

Defining Optimal Triage Strategies for hrHPV Screen-Positive Women—An Evaluation of HPV 16/18 Genotyping, Cytology, and p16/Ki-67 Cytoimmunochemistry



Grazyna A. Stanczuk¹, Gwen J. Baxter¹, Heather Currie², William Forson², James R. Lawrence¹, Kate Cuschieri³, Allan Wilson⁴, Lynne Patterson⁴, Lynn Govan⁴, Janice Black⁴, Tim Palmer⁵, and Marc Arbyn⁶

Abstract

Background: Several options for the triage of high-risk HPV screen-positive (hrHPV⁺) women were assessed.

Methods: This study incorporated CIN2⁺ cases and controls, all of whom tested hrHPV⁺ and whose results of liquid-based cytology (LBC), HPV16/18 genotyping, and p16/Ki67 cytoimmunochemistry were available. Sensitivity and specificity for the CIN2⁺ of these triage tests were evaluated.

Results: Absolute sensitivities of HPV 16/18 typing, LBC, and p16/Ki-67 cytoimmunochemistry for CIN2⁺ detection were 61.7%, 68.3%, and 85.0% for women with hrHPV⁺ clinician-taken samples. Respective specificities were 70.5%, 89.1%, and 76.7%. The absolute accuracy of the triage tests was similar for women with a hrHPV⁺ self-sample. P16/Ki-67 cytoimmunochemistry was significantly more sensitive than LBC although significantly less specific.

Conclusions: All three single-test triage options, if positive, exceed the threshold of 20% risk at which colposcopy would be indicated. However, none of them conferred a post-test probability of CIN2⁺ <2%; which would permit routine recall. P16/Ki-67 cytoimmunochemistry on HPV16/18 negative women had a post-test probability of CIN2⁺ of 1.7% and 0.6% if also LBC negative.

Impact: This is one of the few studies to directly compare the performance of triage strategies of hrHPV⁺ women, in isolation and combinations. It is the only study assessing triage strategies in women who test hrHPV⁺ in self-taken vaginal samples. A combined triage option that incorporated HPV 16/18 typing prior to p16/ki-67 cytoimmunochemistry in HPV 16/18-negative women yielded a post-test probability of CIN2⁺ of >20%, whereas women who tested negative had a probability of CIN2⁺ of <2%. *Cancer Epidemiol Biomarkers Prev*; 26(11); 1629–35. ©2017 AACR.

Introduction

Current evidence indicates that high-risk human papillomavirus (hrHPV) testing is more effective in reducing the burden of cervical intraepithelial neoplasia grade 3 (CIN3) and cervical cancer compared with cytology (1). Consequently, international guidelines support the approach of using HPV molecular testing as the optimal modality for primary cervical screening (2–4). Consensus around what the optimal "second line" triage test(s)

for HPV-positive women should be is, however, less apparent. This is reflected in the variety of triage algorithms proposed across countries committed to the rollout of HPV-based primary screening (4, 5). The optimal triage of screen-positive women is one of the most important considerations in this new era of cervical screening given the high prevalence of transient HPV infection and the implications of unnecessary referrals for colposcopy. Performing cytology on hrHPV-positive women has been recommended as a triage strategy (6, 7). However, even in organized settings, the reliance on morphology and subjective skills may limit the overall effectiveness of this approach—particularly in immunized populations (8, 9).

In addition, data indicate that up to two thirds of HPV-positive women will be cytology negative, thus creating a category of women who need further follow-up (7, 10). The ideal triage test would create a scenario where negative women could be returned to routine recall while positive women could be referred to colposcopy with robust cause.

Other triage options are HPV 16/18 genotyping and detection of the cyclin-dependent kinase inhibitor protein—p16INK4a either in isolation or in combination with Ki-67; the p16/Ki-67 dual-stained cytology is available commercially as CINtec PLUS (Roche Molecular Systems; refs. 10–15). The rationale for these approaches are based on the higher risk that HPV 16 and 18 confer for the development of CIN3⁺ compared with other HR-HPV

¹Department of Research and Development, Dumfries and Galloway Royal Infirmary, Dumfries, United Kingdom. ²Department of Obstetrics and Gynaecology, Dumfries and Galloway Royal Infirmary, Dumfries, United Kingdom. ³Scottish HPV Reference Laboratory, Royal Infirmary of Edinburgh, Edinburgh, United Kingdom. ⁴Department of Pathology, Monklands Hospital, Airdrie, United Kingdom. ⁵Department of Pathology, NHS Highland, Inverness, United Kingdom. ⁶Unit of Cancer Epidemiology/Belgian Cancer Centre, Scientific Institute of Public Health, Brussels, Belgium.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

Corresponding Author: Grazyna A. Stanczuk, Dumfries and Galloway Royal Infirmary, Mac Aulay Rd, Stornoway, Western Isles, HI 2AF. Phone: 4418-5170-4704; Fax: 4418-5170-8235; E-mail: grazyna.stanczuk@nhs.net

doi: 10.1158/1055-9965.EPI-17-0534

©2017 American Association for Cancer Research.

types and the observation that markers of abnormal cell proliferation, that are detectable through cytoimmunochemistry are indicative of potentially transforming infections (16, 17). There is evidence to indicate that the specificity and positive predictive value (PPV) of both these approaches exceeds that of hrHPV testing when used to stratify the risk among women with minor abnormalities (18). A recent triage study conducted in the United States demonstrated higher specificity of p16/Ki-67 dual-stained cytology compared with conventionally stained cytology (where cytologists were aware of the HPV status) for the triage of HPV-positive women (13). However, few studies have compared p16/Ki-67 dual-stained cytology, HPV 16/18 typing, and cytology concurrently within the same population—in the setting of primary screening. Moreover, there are few reports where combinations of the three triage approaches have been considered and none where the performance of p16/Ki-67 dual-stained cytology has been assessed as a triage of women who tested hrHPV positive in a self-taken screening sample.

