Triaging borderline/mild dyskaryotic Pap cytology with p16/Ki-67 dual-stained cytology testing: cross-sectional and longitudinal outcome study

M H Uijterwaal[*,1], B I Witte[2], F J Van Kemenade[3], D Rijkaart[4], R Ridder[5], J Berkhof[2],
G A M A Balfoort-van der Meij[6], M C G Bleeker[1], P J F Snijders[1] and C J L M Meijer[1]

[1]Department of Pathology, VU University Medical Center, PO BOX 7057, 1000 MB Amsterdam, The Netherlands; [2]Department of Epidemiology and Biostatistics, VU University Medical Center, PO BOX 7057, 1000 MB Amsterdam, The Netherlands; [3]Department of Pathology, Erasmus University Medical Center, PO BOX 2040, 3000 CA Rotterdam, The Netherlands; [4]Department of Obstetrics and Gynaecology, VU University Medical Center, PO BOX 7057, 1000 MB Amsterdam, The Netherlands; [5]Ventana Medical Systems, Inc., 1910 E Innovation Park Drive, Tucson, AZ 85755, USA and [6]Saltro Diagnostic Centre, PO BOX 9300, 3506 GH Utrecht, The Netherlands

Background: Women with borderline/mildly dyskaryotic (BMD) cytology smears are currently followed up with repeat testing at 6 and 18 months. The objective of this study is to analyse the cross-sectional and longitudinal performance of p16/Ki-67 dual-stained cytology for the detection of cervical intraepithelial neoplasia (CIN) grade 3 or worse (CIN3 +) and CIN2 + in women with BMD, and to compare the results with baseline human papillomavirus (HPV) testing.

Methods: Conventional Pap cytology specimens of 256 women with BMD were dual stained for p16/Ki-67 retrospectively, and compared with baseline HPV results and long-term follow-up results.

Results: p16/Ki-67 dual-stained cytology showed a sensitivity of 100%, a specificity of 64.4% and a negative predictive value (NPV) of 100.% for CIN3 + . Human papillomavirus testing demonstrated similar sensitivity (96.3%), and NPV (99.1%), but a significantly lower specificity (57.6%; $P = 0.024$) for CIN3 + . Sensitivity, specificity and NPV for CIN2 + of dual-stained cytology were 89.7%, 73.1% and 95.1%, respectively, which was similar when compared with HPV testing. Dual-stained cytology showed a significant lower referral rate than HPV testing (43.6% vs 49.1%; $P = 0.043$). During long-term follow-up, no CIN3 + lesions developed in HPV-positive, dual-stained negative women.

Conclusions: Comparable sensitivity and NPV of dual-stained cytology for CIN3 + , combined with a significantly higher specificity, makes p16/Ki-67 dual-stained cytology a viable alternative to HPV testing for triaging BMD.

In the Netherlands cytology screening programme, ~2–3% of women are diagnosed with borderline/mildly dyskaryotic (BMD) Pap cytology. (Health Council of the Netherlands, 2011; Denise et al, 2001). Even though this percentage is relatively low compared with other screening programs, it corresponds to almost 15 000 women per year. Because only 5–15% of these women have high-grade cervical disease, further triage is necessary (Arbyn et al, 2006; Berkhof et al, 2006). Currently, women with BMD are recalled for repeat cytology after 6 months, and if the cytological abnormality persists they are referred for colposcopy-directed biopsy. The disadvantage of repeat testing is that women may become lost to follow-up (Lytwyn et al, 2003) and experience

distress (Melnikow et al, 2002; Dietsch and Davies, 2007). By testing women with BMD for the presence of human papillomavirus (HPV) results in a much higher sensitivity for the detection of cervical intraepithelial neoplasia (CIN)2+/3+ compared with repeat Pap cytology testing, and consequently less follow-up visits (Bulk et al, 2007; Bulkmans et al, 2007; Ronco et al, 2010; Kelly et al, 2011; Kitchener et al, 2013). However, HPV testing has a 2–5% lower specificity for CIN2+/3+ than cytology. Because of this limited specificity, HPV triaging at baseline results in higher colposcopy referral rates than repeat cytology.

Recently, cross-sectional studies have shown a significantly higher specificity for p16/Ki-67 dual-stained cytology compared with HPV triage, without a substantial loss of sensitivity for high-grade cervical lesions (Denton et al, 2010; Schmidt et al, 2011; Wentzensen et al, 2012). Normally, the p16 protein triggers cell cycle arrest in the course of cellular differentiation processes and is rarely observed simultaneously with Ki-67. However, in transforming HPV infections, p16 is strongly overexpressed in proliferating cells, and the p16-induced cell cycle arrest is circumvented by inactivation of pRb, which is mediated by the expression of the viral E7 oncoprotein of high-risk HPV types. Observing dual expression therefore suggests HPV-induced deregulation of the cell cycle and may be used as an indicator for the presence of CIN2+/3+ lesions (Schmidt et al, 2011; Loghavi et al, 2013; Waldstrom et al, 2013).

The aim of this study was to evaluate the use of p16/Ki-67 dual-stained cytology as a triage test at baseline for identifying women with BMD, and underlying CIN3+ and CIN2+ lesions within a population-based screening setting with long-term follow-up (VUSA-Screen study). The results of p16/Ki-67 dual-stained cytology, which was performed on the original, de-stained conventional Pap cytology specimens prepared at patient enrolment into the VUSA-Screen study, were compared with the baseline performance of the Hybrid Capture 2 (HC2) high-risk HPV test. Furthermore, as one of the first studies, we examined the long-term predictive values of the p16/Ki-67 dual staining for the presence or absence of CIN3+ and CIN2+ lesions.

