 24, 2012; accepted May 29, 2012; published OnlineFirst June 6, 2012.



18. McCredie MR, Sharples KJ, Paul C, Baranyai J, Medley G, Jones RW, et al. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. *Lancet Oncol* 2008;9:425–34.
19. Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. *CA Cancer J Clin* 2012;62:147–72.
20. Carozzi F, Confortini M, Dalla PP, Del MA, Gillio-Tos A, De ML, et al. Use of p16-INK4A overexpression to increase the specificity of human papillomavirus testing: a nested substudy of the NTCC randomised controlled trial. *Lancet Oncol* 2008;9:937–45.
21. Grabe N, Lahrmann B, Pommerencke T, von Knebel DM, Reuschenbach M, Wentzensen N. A virtual microscopy system to scan, evaluate and archive biomarker enhanced cervical cytology slides. *Cell Oncol* 2010;32:109–19.