Previously, we assessed the performance of hrHPV DNA testing using the cobas 4800 HPV assay on clinician-collected cervical samples and self-collected vaginal samples in more than 5,000 women recruited to the Papillomavirus Dumfries and Galloway (PaVDaG) study (19). This study was nested into the PaVDaG framework with the objective of comparing the accuracy of HPV16/18 genotyping, liquid-based cytology (LBC), p16/Ki-67 dual-stained cytology and combinations thereof, for the detection of CIN2⁺ in a sample of hrHPV-positive women in one round of screening. In addition, the accuracy of triage with p16/Ki-67 dual-stained cytology and other markers in women who were hrHPV positive on a self-collected vaginal sample was also assessed.

Material and Methods

Overview of PaVDaG cohort/study

A detailed description of the PaVDaG cross-sectional study can be found in Stanczuk and colleagues (19). Briefly, women ages 20 to 60 attending for routine cervical screening were invited to participate in the study and asked to provide a routine LBC taken by a clinician, a self-taken vaginal swab, and a random void urine specimen. HrHPV testing on all specimens was performed by cobas 4800 PCR DNA test that detects 14 hrHPV types and provides a read out that includes separation of HPV 16/18. Cervical screening history and follow-up information was accessed through the national Scottish Cytology Call-Recall System (SCCRS). The British Society for Clinical Cytopathology reporting guidelines and CIN nomenclature were used to classify cytology findings and histologic outcomes respectively (20, 21). Histology was performed as per NHS standard care—review histology was not done unless routinely indicated. Management of women with abnormal cytology results was performed according to national guidelines and complemented with additional diagnostic interventions for hrHPV-positive women who were cytology negative as described previously (19–22).

Subset of PaVDaG cohort used to evaluate the diagnostic accuracy of tests

A diagnostic test accuracy study was set up that incorporated 61 cases with histologic-confirmed CIN2⁺ and 279 controls (\leq CIN1), all being hrHPV positive and selected at random from the PaVDaG study. Controls incorporated LBC samples from women with at least two cytology negative screening results

(3 years apart) and no previous evidence of CIN2⁺. Women who had two follow-up negative LBC results after a borderline squamous cytology result as well as three negative LBC results after low-grade cytology and no previous evidence of CIN2⁺ were also eligible for inclusion as controls, as were women who had normal colposcopy without biopsy or who had biopsies showing \leq CIN1. All selected cases and controls were submitted for p16/Ki-67 double staining cytology. In addition, conventionally stained LBC and HPV16/18 results were available for all the women. A total of 61 CIN2⁺ cases and 279 control women were selected at random from the whole study population. The accuracy of the aforementioned triage tests was also assessed for 57 CIN2⁺ cases and 335 controls associated with self-taken samples.

For both, cervical and self-taken samples, the sensitivity and specificity for CIN2⁺ and CIN3⁺ of p16/Ki-67 dual-stained cytology, HPV 16/18 genotyping and cytology were assessed (as a triage of hrHPV-positive women), as single or combined test strategies.

P16/Ki-67 cytoimmunochemistry

New slides were made for the p16/Ki-67 dual-stained cytology analysis using the CINtec PLUS Cytology Kit (Roche MTM Laboratories) from residual LBC material. CINtec PLUS interpretation was blinded to the hrHPV and cytology results. Samples with one or more cervical epithelial cells that simultaneously showed brown cytoplasmic immunostaining (p16) and red nuclear immunostaining (Ki-67) were classified as positive regardless of the morphologic appearance of the cells. Slides with excessive background staining were considered not evaluable and excluded from analysis.

Ethical approval

The study was approved by the NHS West of Scotland Research Ethics Service (ref: 12/WS/0085). The study sponsor was NHS Dumfries and Galloway. Roche provided cobas 4800 collection and testing kits and CINtec PLUS kits but was not involved in study design and all analysis was performed independently.

Data analysis and power calculation

Absolute accuracy (for detection of CIN2⁺) of HPV16/18 genotyping, LBC and p16/Ki-67 dual-stained cytology was measured in terms of sensitivity, specificity. PPV and compliment of the negative predictive value (cNPV) = 1-NPV were computed using the prevalence of CIN2⁺ among hrHPV⁺ women observed in the PaVDaG screening study (16.2%; ref. 19). The number of tests used in a particular triage option, the number of colposcopies per 1,000 hrHPV⁺ patients, as well as the number of referrals needed to find one CIN2⁺ case [NNR (number needed to refer = 1/PPV)] were also assessed as measures of cost and efficiency.

Relative sensitivity and specificity of each triage test or combination of triage tests (concurrent or sequential) compared with another was also determined. The McNemar test was used to assess concordance between tests. Statistical significance was defined as $P < 0.05$ and the CI not spanning 1.

The utility of the triage tests can be displayed on pre-test–post-test probability (PPP) plots (23). The left y -axis gives the prevalence of CIN2⁺ in the given population before triage testing (pre-test probability) and the right y -axis gives the post-test probability (PPV) of CIN2⁺ when the triage test is positive or negative (cNPV). Decision thresholds are defined by colored zones, where the red zone indicates immediate referral for colposcopy (defined as a risk

of CIN2⁺ of >20%), the yellow zone indicates further surveillance (risk of CIN2⁺ between 2% and 20%), and the green zone indicates routine screening/recall (risk CIN2⁺ <2%). Knowledge of the positive hrHPV status has been demonstrated to increase sensitivity of cytologic interpretation (24). An assumption was made that p16/Ki-67 dual-stained cytology would allow more objective interpretation and increase specificity of triage without affecting the sensitivity, compared with traditional cytology at cutoff value of borderline squamous changes (13). A matched design was used for sample size computation (25). We hypothesized noninferior sensitivity accepting a minimal value for the lower 95% confidence interval bound of the relative sensitivity of 0.90, which yielded a sample size of 60 CIN2⁺ cases of CIN2⁺ as per the validation protocol of Meijer and colleagues (3). The number of nondiseased subjects (\leq CIN1) needed was based on an assumed difference in specificity between 71% (double staining) and 60% (cytology at ASC-US⁺) with a power of 90%, which yielded 230 subjects.