SUBJECTS AND METHODS

Study population. The study was conducted as a sub-study nested in the VUSA-Screen study, a cohort study within the setting of the Dutch population-based cervical cancer screening programme. The VUSA-Screen study took place between 2003 and 2005. The study was designed to evaluate the effectiveness of HPV testing using the Digene HC2 assay (Qiagen, Venlo, The Netherlands) in a population-based screening cohort. The design of the study has been described previously (Rijkaart et al, 2010). In the intervention arm of the VUSA-Screen study, women with BMD test results were subjected to HPV triage testing. In case of HPV-positive test results, women were directly referred to colposcopy. Women with BMD and a negative HPV test result were offered cytology at 6 and 18 months and were referred to colposcopy upon abnormal cytology (⩾BMD) during these follow-up visits (Figure 1).

In the sub-study reported here, we selected all archived conventional cytology slides of the 337 women with BMD participating in the intervention arm of the VUSA-Screen study. We retrieved 305 of these slides for conducting p16/Ki-67 dual-stained cytology, retrospectively. Histology and/or cytology follow-up results were collected until December 2012, providing up to 9 years of follow-up. In this timeframe, all women should have had a follow-up visit (i.e., more than a full round in population-based screening). These follow-up results were retrieved from PALGA, the nationwide registry and network of histological and cytological results in the Netherlands (Casparie et al, 2007). The results of

HPV test triage and p16/Ki-67 triage were compared at baseline and after the entire follow-up period. For each woman, the most severe diagnosis during follow-up was used.

p16/Ki-67 dual-stained cytology. A commercial kit specifically designed for the simultaneous detection of p16 and Ki-67 in cervical cytology preparations was used (CINtec PLUS, Roche mtm laboratories AG, Mannheim, Germany). All slides were de-stained and subjected to p16/Ki-67 dual staining according to the instructions of the manufacturer. All immunostained slides were analysed and scored by an experienced cytotechnologist blinded to all study data. Samples were considered positive when p16 and Ki-67 immunoreactivities were revealed within the same cell (i.e., a cytoplasmic brown staining for p16, together with a nuclear red staining for Ki-67). The presence of at least one dual-stained cell(s) was used as a cutoff to rate the sample as positive for the CINtec PLUS test (Figure 2). As a sub-analysis, we also evaluated the performance of p16/Ki-67 dual-stained cytology at cutoffs of 2–4, 5–49, and 50 or more dual-stained cells.

HPV detection. High-risk HPV testing was performed by HC2 high-risk HPV DNA test in an automated format on a rapid capture system according to manufacturer's instructions (Qiagen). This test uses a cocktail probe to detect 13 high-risk HPV types: 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68. Samples with HC2 outcome of ⩾1 RLU/CO were considered as HPV positive. Results of this analysis, which was carried out at baseline, were used for the comparison with p16/Ki-67 dual-stained cytology in the current study.

Cytology. From all women, a conventional cytological scrape was taken with a cytobrush (Rovers, Oss, The Netherlands). After preparation of the smear on a glass slide, the brush was placed in a vial containing 1 ml Universal Collection Medium (Qiagen) for HPV testing. Cervical cytology results were reported, blinded to the HPV testing results, according to the CISOE-A classification based on Dutch guidelines. This classification is a five-tiered cytology classification system that can be easily converted into either the British or the 2001 Bethesda system (Hanselaar, 2002). Borderline/mildly dyskaryotic corresponds to atypical squamous cells of undetermined significance, atypical squamous cells, and cannot exclude high-grade squamous intraepithelial lesions or low-grade squamous intraepithelial lesions. Moderate- or worse dyskaryosis corresponds to high-grade squamous glandular intraepithelial lesions.

Evaluation of the p16/Ki-67 dual-stained cytology results were compared with the baseline results of the Pap cytology smears.

Histology. Of the women who were referred to colposcopy, colposcopy-directed biopsies were taken from suspicious areas from their cervix according to standard procedures in the Netherlands (Hopman et al, 2000). Histological examination of these biopsies was carried out at local pathology laboratories and classified as normal, CIN grade 1, 2 or 3, or as invasive cancer according to international criteria (Hopman et al, 1995; Anderson 2013). High-grade lesions were all reviewed by two independent pathologists at the VU University Medical Center.