Results

Absolute accuracy of triage tests to detect CIN2⁺

The absolute accuracy of the triage tests (LBC, p16/Ki-67 dual-stained cytology and HPV 16/18 typing) in clinician-taken samples for the detection of CIN2⁺ are presented singly and in various combinations in Table 1. Unless otherwise stated, "LBC" is based on a threshold of borderline squamous changes. As stand-alone assays, the triage with the highest sensitivity was p16/Ki-67 dual-stained cytology: 85% (95% CI, 73–93) followed by LBC: 68% (95% CI, 55–80) and HPV 16/18 typing: 62% (95% CI, 48–74). Comparatively, the triage test with the highest specificity was LBC: 89% (95%CI, 85–93) followed by p16/Ki-67 dual-stained cytology: 76.7% (95% CI, 71.1–81.8), and HPV 16/18 typing: 70.5% (95% CI, 64.6–76.0). The PPV of LBC was the highest at 54.9%

compared to p16/Ki-67 dual-stained cytology and HPV 16/18 typing (41.4% and 28.8%, respectively). HPV 16/18 typing generated the greatest number of colposcopies and the highest number of referrals to detect one CIN2⁺.

The corresponding accuracy parameters for triage of women with an hrHPV-positive self-sample were similar (see Supplementary Table S1).

Relative accuracy of triage tests to detect CIN2⁺

The relative sensitivity and specificity of p16/Ki-67 dual-stained cytology and HPV 16/18 typing compared with LBC at three thresholds [borderline (ASC-US), low-grade dyskaryosis and high-grade dyskaryosis] is presented in Table 2 for women who were hrHPV positive on clinician-taken samples. P16/Ki-67 dual-stained cytology was more sensitive than LBC and HPV16/18 genotyping at all thresholds (p McN < 0.05; p McN = 0.0060). However, relative specificity of p16/Ki-67 dual-stained cytology versus LBC was significantly lower at all thresholds (p McN < 0.001 for all). The relative sensitivity of HPV 16/18 typing was not significantly different to LBC at a threshold of borderline and low grade squamous changes although it was significantly higher than LBC at a threshold of high-grade dyskaryosis (p McN < 0.0001). The specificity of HPV 16/18 typing was lower than that of LBC at all thresholds (P = < 0.001)

Relative accuracy findings for triage of women who were hrHPV positive on a self-sample were slightly different in relation to the specificity of p16/Ki-67 dual-stained cytology which was significantly higher than HPV 16/18 typing (see Supplementary Table S2).

Combination of triage tests

A comprehensive list of options with two or three triage tests used in parallel or sequentially is available for cervical samples

Table 1. Accuracy and diagnostic efficiency of triage tests to detect CIN2⁺ among women with an hrHPV⁺ clinician-taken cervical sample

Triage tests	Sensitivity (95% CI)	Specificity (95% CI)	PPV	cNPV	Triage tests		Colposcopies	NNR
					(1) ^a	(2) ^b		
One single triage test								
Dual-stained cytology	85.0% (73.4%–92.9%)	76.7% (71.1%–81.8%)	41.4%	3.6%	1,000	1,000	333	2.4
HPV16/18	61.7% (48.2%–73.9%)	70.5% (64.6%–76.0%)	0.288	0.095	1,000	0	347	3.5
LBC (ASC-US ⁺) ^c	68.3% (55.0%–79.7%)	89.1% (84.7%–92.7%)	0.549	0.064	1,000	1,000	202	1.8
LBC (LSIL ⁺)	56.7% (43.2%–69.4%)	95.0% (91.5%–97.3%)	0.685	0.081	1,000	1,000	134	1.5
LBC (HSIL ⁺)	28.3% (17.5%–41.4%)	100.0% (98.6%–100.0%)	1.000	0.122	1,000	1,000	46	1.0
Cotesting with two triage tests								
HPV16/18 & Dual-stained cytology	95.0% (86.1%–99.0%)	56.2% (49.9%–62.3%)	0.295	0.017	2,000	1,000	521	3.4
LBC (ASC-US ⁺) & Dual-stained cytology	90.0% (79.5%–96.2%)	72.5% (66.6%–77.8%)	0.387	0.026	2,000	2,000	376	2.6
HPV16/18 & LBC (ASC-US ⁺)	93.3% (83.8%–98.2%)	62.8% (56.6%–68.7%)	0.327	0.020	2,000	1,000	463	3.1
Cotesting with three triage tests								
HPV16/18 & LBC (ASC-US ⁺) & Dual-stained cytology	98.3% (91.1%–100.0%)	52.7% (46.4%–58.9%)	0.287	0.006	3,000	2,000	556	3.5
Multistep triage								
HPV16/18 → LBC (ASC-US ⁺) if HPV16/18-	93.3% (83.8%–98.2%)	62.8% (56.6%–68.7%)	0.327	0.020	1,653	653	463	3.1
HPV16/18 → Dual-stained cytology if HPV16/18-	95.0% (86.1%–99.0%)	56.2% (49.9%–62.3%)	0.295	0.017	1,653	653	521	3.4
HPV16/18 → LBC (ASC-US ⁺) if HPV16/18- → Dual-stained cytology if LBC-	98.3% (91.1%–100.0%)	52.7% (46.4%–58.9%)	0.287	0.006	2,175	1,175	556	3.5

NOTE: This includes the absolute sensitivity and the specificity for CIN2⁺ (and 95% CI), the computed PPV and cNPV (complement of the negative predictive value 1-NPV), the number of tests used in the triage strategy, including 1,000 hrHPV⁺ women, the number of colposcopy referrals and the NNR (number needed to refer to colposcopy to find 1 case of CIN2⁺). Three single triage tests include p16/Ki-67 dual-stained cytology, HPV16/18 genotyping and LBC triage considering three possible cutoff values for LBC positivity (LSIL, low-grade; HSIL, high-grade dyskaryosis). Several other triage test combinations include cotesting with 2 or 3 tests or multistep triage when the subsequent test is invoked after a negative result by the previous test.

^aHPV16/18 genotyping considered as a separate triage test.

^bHPV16/18 genotyping not counted as a separate triage test since integrated in the screening test.

^cASC-US corresponds to borderline cytology.

Table 2. Relative accuracy for cervical intraepithelial neoplasia of grade 2 or worse of tests using LBC and HPV16/18 as comparator in triage of women with an hrHPV-positive clinician-taken sample

Index test	Comparator test	Relative sensitivity	P _{McN}	Relative specificity	P _{McN}
Dual-stained HPV16/18	LBC (ASC-US ⁺)	1.24 (1.05-1.48)	0.0213	0.86 (0.80-0.92)	<0.0001
Dual-stained HPV16/18	LBC (ASC-US ⁺)	0.90 (0.67-1.21)	0.4927	0.79 (0.72-0.87)	<0.0001
Dual-stained HPV16/18	LBC (LSIL ⁺)	1.50 (1.20-1.88)	0.0005	0.81 (0.75-0.87)	<0.0001
Dual-stained HPV16/18	LBC (LSIL ⁺)	1.09 (0.79-1.49)	0.6015	0.74 (0.68-0.81)	<0.0001
Dual-stained HPV16/18	LBC (HSIL ⁺)	3.00 (2.03-4.42)	<0.0001	0.77 (0.72-0.82)	<0.0001
Dual-stained HPV16/18	LBC (HSIL ⁺)	2.18 (1.36-3.48)	0.0009	0.71 (0.65-0.76)	<0.0001
Dual-stained HPV16/18	HPV16/18	1.38 (1.10-1.73)	0.0060	1.09 (0.99-1.20)	0.0917

(Tables 1 and 3) and self-samples (Supplementary Tables S1 and S3).

In clinician-taken samples, application of two triage tests as co-tests [HPV 16/18 typing and p16/Ki-67 dual-stained cytology, LBC plus p16/Ki-67 dual-stained cytology and LBC (\geq borderline) plus HPV 16/18 typing] were all associated with absolute sensitivities of over $\geq 90\%$. Respective specificities of the above 2-test approaches ranged from 56% to 72%. For full details and CIs refer to Table 1.

The relative sensitivity and specificity of the 2-test approaches are detailed in Table 3 using LBC at a threshold of borderline as a comparator. All 2-test approaches that included p16/Ki-67 dual-stained cytology were significantly more sensitive than the comparator. HPV 16/18 typing with LBC at borderline and low-grade was significantly more sensitive than the comparator, although the gain in sensitivity was lost when the combination of HPV 16/18 plus LBC at cutoff value of \geq moderate was considered. Regarding specificity, all of the dual co-test triage approaches described above was significantly less specific than the comparator test.

Multistep triage options

The lower part of Table 1 shows the accuracy estimates of three multistep (sequential) triage strategies that all start with

HPV 16/18 typing. The strategy with the highest sensitivity of 98% (95% CI, 91-100) and the lowest post-test risk of CIN2⁺ (0.6%) involved triage of HPV16/18-negative women with LBC and if negative then a further triage with p16/Ki-67 dual-stained cytology. However, this approach was associated with the highest number of colposcopies and a specificity that was lower than any of the stand-alone and combined options.

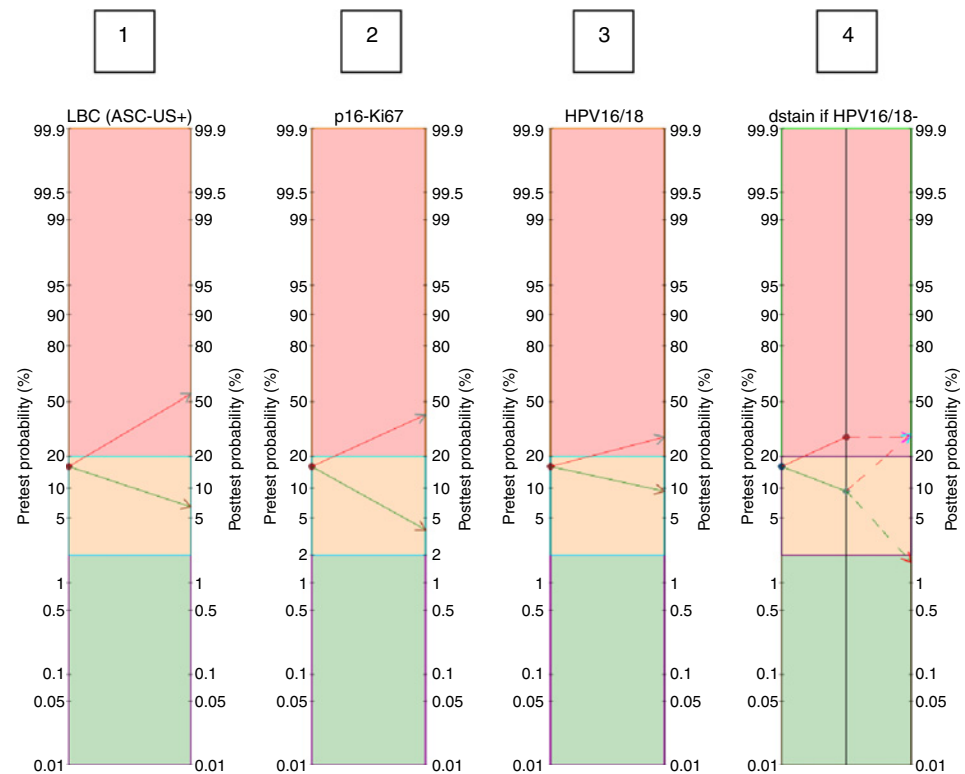
Figure 1 illustrates the probability of CIN2⁺ according to the three single-triage scenarios (i.e., LBC at borderline or worse, p16/Ki-67 dual-stained cytology, 16/18 typing; columns 1-3 respectively). The option of p16/Ki-67 dual-stained cytology subsequent to HPV 16/18 typing is incorporated into column 4. In the absence of triage, the pre-test probability of CIN2⁺ associated with hrHPV-positive status resides in the yellow zone (2%-20% indicating that surveillance is warranted). All three individual triage options, if positive, exceed the threshold of 20% risk at which colposcopy would be indicated. However, the post-test risk in women who are negative by any of the three individual triage options, does not fall into the green zone where routine screening can be advised. However, if the p16/Ki-67 dual-stained cytology is performed in women who test negative for HPV 16/18, the post-test(s) probability does fall into the green zone.