Data and statistical analysis. The association between p16/Ki-67 positivity, HPV positivity and CIN3+ lesions detected within 3 years of follow-up were tested using the χ^2-test. Cervical intraepithelial neoplasia (CIN2+) was used as an additional, secondary outcome. Adjusted estimates for sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and referral rates were calculated from cross-tabulation of test results and adjusted end points. Significant differences between tests with regard to sensitivity, specificity and referral rates were calculated using the McNemar test. The two tests were compared
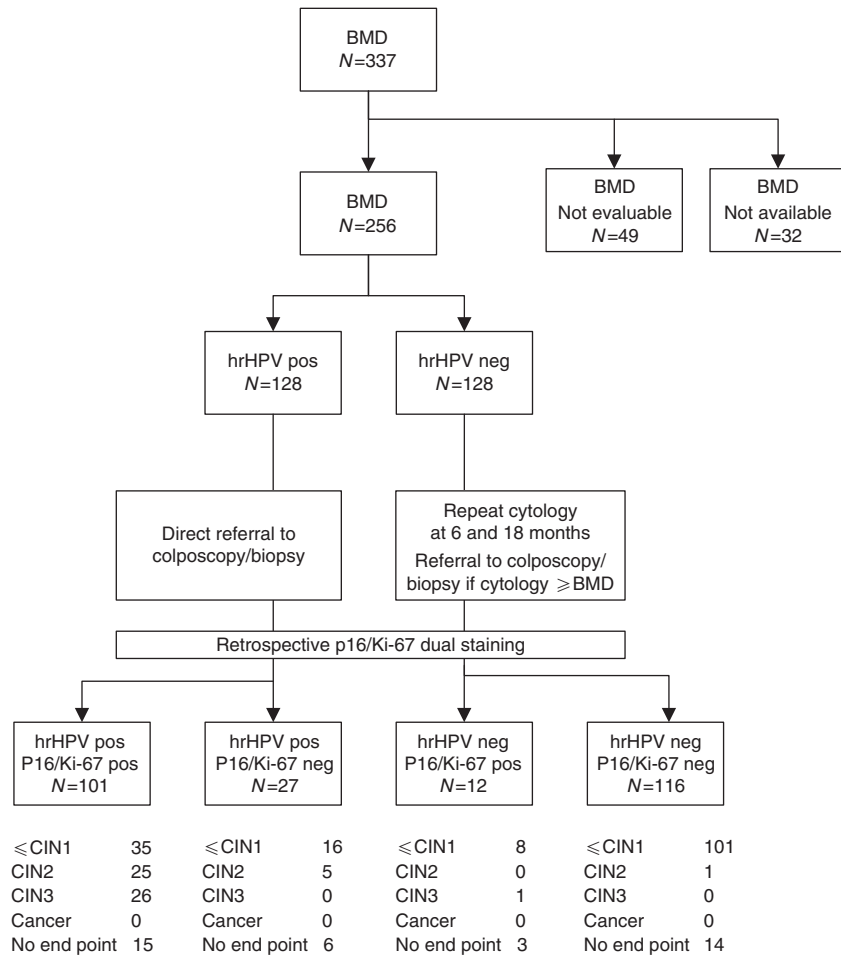
Figure 1. Flowchart design of VUSA-Screen study including p16/Ki-67 dual-stain cytology test results in 256 women with BMD – 3-years follow-up.
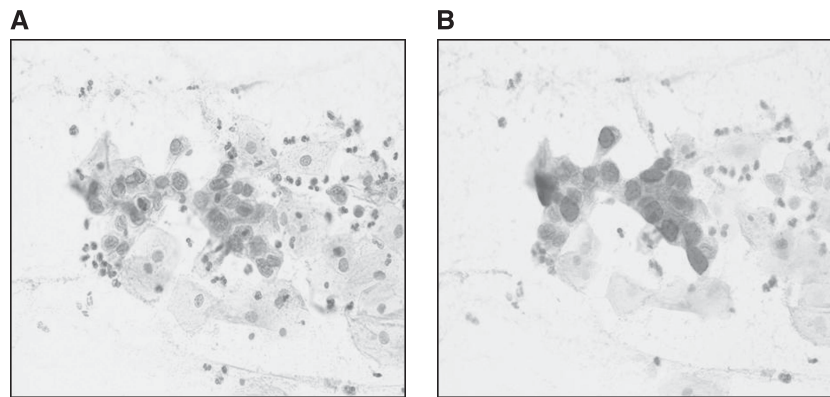

Figure 2. (A) Pap-stained BMD smear (×40) without p16/Ki-67 dual staining. (B) The same smear as in picture A dual stained for p16/Ki-67. Brown cytoplasmic signal for p16 overexpression and red nuclear signal for Ki-67 expression within the same cell points to dysplasia.

with respect to PPV and NPV using the method of Leisenring *et al* (2000). Exact 95% confidence intervals (CIs) were calculated.

Women reached their study-end point if they had a histological outcome or an adjusted end point. Different adjusted end points were defined for HPV-positive and HPV-negative women with BMD, given their different risks of high-grade lesions (i.e., cumulative 3-year CIN3 + risk of 20.4% (HPV-positive women) *vs* 1.3% (HPV-negative women); Rijkaart *et al*, 2010, Katki *et al*, 2013). For both HPV-positive and HPV-negative women with BMD, the adjusted end point was defined as ≤CIN1 when they

had normal cytology and a negative HPV result as a repeat test. Alternatively, HPV-positive BMD women were considered to have ≤CIN1 when they had no HPV test result in follow-up but twice normal cytology follow-up test results. The latter women have a CIN3 + risk of 1.2% that is in the Netherlands sufficiently low to dismiss them from further follow-up. As BMD women with a HPV-negative result have a low intrinsic risk of high-grade disease, they were already categorised as having ≤CIN1 when they had no HPV follow-up result, but had a single smear read as normal cytology within 1 year after the BMD result. Women with

other (combinations of) repeat-test results were considered as having no adjusted end point and were therefore excluded from the analysis.

For the long-term predictive value of both the p16/Ki-67 dual staining and the HPV triage testing, the Kaplan–Meier analysis with the log-rank test was carried out. All women in this study with BMD were followed up until 9 years after inclusion. In this timeframe, all women should have had a follow-up visit (i.e., more than a full round in population-based screening).

Analyses were carried out with IBM SPSS version 20 (International Business Machines Corp., Armonk, NY, USA).

RESULTS

Study population. Of the 337 women with BMD who participated in the VUSA-Screen study, a total of 305 archived conventional cytological slides were available for p16/Ki-67 dual-stained cytology. Following removal of the original cover slide by xylene treatment and subsequent de-staining, slides were stained in two separate, subsequent sets. There were 49 out of the 305 cases excluded from the analysis, as these slides were not evaluable. Owing to inhomogeneous staining, 25 (51%) cases were excluded and 17 (35%) cases were excluded because of background staining. Furthermore, five slides (10%) were broken during processing and two (4%) slides had a very low cellularity. Forty-three out of these 49 slides (88%) were stained as part of the initial staining set. A total of 256 cases were included in the final analysis.

Patients' characteristics. The median age of the women included in this study was 40.0 years (range 29–60 years). Half of the women were HC2 high-risk HPV positive (128/256, 50.0%), whereas the remaining 128 women tested HPV negative. Among the women who attended a follow-up visit, the median follow-up time was 61 months (range 1–103 months). For this analysis, 106/128 (82.8%) of the HPV-positive women had a histological end point.

One HPV-positive woman, having normal cytology and a negative HPV test, was considered to have ⩽CIN1 as an adjusted end point. For the remaining 21 women, no adjusted end points could be determined. Seven of them (7/128; 5.5%) did not have any follow-up test result between inclusion and 3 years later. Another 12 women (9.4%) had one normal cytology follow-up test result without information about their HPV status. They did not meet the criteria for an adjusted end point and were excluded. Finally, two women had a BMD smear as follow-up outcome (1.6%) and thus no adjusted end point as well.