Table 3. Relative sensitivity and specificity for cervical intraepithelial neoplasia of grade 2 or worse of combinations of triage tests compared with liquid-base cytology (LBC)-based triage (at cutoff ASC-US⁺/borderline+) used to triage women with an hrHPV-positive result in clinician-taken samples

Index test	Relative sensitivity	P _{McN}	Relative specificity	P _{McN}
Cotesting with 2 tests				
HPV16/18 & dual-stained cytology	1.39 (1.16-1.67)	0.0004	0.63 (0.57-0.70)	<0.0001
LBC (ASC-US ⁺) & dual-stained cytology	1.32 (1.13-1.53)	0.0002	0.81 (0.76-0.87)	<0.0001
LBC (LSIL ⁺) & dual-stained cytology	1.32 (1.13-1.53)	0.0002	0.83 (0.78-0.89)	<0.0001
LBC (HSIL ⁺) & dual-stained cytology	1.24 (1.05-1.48)	0.0213	0.86 (0.80-0.92)	<0.0001
HPV16/18 & LBC (ASC-US ⁺)	1.37 (1.17-1.60)	0.0001	0.70 (0.65-0.77)	<0.0001
HPV16/18 & LBC (LSIL ⁺)	1.27 (1.05-1.53)	0.0192	0.75 (0.69-0.82)	<0.0001
HPV16/18 & LBC (HSIL ⁺)	1.10 (0.87-1.39)	0.5572	0.79 (0.72-0.87)	<0.0001
Cotesting with 3 tests				
HPV16/18 & LBC (ASC-US ⁺) & dual-stained cytology	1.44 (1.21-1.70)	<0.0001	0.59 (0.53-0.66)	<0.0001
HPV16/18 & LBC (LSIL ⁺) & dual-stained cytology	1.44 (1.22-1.70)	<0.0001	0.61 (0.55-0.68)	<0.0001
HPV16/18 & LBC (HSIL ⁺) & dual-stained cytology	1.39 (1.16-1.67)	0.0004	0.63 (0.57-0.70)	<0.0001
Two-step triage				
HPV16/18 \rightarrow LBC (ASC-US ⁺) if HPV16/18-	1.37 (1.17-1.60)	0.0001	0.70 (0.65-0.77)	<0.0001
HPV16/18 \rightarrow LBC (LSIL ⁺) if HPV16/18-	1.27 (1.05-1.53)	0.0192	0.75 (0.69-0.82)	<0.0001
HPV16/18 \rightarrow LBC (HSIL ⁺) if HPV16/18-	1.10 (0.87-1.39)	0.4328	0.79 (0.72-0.87)	<0.0001
HPV16/18 \rightarrow dual-stained cytology if HPV16/18-	1.39 (1.16-1.67)	0.0004	0.63 (0.57-0.70)	<0.0001
LBC (ASC-US ⁺) \rightarrow HPV16/18 if <ASC-US	1.37 (1.17-1.60)	0.0001	0.70 (0.65-0.77)	<0.0001
LBC (LSIL ⁺) \rightarrow HPV16/18 if <LSIL	1.27 (1.05-1.53)	0.0192	0.75 (0.69-0.82)	<0.0001
LBC (HSIL ⁺) \rightarrow HPV16/18 if <HSIL	1.10 (0.87-1.39)	0.5572	0.79 (0.72-0.87)	<0.0001
LBC (ASC-US ⁺) \rightarrow dual-stained cytology if <ASC-US	1.32 (1.13-1.53)	0.0002	0.81 (0.76-0.87)	<0.0001
LBC (LSIL ⁺) \rightarrow dual-stained cytology if <LSIL	1.32 (1.13-1.53)	0.0002	0.83 (0.78-0.89)	<0.0001
LBC (HSIL ⁺) \rightarrow dual-stained cytology <HSIL	1.24 (1.05-1.48)	0.0213	0.86 (0.80-0.92)	<0.0001
Dual-stained cytology \rightarrow HPV16/18 if dual-stained	1.39 (1.16-1.67)	0.0004	0.63 (0.57-0.70)	<0.0001

Figure 1.