Of the 128 HPV-negative women with BMD, 20 women (15.6%) had a histological end point. Adjusted end points (⩽CIN1) could be established for 91 other HPV-negative women (71.1%); 20 of them based on a double-negative cytology and HPV test result in follow-up, and 71 of them because of a repeat normal cytology test result in follow-up. For the remaining 17 women, no adjusted end points were determined. Fourteen of them did not have any follow-up test result (11.0%) between inclusion and 3 years later. Another three women (2.0%) had a BMD outcome during follow-up.

p16/Ki-67 positivity rates. The overall p16/Ki-67 positivity for women with and without an end point was 44.1% (113/256). The main outcomes between the different groups are presented in Figure 1.

There was a significant trend ($P < 0.001$) for p16/Ki-67 dual-staining positivity and increasing severity of disease: positivity varied from 44.1% (15/34) in women without dysplasia, 58.8% (20/34) in women with CIN1, 80.6% (25/31) in CIN2 to 100.0% (27/27) in women with CIN3. The prevalence of positive p16/Ki-67 test results in the HPV-positive group was 78.9% *vs* 9.4% in the HPV-negative cohort.

Sensitivity and specificity of p16/Ki-67 dual-stained cytology in comparison with HPV testing results. Sensitivities, specificities, PPVs and NPVs as well as referral rates were calculated for p16/Ki-67 and HPV (Table 1).

Table 1. Clinical performance of p16/Ki-67 and HPV testing to detect CIN2+ or CIN3+ in women with BMD using the standard positivity cutoff of ⩾1 dual-stained cell in the first half and alternative cutoffs in the second half

	n1/N1	Sensitivity	(95% CI)	n2/N2	Specificity	(95% CI)	PPV	(95% CI)	NPV	(95% CI)	Referral rate
p16/Ki-67 ⩾1 pos											
CIN2+	52/58	89.7%	78.8–96.1%	117/160	73.1%	65.6–79.8%	54.7%	44.2–65.0%	95.1%	89.7–98.2%	
CIN3+	27/27	100.0%	87.2–100.0%	123/191	64.4%	57.2–71.2%[a]	28.4%	19.6–38.6%[a]	100.0%	97.0–100.0%	43.6%[a]
hrHPV											
CIN2+	56/58	96.6%	88.1–99.6%	109/160	68.1%	60.3–75.3%	52.3%	42.5–62.1%	98.2%	93.6–99.8%	
CIN3+	26/27	96.3%	81.0–99.9%	110/191	57.6%	50.2–64.7%	24.3%	16.5–33.5%	99.1%	95.1–100.0%	49.1%
Alternative cutoffs											
p16/Ki-67 ⩾2 pos											
CIN2+	51/58	87.9%	76.7–95.0%	120/160	75.0%	67.6–81.5%	56.0%	45.2–66.4%	94.5%	89.0–97.8%	
CIN3+	26/27	96.3%	81.0–99.9%	126/191	66.0%	58.8–72.7%	28.6%	19.6–39.0%	99.2%	96.3–100.0%	41.7%
p16/Ki-67 ⩾5 pos											
CIN2+	45/58	77.6%	64.7–87.5%	134/160	83.8%	77.1–89.1%	63.4%	51.1–74.5%	91.2%	85.4–95.2%	
CIN3+	26/27	96.3%	81.0–99.0%	146/191	76.4%	69.8–82.3%	36.6%	25.5–48.9%	99.3%	96.3–100.0%	32.6%
p16/Ki-67 ⩾50 pos											
CIN2+	24/58	41.4%	28.6–55.1%	151/160	94.4%	89.6–97.4%	72.7%	54.5–86.7%	81.6%	75.3–86.9%	
CIN3+	15/27	55.6%	35.3–74.5%	173/191	90.6%	85.5–94.3%	45.5%	28.1–63.6%	93.5%	88.9–96.6%	15.1%

Abbreviations: CI = confidence interval; CIN = cervical intraepithelial (grade 2 or 3 or higher); HPV = human papillomavirus; n1 = number of test positive disease cases; N1 = total number of disease cases; n2 = number of test negative non-disease cases; N2 = total number of non-disease cases; NPV = negative predictive value; PPV = positive predictive value.
[a]Significant difference compared with HPV ($P < 0.05$).

For CIN3 + , the sensitivity of p16/Ki-67 dual-stained cytology was similar to HPV testing (100.0% vs 96.3%; P = 1.00). Also for CIN2 + , sensitivity of both tests were similar (89.7% vs 96.6%; P = 0.219).

Specificity of p16/Ki-67 for CIN3 + was significantly higher in comparison with HPV testing (64.4% vs 57.6%; P = 0.024). Also, the PPV of p16/Ki-67 dual-stained cytology was higher at the CIN3 + threshold (28.4% vs 24.3%; P = 0.013). For CIN2 + , the specificity for p16/Ki-67 dual-stained cytology was comparable to HPV testing (73.1 and 68.1%; P = 0.08). Also, the PPVs for CIN2 + were similar (54.7 and 52.3%; P = 0.396) (Table 1).

Referral rates of BMD women on the basis of p16/Ki-67 dual-stained cytology would have resulted in lower referral rates of these women for colposcopy (43.6% vs 49.1%; P = 0.043) compared with referral on the basis of HPV testing.

Long-term follow-up data. The 5-year cumulative incidence risks (CIRs) of the BMD women stratified by p16/Ki-67 dual-stained cytology results, HPV results or both, are presented in Table 2 with corresponding Kaplan–Meier curves for CIN2 + and CIN3 + in Figure 3.

Women with a BMD test result at baseline and a negative HPV test had a 5-year CIR of 0.8% (95% CI 0.0–2.4%) for CIN3 + , whereas in those with a negative p16/Ki-67 dual-stain test result no CIN3 lesions were found. When analysed for CIN2 + outcome, women with BMD and a negative HPV test had a 5-year CIR of 1.6%, whereas for those with a negative p16/Ki-67 test result, the five-year CIR was 4.3% (95% CI: 1.0–7.6%). These risks were not significantly different (P = 0.326).

The highest 5-year CIR was found for women with BMD and a positive HPV, as well as a positive p16/Ki-67 dual-stained cytology test result. For CIN3 + this risk was 26.9% (95% CI: 18.1–35.7%) and for CIN2 + this risk was 53.3% (95% CI: 43.7–63.7%). Lowest 5-year CIRs were achieved for women with a negative HPV and a negative p16/Ki-67 dual-stain result. For CIN2 + , this CIR was 0.9% (0.0–2.7%).

The 5-year CIR for CIN2 + for women with BMD, a positive HPV test and a negative p16/Ki-67 dual-stained cytology result was 18.8% (95% CI: 3.9–33.7%). A negative HPV test and a positive p16/Ki-67 dual-stained cytology result gave a 5-year CIR of 9.1% (0.0–26.2%) for CIN2 + .

p16/Ki-67 dual-stained cytology testing at different cutoffs. In all previous studies, a cutoff of 'one or more' dual-stained-positive cells was applied for calling a slide positive for the p16/Ki-67 dual-stained cytology test. This threshold was also used in this study. In addition, we tested the effects of raising the positivity threshold on the sensitivity and specificity.

Raising the cutoff to two or more dual-stained-positive cells resulted in an almost similar specificity for CIN3 + (66.0%) and CIN2 + (75.0%) compared with a cutoff of one or more dual-stained-positive cells. At this cutoff, one CIN3 lesion and seven CIN2 lesions were missed. Raising the cutoff further to five or more positive cells resulted in an increased specificity for CIN3 + (76.4%) and CIN2 + (83.8%). However, although at this cutoff, no more CIN3 lesions were missed compared with a cutoff of two or more dual-stained-positive cells or compared with HPV triaging, the sensitivity for CIN2 + was reduced (77.6%). Details of this analysis are shown in Table 1.

DISCUSSION

This is one of the first studies that evaluates both the cross-sectional and the longitudinal performance of p16/Ki-67 dual-stained cytology as a triage tool for women with BMD compared with HPV testing. Results indicate that for triaging women with BMD cytology, p16/Ki-67dual staining, even on conventional slides, has a small but significantly higher specificity for CIN3 + than HPV testing (64.4% vs 57.6%). For both tests, the sensitivity for CIN3 + was similar (100.0% vs 96.3%). At CIN2 + level, p16/Ki-67 dual staining and HPV testing reached similar specificity (73.1% vs 68.1%) and similar sensitivity (89.7% vs 96.6%) levels. In addition, the results show that fewer women would be referred to colposcopy in case of dual staining compared with HPV testing (43.6% vs 49.1%).

Our results are generally comparable with data from other cross-sectional studies (Schmidt et al, 2011; Wentzensen et al, 2012; Waldstrom et al, 2013), which have demonstrated that p16/Ki-67 dual staining is superior to HPV testing at baseline with respect to specificity and similar in terms of sensitivity for CIN3 + and CIN2 + . Wentzensen et al (2012) showed that the specificity of p16/Ki-67 dual staining for CIN2 + and CIN3 + was significantly improved compared with HPV testing. In this study, differences in specificity were only significant for CIN3 + . The p16/Ki-67 detection rate for CIN2 was 80.6% (25/31). Six CIN2 detected within 3 years after baseline were negative for p16/Ki-67 at baseline. It is unclear whether these represent regressive or persistent lesions. In comparison, there was only one woman with CIN2 who tested negative for HPV at baseline. Mesher et al (2013) compared different triage strategies, including p16 cytology among

Table 2. Cumulative 5-year incidence risks stratified by p16/Ki-67 dual-stained cytology (cutoff ⩾1 dual-stained cell) or HPV or both

Baseline group	Number of women	CIR CIN3 + t = 60	95% CI	CIR CIN2 + t = 60	95% CI
BMD p16/Ki-67 pos	113	25.2%	17.0–33.4	49.6%	40.0–59.2
BMD p16/Ki-67 neg	143	0.0%		4.3%	1.0–7.6
BMD HPV pos	128	21.1%	13.8–28.4	46.2%	37.4–55.0
BMD HPV neg	128	0.8%	0.0–2.4	1.6%	0.0–4.0
BMD HPV pos p16/Ki-67 pos	101	26.9%	18.1–35.7	53.7%	43.7–63.7
BMD HPV pos p16/Ki-67 neg	27	0.0%		18.8%	3.9–33.7
BMD HPV neg p16/Ki-67 pos	12	9.1%	0.0–26.2	9.1%	0.0–26.2
BMD HPV neg p16/Ki-67 neg	116	0.0%		0.9%	0.0–2.7
BMD total	256	11.0%	7.1–14.9	24.1%	18.8–29.4