Pre-test-post-test probability plot associated with four triage options. (Adopted from Arbyn 2017, *Ann Intern Med*). Pre-test (left y-axis) and post-test probabilities (right y-axis) of CIN2⁺ after triage of hrHPV⁺ women (in cervical samples) using (i) cytologic triage using ASC-US (borderline) or worse as cutoff value; (ii) p16/Ki-67 dual-stained cytology; (iii) HPV16/18 genotyping; or (iv) a double-step triage with HPV16/18 genotyping followed by p16/Ki-67 dual-stained cytology of HPV16/18 negative. Decision "benchmarks" are defined at CIN2⁺ risk levels of 2% and 20%. When the risk is in the red zone, referral to colposcopy is indicated; in the yellow zone, the patient needs further surveillance; whereas in the green zone, the woman may be released to the normal screening schedule.



Discussion

Now that hrHPV testing with a validated assay is accepted as the optimal modality for primary screening (26); one of the key challenges is to define and implement robust triage strategies, to avoid unnecessary referral to colposcopy (27). The lack of consensus around optimal triage tests is evidenced through the variety of approaches considered across the various HR-HPV-based screening programs (7). Arguably, the most evidenced triage tests involve cytology, repeat HPV testing, HPV 16/18 and p16 staining (with or without Ki-67) and various combinations thereof (5). However, there are comparatively few studies in which these have been compared with each other within the same population and fewer still where they have been evaluated in both clinician and self-taken samples. The present analysis attempted to address this.

P16/Ki-67 dual-stained cytology was more sensitive than traditional LBC and HPV16/18 genotyping, although the specificity compared to LBC was significantly lower. This is at variance with findings of Wentzensen and colleagues (13), who showed similar sensitivity and a higher specificity of p16/Ki-67 triage compared with traditional cytology. However, it is well established that the performance of cytology varies between settings (8) and the cytology associated with the present study was performed as part of an organized call-recall UK-wide program, which is subject to nationally observed quality assurance and guidelines, whereas the United States provides opportunistic screening. Importantly, unlike the U.S. study, cytologists were unaware of HPV status when reporting results. Previous research has shown that knowledge of HPV positivity can lead to a higher sensitivity at the

compromise of specificity (24). Ideally, a triage test, if negative, should provide sufficient reassurance to permit a safe return to routine recall while conferring sufficient risk, if positive, to warrant referral to colposcopy. Figures of <2% probability of CIN2⁺ for a negative result + and ≥20% probability of CIN2⁺ of a positive triage result for CIN2 have been proposed (23, 28). It is notable that no single-triage approach satisfied these criteria.

HPV 16/18 typing alone had the lowest sensitivity and specificity as a single triage; however, one practical advantage of typing is that it occurs concurrently with the screening test. In settings where cytology is impractical and/or of low quality, this may be a more workable solution—particularly as the post-test probability of a positive result exceeds 20% for CIN2⁺. Managing the 16/18 negatives is more challenging; a subsequent negative dual stain led to a post-test probability of <2%; however, the number of additional colposcopies associated with this scenario was relatively high.

Screening programs that enfranchise hard to reach populations with a self-sampling option are increasing as is the evidence that this may be a practical and technically sound option for women who do attend. Although the focus of this work was to evaluate triage approaches in clinician-taken samples, it was notable that triage test performance was similar in self-taken samples. The only significant difference was observed in relation to HPV 16/18 typing that showed a lower relative specificity compared with both LBC and the dual stain. The higher prevalence and diversity of hrHPV in the vagina compared with the cervix is relatively well documented and may account for this observation (29, 30). Previous studies have also demonstrated that self-taken samples

are valid biospecimens for host methylation biomarker detection (19, 31, 32).

The main limitation of this study is that all analyses were performed retrospectively and women were not triaged in real-time. This may have accounted for the number of technically invalid samples by the dual-staining, which was relatively high (between 7% and 8%). Comparatively, in the original PaVdAG study only around 2% of samples were unsatisfactory (19). In addition, if the study was prospective, women who were hrHPV positive on their self-taken sample would have been invited for a subsequent cytology sample, whereas we used the contemporaneous LBC sample (collected at the same time as the self-sample) as a proxy.

Another limitation is that not all controls/negative outcomes were associated with the gold standard (histology). However, using negative cytology across screening rounds has precedent (33). We are also aware that the overarching PaVdAG design may have somewhat biased the absolute specificity estimates. This was addressed by including relative specificity estimates that are less susceptible to bias (34).

The focus of this study was to evaluate the performance of the three most-evidenced triage strategies and this informed our selection of cytology, dual-staining and 16/18 typing. However, the evidence for other triage markers, including host and viral methylation targets and sequential viral load measurement, is increasing (3, 35, 36). Arguably the optimal triage test would be delivered at a single time point, as a single test. If the said test was amenable to self-collected samples, avoided subjective interpretation and was also impervious to the impact of immunization this would be of significant value. A recent study that measured the prevalence of HPV 16/18, using clinically validated assays, in immunized young women attending for first screen showed that 16/18 prevalence was 75% lower than that of unimmunized cohorts. The changing pattern of 16/18 infection will clearly have implications for the predictive value of 16/18 typing in immunized women (37). A further report also showed significant reductions in the predictive value of cytology in immunized women, particularly of low-grade cytology for CIN2⁺ (9). Objective triage tests that are positive irrespective of HPV type would have greater durability for immunized populations.

We report that LBC and dual-staining out-perform HR-HPV16/18 genotyping as single time point triage tests in terms of sensitivity and specificity. Although dual-staining showed significantly higher sensitivity than LBC its specificity was significantly lower.