Abbreviations: BMD = borderline/mildly dyskaryosis; CI = confidence interval; CIR = cumulative incidence risk; CIN = cervical intraepithelial neoplasia (grade 2 or 3); HPV = human papillomavirus; neg = negative; pos = positive; t = follow-up time in months.
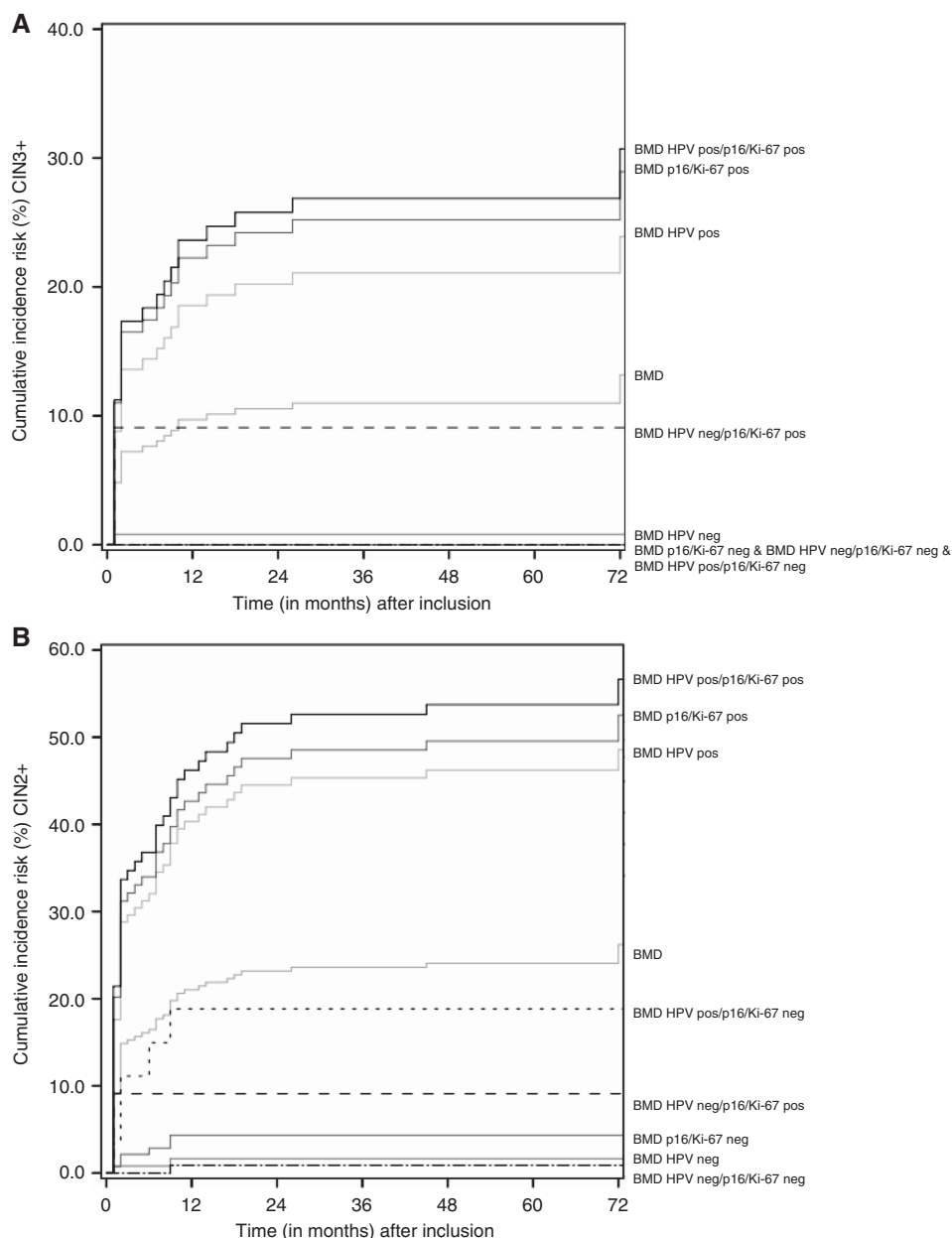
Figure 3. (A) Cumulative incidence risk (CIR) of CIN3+ in women with BMD Pap cytology stratified by p16/Ki-67 dual-stained cytology or HPV or both. **(B)** Cumulative incidence risk (CIR) of CIN2+ in women with BMD Pap cytology stratified by p16/Ki-67 dual-stained cytology or HPV or both. For comparison, the cumulative disease in the whole cohort is also shown. BMD = borderline/mildly dyskaryotic cytology smears; HPV = human papillomavirus; CIN2+ = cervical intraepithelial neoplasia grade 2 or worse.

women who were referred for colposcopy with low-grade cytology. For CIN2+ as well as CIN3+, that study also found an improved specificity for p16 cytology compared with HPV testing (for CIN2+, 50.7% *vs* 23.3%). They also reported the disadvantage of missed lesions (for CIN3+: 15–30%), where they advised early recall for women with a negative test result for p16. One possibility why in this study the specificity for p16/Ki-67 dual staining was only significant for CIN3+ might be an overcall of CIN2 lesions on biopsies.

To estimate the utility of p16/Ki-67 dual staining as a triage test, we should also compare results with present cytology triage in the 5-year based screening programme, where a remaining risk of CIN3+ of 1.2% is regarded to be acceptable for women with BMD at baseline, and normal cytology at 6 and 18 months follow-up (Rijkaart *et al*, 2012b). In this study, none of the women with BMD and a negative dual-stained cytology result developed CIN3+ over

a period of up to 9 years. However, the 5-year cumulative CIN2+ incidence risk of a woman with BMD and a negative dual-stain result was 4.3% (95% CI: 1.0–7.6%) compared with 0.8% (95% CI: 0.0–2.4%) in women with BMD and a negative HPV test result. Because in clinical practice women with a CIN2+ lesion require treatment, we would advise women with BMD cytology and a negative p16/Ki-67 dual staining result to attend a follow-up visit within 2 years.

In order to determine the utility of the dual staining, more considerations can be made. p16/Ki-67 dual staining should not require morphological interpretation. This is an improvement of the evaluation of p16 alone, which is still morphology dependent. However, a trained cytotechnologist or pathologist will also be influenced by the morphology as it is their expertise. Therefore, in contrast to the HPV test, the evaluation of the dual staining is not totally objective. It is a consideration of the clinician whether a

cell-based triage with p16/Ki-67 with 5.5% less referrals or an objective molecular HPV test with a somewhat higher referral rate is preferred.

The New Technologies for Cervical Cancer Screening trial investigated the usefulness of an earlier version of the p16 cytology test for triaging HPV-positive women that was based on the p16 biomarker only (i.e., no p16/Ki-67 dual staining) in combination with morphology interpretation of immunoreactive cells (Carozzi et al, 2013). The authors concluded that p16 testing could be used as a triage test for HPV-positive women and advised that HPV-positive women with a negative p16 cytology test result should be re-screened after 2–3 years, given a 3-year CIN2 + risk of 5.2% (95% CI: 3.4–7.0%). Our data are similar to The New Technologies for Cervical Cancer Screening at the CIN3 + threshold, but a relevant number of CIN2 lesions remained undetected by p16/Ki-67 dual-stained cytology. This leads to a 5-year CIR for CIN2 + of 18.8% among HPV-positive women with BMD. Therefore, a follow-up of HPV-positive, p16/Ki-67 dual-stained cytology negative women with initial BMD cytology results after 2 years appears to be more appropriate based on the data of this study.

Currently, the p16/Ki-67 dual staining cutoff is one or more dual-stained-positive cells. As reading p16/Ki-67 staining cells requires experienced cytotechnicians/pathologists, and in order to achieve a higher reproducible cutoff, we also evaluated the performance of p16/Ki-67 dual-stained cytology at different cutoffs. We measured a higher specificity for CIN2 + and CIN3 + (83.8% and 76.4%) at a p16/Ki-67 positivity cutoff of five or more dual-stained-positive cells when used in this BMD cohort. At this cutoff, one CIN3 lesion and seven more CIN2 lesions would have been missed when compared with using a cutoff of one or more dual-stained cells. Therefore, given the high number of missed CIN2 lesions and one CIN3 lesions that is missed, we did not consider adjustment of the cutoff as improvement.

The strengths of this study were the large sample size, the long follow-up of up to 9 years and the setting within a population-based screening programme. An important limitation of this study was the inability to assess 49 out of 305 samples due to inhomogeneous re-staining or background staining of these conventional smears.

Dual staining of the conventional Pap cytology slides in this study was performed using a de-staining and dual-staining protocol of two separate slide sets, in each of which about half of the samples were stained. The de-staining protocol step of the conventional slides in the initial set of slides processed may have been too long, which led to excessive background staining. In that initial set, 43 of the total of 49 non-evaluable slides were present. In the second set, an optimisation of the de-staining protocol step with respect to the length of the de-staining procedure was implemented to prevent avoidable background staining, and thus inadequate staining results. Because the two batches were blindly assembled, we have no indication that re-staining had a biasing effect on our outcomes. Also, the sensitivity and specificity for the two batches were similar. It is important to note that these initial technical issues related to the de-staining protocol are specific study-related limitations owing to the use of archived, Pap-stained conventional smears that are not applicable in routine clinical practice, where either liquid-based cytology or co-collected conventional smears or fresh Pap-stained conventional smears might be used.

At last, there is a possibility of verification bias as, because of the study protocol, not all of the 256 women had a histological end point. Although studies have shown that a negative HPV test and normal cytology is associated with an extremely low risk for CIN3 + (Cuzick et al, 2006; Bulkmans et al, 2007; Mayrand et al, 2007; Dillner et al, 2008; Rijkaart et al, 2012a), we cannot exclude yet that this may have biased our estimations (Ransohoff and Feinstein, 1978; Choi, 1992).

In summary, our study shows that women with BMD cytology result can be triaged with p16/Ki-67 dual-stained cytology testing, with less referrals compared with HPV triage testing. Based on the cumulative 5-year CIN2 + risk of 4.3% for women with BMD Pap cytology that show negative p16/Ki-67 dual-stain results, we advise these women to a follow-up visit within 2 years.

REFERENCES

Anderson MC (2013) Premalignant and malignant squamous lesions of the cervix. In Obstetrical and Gynaecological Pathology, Fox H, Wells M Haines Taylor (eds) pp 292–297. Churchill Livingstone: New York, USA.

Arbyn M, Sasieni P, Meijer CJ, Clavel C, Koliopoulos G, Dillner J (2006) Chapter 9: Clinical applications of HPV testing: a summary of meta-analyses. Vaccine 24(Suppl 3): S3/78–89.

Berkhof J, de Bruijne MC, Zielinski GD, Bulkmans NW, Rozendaal L, Snijders PJ, Verheijen RH, Meijer CJ (2006) Evaluation of cervical screening strategies with adjunct high-risk human papillomavirus testing for women with borderline or mild dyskaryosis. Int J Cancer 118(7): 1759–1768.

Bulk S, Bulkmans NW, Berkhof J, Rozendaal L, Boeke AJ, Verheijen RH, Snijders PJ, Meijer CJ (2007) Risk of high-grade cervical intra-epithelial neoplasia based on cytology and high-risk HPV testing at baseline and at 6-months. Int J Cancer 121(2): 361–367.

Bulkmans NW, Berkhof J, Rozendaal L, van Kemenade FJ, Boeke AJ, Bulk S, Voorhorst FJ, Verheijen RH, van GK, Boon ME, Ruitinga W, van BM, Snijders PJ, Meijer CJ (2007) Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. Lancet 370(9601): 1764–1772.

Carozzi F, Gillio-Tos A, Confortini M, Del MA, Sani C, De ML, Girlando S, Rosso S, Naldoni C, Dalla PP, Zorzi M, Giorgi-Rossi P, Segnan N, Cuzick J, Ronco G (2013) Risk of high-grade cervical intraepithelial neoplasia during follow-up in HPV-positive women according to baseline p16-INK4A results: a prospective analysis of a nested substudy of the NTCC randomised controlled trial. Lancet Oncol 14(2): 168–176.

Casparie M, Tiebosch AT, Burger G, Blauwgeers H, van de PA, van Krieken JH, Meijer GA (2007) Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. Cell Oncol 29(1): 19–24.

Choi BC (1992) Sensitivity and specificity of a single diagnostic test in the presence of work-up bias. J Clin Epidemiol 45(6): 581–586.

Cuzick J, Clavel C, Petry KU, Meijer CJ, Hoyer H, Ratnam S, Szarewski A, Birembaut P, Kulasingam S, Sasieni P, Iftner T (2006) Overview of the European and North American studies on HPV testing in primary cervical cancer screening. Int J Cancer 119(5): 1095–1101.

Denise ZG, Snijders PJ, Rozendaal L, Voorhorst FJ, Runsink AP, de Schipper FA, Meijer CJ (2001) High-risk HPV testing in women with borderline and mild dyskaryosis: long-term follow-up data and clinical relevance. J Pathol 195(3): 300–306.

Denton KJ, Bergeron C, Klement P, Trunk MJ, Keller T, Ridder R (2010) The sensitivity and specificity of p16(INK4a) cytology vs HPV testing for detecting high-grade cervical disease in the triage of ASC-US and LSIL pap cytology results. Am J Clin Pathol 134(1): 12–21.