The lowest risk of CIN2⁺ in screen negative women (0.6%) was achieved by offering colposcopy to all HPV16/18-positive

women, and those who were hrHPV positive for non-16/18 types who had positive LBC and/or dual-staining. Such low cNPV is likely to remain below 2% even with the anticipated reduced performance of screening tests in immunized population. Although we did not perform a formal cost/benefit analysis on the various approaches, clearly the relative cost per test (including patient and clinician time) will be influential for future decision-making. Furthermore, given the rate of development of self-sampling and rapid molecular diagnostics, a purely molecular pathway for cervical screening may be a not too distant proposition.

Disclosure of Potential Conflicts of Interest

G.A. Stanczuk has received speakers honoraria from Roche. K. Cuschieri reports receiving a commercial research grant from GeneFirst, EuroImmun, Genomica, LifeRiver, SelfScreen, and Hologic. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: G.A. Stanczuk, H. Currie, A. Wilson

Development of methodology: G.A. Stanczuk, A. Wilson

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): G.A. Stanczuk, G.J. Baxter, H. Currie, W. Forson, A. Wilson

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): G.A. Stanczuk, A. Wilson, L. Patterson, M. Arbyn

Writing, review, and/or revision of the manuscript: G.A. Stanczuk, H. Currie, K. Cuschieri, A. Wilson, T. Palmer, M. Arbyn

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): G.A. Stanczuk, G.J. Baxter, J.R. Lawrence, K. Cuschieri, L. Patterson, L. Govan, J. Black

Study supervision: G.A. Stanczuk, G.J. Baxter

Grant Support

The study was sponsored by NHS Dumfries and Galloway and supported by funding from the CSO, Scotland. Roche donated Cobas 4800 collection and testing kits and CINtec PLUS kits. M. Arbyn was supported by the Seventh Framework Program of the Directorate-General for Research of the European Commission through the comparing health services interventions for the prevention of HPV-related cancer (coheaHr) network (grant 603019) and by the Joint Action on Comprehensive Cancer Control, which has received funding from the European Union within the framework of the Health Program (2008–2013).

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 June 14, 2017; revised July 20, 2017; accepted September 1, 2017; published OnlineFirst September 8, 2017.

References

- Arbyn M, Ronco G, Anttila A, Meijer CJLM, Poljak M, Ogilvie G, et al. Evidence regarding HPV testing in secondary prevention of cervical cancer. *Vaccine* 2012;30(Suppl 5):F88–F99.
- von Karsa L, Arbyn M, De Vuyst H, Dillner J, Dillner L, Franceschi S, et al. European guidelines for quality assurance in cervical cancer screening. Summary of the Supplements on HPV Screening and Vaccination (Review) *Papillomavirus Research* 2015;1:22–31.
- Meijer CJ, Berkhof J, Castle PE, Hesselink AT, Franco EL, Ronco G, et al. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. *Int J Cancer* 2009;124:516–20.
- Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, et al. ACS-ASCCP-ASCP Cervical Cancer Guideline Committee. 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.
- Luttmer R, De Strooper LM, Steenberg RD, Berkhof J, Snijders PJ, Heide-man DA, et al. Management of high-risk HPV-positive women for detection of (pre)cancer. *Expert Rev Mol Diagn* 2016;16:961–74.
- Wentzensen N, Schiffman M, Palmer T, Arbyn M. Triage of HPV positive women in cervical cancer screening. *J Clin Virol* 2016;76(Suppl 1):S49–55.
- Rijkaart DC, Berkhof J, van Kemenade FJ, Coupe VM, Hesselink AT, Rozendaal L, et al. Evaluation of 14 triage strategies for HPV DNA-positive women in population-based cervical screening. *Int J Cancer* 2012;130:602–10.
- Wright TC Jr, Stoler MH, Behrens CM, Sharma A, Sharma K, Apple R. Interlaboratory variation in the performance of liquid-based cytology: insights from the ATHENA trial. *Int J Cancer* 2014;134:1835–43.