Dietsch E, Davies C (2007) The nocebo effect for women in waiting. Collegian 14(3): 9–14.

Dillner J, Rebolj M, Birembaut P, Petry KU, Szarewski A, Munk C, de SS, Naucler P, Lloveras B, Kjaer S, Cuzick J, van BM, Clavel C, Iftner T (2008) Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. BMJ 337: a1754.

Hanselaar AG (2002) Criteria for organized cervical screening programs. Special emphasis on The Netherlands program. Acta Cytol 46(4): 619–629.

Health Council of the Netherlands (2011) Population screening for cervical cancer. [Report]. The Hague: Health Council of the Netherlands. Publication no 2011/07. publication no. 2011/07. 2011.

Hopman EH, Rozendaal L, Voorhorst FJ, Walboomers JM, Kenemans P, Helmerhorst TJ (2000) High risk human papillomavirus in women with normal cervical cytology prior to the development of abnormal cytology and colposcopy. BJOG 107(5): 600–604.

Hopman EH, Voorhorst FJ, Kenemans P, Meyer CJ, Helmerhorst TJ (1995) Observer agreement on interpreting colposcopic images of CIN. Gynecol Oncol 58(2): 206–209.

Katki HA, Schiffman M, Castle PE, Fetterman B, Poitras NE, Lorey T, Cheung LC, Raine-Bennett T, Gage JC, Kinney WK (2013) Five-year risks of CIN 3 + and cervical cancer among women with HPV testing of ASC-US Pap results. *J Low Genit Tract Dis* **17**(5 Suppl 1): S36–S42.

Kelly RS, Patnick J, Kitchener HC, Moss SM (2011) HPV testing as a triage for borderline or mild dyskaryosis on cervical cytology: results from the Sentinel Sites study. *Br J Cancer* **105**(7): 983–988.

Kitchener HC, Denton K, Soldan K, Crosbie EJ (2013) Developing role of HPV in cervical cancer prevention. *BMJ* **347**: f4781.

Leisenring W, Alonzo T, Pepe MS (2000) Comparisons of predictive values of binary medical diagnostic tests for paired designs. *Biometrics* **56**(2): 345–351.

Loghavi S, Walts AE, Bose S (2013) CINtec(R) PLUS dual immunostain: a triage tool for cervical pap smears with atypical squamous cells of undetermined significance and low grade squamous intraepithelial lesion. *Diagn Cytopathol* **41**(7): 582–587.

Lytwyn A, Sellors JW, Mahony JB, Daya D, Chapman W, Howard M, Roth P, Lorincz AT, Gafni A, Walter SD (2003) Adjunctive human papillomavirus testing in the 2-year follow-up of women with low-grade cervical cytologic abnormalities: a randomized trial and economic evaluation. *Arch Pathol Lab Med* **127**(9): 1169–1175.

Mayrand MH, Duarte-Franco E, Rodrigues I, Walter SD, Hanley J, Ferenczy A, Ratnam S, Coutlee F, Franco EL (2007) Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. *N Engl J Med* **357**(16): 1579–1588.

Melnikow J, Kuppermann M, Birch S, Chan BK, Nuovo J (2002) Management of the low-grade abnormal Pap smear: What are women's preferences? *J Fam Pract* **51**(10): 849–855.

Mesher D, Szarewski A, Cadman L, Austin J, Ashdown-Barr L, Ho L, Terry G, Young M, Stoler M, Bergeron C, McCarthy J, Wright C, Liddle S, Soutter WP, Lyons D, Cuzick J (2013) Comparison of human papillomavirus testing strategies for triage of women referred with low-grade cytological abnormalities. *Eur J Cancer* **49**(9): 2179–2186.

Ransohoff DF, Feinstein AR (1978) Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* **299**(17): 926–930.

Rijkaart DC, Berkhof J, Rozendaal L, van Kemenade FJ, Bulkmans NW, Heideman DA, Kenter GG, Cuzick J, Snijders PJ, Meijer CJ (2012a) Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: final results of the POBASCAM randomised controlled trial. *Lancet Oncol* **13**(1): 78–88.

Rijkaart DC, Berkhof J, van Kemenade FJ, Coupe VM, Rozendaal L, Heideman DA, Verheijen RH, Bulk S, Verweij W, Snijders PJ, Meijer CJ (2012b) HPV DNA testing in population-based cervical screening (VUSA-Screen study): results and implications. *Br J Cancer* **106**(5): 975–981.

Rijkaart DC, Berkhof J, van Kemenade FJ, Rozendaal L, Verheijen RH, Bulk S, Herreilers ME, Verweij WM, Snijders PJ, Meijer CJ (2010) Comparison of HPV and cytology triage algorithms for women with borderline or mild dyskaryosis in population-based cervical screening (VUSA-screen study). *Int J Cancer* **126**(9): 2175–2181.

Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla PP, Del MA, Ghiringhello B, Girlando S, Gillio-Tos A, De ML, Naldoni C, Pierotti P, Rizzolo R, Schincaglia P, Zorzi M, Zappa M, Segnan N, Cuzick J (2010) Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. *Lancet Oncol* **11**(3): 249–257.

Schmidt D, Bergeron C, Denton KJ, Ridder R (2011) 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* **119**(3): 158–166.

Waldstrom M, Christensen RK, Ornskov D (2013) Evaluation of p16(INK4a)/Ki-67 dual stain in comparison with an mRNA human papillomavirus test on liquid-based cytology samples with low-grade squamous intraepithelial lesion. *Cancer Cytopathol* **121**(3): 136–145.

Wentzensen N, Schwartz L, Zuna RE, Smith K, Mathews C, Gold MA, Allen RA, Zhang R, Dunn ST, Walker JL, Schiffman M (2012) Performance of p16/Ki-67 immunostaining to detect cervical cancer precursors in a colposcopy referral population. *Clin Cancer Res* **18**(15): 4154–4162.