9. Palmer TJ, McFadden M, Pollock KG, Kavanagh K, Cuschieri K, Cruickshank M, et al. HPV immunisation and cervical screening—confirmation of changed performance of cytology as a screening test in immunised women: a retrospective population-based cohort study. *Br J Cancer* 2016;114:582–9.
10. Uijterwaal MH, Polman NJ, Witte BI, van Kemenade FJ, Rijkaart D, Berkhof J, et al. Triaging HPV-positive women with normal cytology by p16/Ki-67 dual-stained cytology testing: baseline and longitudinal data. *Int J Cancer* 2015;136:2361–8.
11. Luttmer R, Berkhof J, Dijkstra MG, van Kemenade FJ, Snijders PJ, Heideman DA, et al. Comparing triage algorithms using HPV DNA genotyping, HPV E7 mRNA detection and cytology in high-risk HPV DNA-positive women. *J Clin Virol* 2015;67:59–66.
12. Carozzi F, Gillio-Tos A, Confortini M, Del Mistro A, Sani C, De Marco L, et al. 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* 2013;14:168–76.
13. Wentzensen N, Fetterman B, Castle PE, Schiffman M, Wood SN, Stiemerling E, et al. p16/Ki-67 dual stain cytology for detection of cervical precancer in HPV-positive women. *J Natl Cancer Inst* 2015;107:djv257.
14. Bergeron C, Ronco G, Reuschenbach M, Wentzensen N, Arbyn M, Stoler M, et al. The clinical impact of using p16(INK4a) immunocytochemistry in cervical histopathology and cytology: an update of recent developments. *Int J Cancer* 2015;136:2741–51.
15. Chatzistamatiou K, Moysiadi T, Moschaki V, Panteleris N, Agorastos T. Comparison of cytology, HPV DNA testing and HPV 16/18 genotyping alone or combined targeting to the more balanced methodology for cervical cancer screening. *Gynecol Oncol* 2016;142:120–7.
16. Schiffman M, Burk RD, Boyle S, Raine-Bennett T, Katki HA, Gage JC, et al. A study of genotyping for management of human papillomavirus-positive, cytology-negative cervical screening results. *J Clin Microbiol* 2015;53:52–9.
17. von Knebel Doeberitz M, Reuschenbach M, Schmidt D, Bergeron C. Biomarkers for cervical cancer screening: the role of p16(INK4a) to highlight transforming HPV infections. *Expert Rev Proteomics* 2012;9:149–63.
18. Roelens J, Reuschenbach M, von Knebel-Doeberitz M, Wentzensen N, Bergeron C, Arbyn M. p16INK4a immunocytochemistry versus HPV testing for triage of women with minor cytological abnormalities: a systematic review and meta-analysis. *Cancer* 2012;120:294–307.
19. Stanczuk G, Baxter G, Currie H, Lawrence J, Cuschieri K, Wilson A, et al. Clinical validation of hrHPV testing on vaginal and urine self-samples in primary cervical screening (cross-sectional results from the Papillomavirus Dumfries and Galloway-PaVDaG study). *BMJ Open* 2016;6:e010660.
20. Smith JHF, Patnick J. Achievable standards, benchmarks for reporting and criteria for conducting cervical cytopathology. Second edition. Sheffield, United Kingdom: NHS Cancers Screening Programme; 2013. Available from: <http://www.cancerscreening.nhs.uk/cervical/publications/nhscsp01.html>.
21. In: Luesley D, Leeson S (editors). 2010. Colposcopy and programme management, 2nd ed. NHS Cancers Screening Programme, Sheffield, United Kingdom. Available from: <http://www.cancerscreening.nhs.uk/cervical/publications/nhscsp20.html>.
22. Hirschowitz L. Histopathology reporting in cervical screening—an integrated approach. Second edition. Sheffield, United Kingdom: NHS Cancers Screening Programme; 2012. Available from: <http://www.cancerscreening.nhs.uk/cervical/publications/cc-04.html>.
23. Arbyn M, Xu L, Verdoodt F, Cuzick J, Szarewski A, Belinson J, et al. Genotyping for HPV16 and HPV18 in women with minor cervical lesions: a systematic review and meta-analysis. *Ann Intern Med* 2017;166:118–27.
24. Benoy IH, Vanden Broeck D, Ruymbeke MJ, Sahebali S, Arbyn M, Bogers JJ, et al. Prior knowledge of HPV status improves detection of CIN2+ by cytology screening. *Am J Obstet Gynecol* 2011;205:569.
25. Tang NS, Tang ML, Chan IS. On tests of equivalence via non-unity relative risk for matched-pair design. *Stat Med* 2003;22:1217–33.
26. Arbyn M, Snijders PJ, Meijer CJ, Berkhof H, Cuschieri K, Kocjan BJ, et al. Which high-risk HPV assays fulfil criteria for use in primary cervical cancer screening? *Clin Microbiol Infect* 2015;21:817–26.
27. Huh WK, Ault KA, Chelmow D, Davey DD, Goulart RA, Garcia FA, et al. Use of primary high-risk human papillomavirus testing for cervical cancer screening: interim clinical guidance. *Obstet Gynecol* 2015;125:330–7.
28. Arbyn M, Roelens J, Martin-Hirsch P, Leeson S, Wentzensen N. Use of HC2 to triage women with borderline and mild dyskaryosis in the UK. *Br J Cancer* 2011;105:877–80.
29. Verdoodt F, Jentschke M, Hillemanns P, Racey CS, Snijders PJ, Arbyn M. Reaching women who do not participate in the regular cervical cancer screening programme by offering self-sampling kits: a systematic review and meta-analysis of randomised trials. *Eur J Cancer* 2015;51:2375–85.
30. Zhang SK, Ci PW, Velicer C, Kang LN, Liu B, Cui JF, et al. Comparison of HPV genotypes and viral load between different sites of genital tract: the significance for cervical cancer screening. *Cancer Epidemiol* 2014;38:168–73.
31. De Strooper LM, Verhoef VM, Berkhof J, Hesselink AT, de Bruin HM, van Kemenade FJ, et al. Validation of the FAM19A4/mir124-2 DNA methylation test for both lavage- and brush-based self-samples to detect cervical (pre)cancer in HPV-positive women. *Gynecol Oncol* 2016;141:341–7.
32. Hesselink AT, Heideman DA, Steenbergen RD, Gök M, van Kemenade FJ, Wilting SM, et al. Methylation marker analysis of self-sampled cervico-vaginal lavage specimens to triage high-risk HPV-positive women for colposcopy. *Int J Cancer* 2014;135:880–6.
33. Arbyn M, Depuydt C, Benoy I, Bogers J, Cuschieri K, Schmitt M, et al. VALGENT: A protocol for clinical validation of human papillomavirus assays. *J Clin Virol* 2016;76(Suppl 1):S14–21.
34. Arbyn M, Verdoodt F, Snijders PJ, Verhoef VM, Suonio E, Dillner L, et al. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. *Lancet Oncol* 2014;15:172–83.
35. Lorincz AT, Brentnall AR, Scibior-Bentkowska D, Reuter C, Banwait R, Cadman L, et al. Validation of a DNA methylation HPV triage classifier in a screening sample. *Int J Cancer* 2016;138:2745–51.
36. Verhelst S, Poppe WA, Bogers JJ, Depuydt CE. Serial measurement of type-specific human papillomavirus load enables classification of cervical intraepithelial neoplasia lesions according to occurring human papillomavirus-induced pathway. *Eur J Cancer Prev* 2017;26:156–64.
37. Bhatia R, Kavanagh K, Cubie HA, Serrano I, Wennington H, Hopkins M, et al. Use of HPV testing for cervical screening in vaccinated women—Insights from the SHEVa (Scottish HPV Prevalence in Vaccinated Women) study. *Int J Cancer* 2016;138:2922–31.