

Defining optimal triage strategies for hrHPV screen positive women– an evaluation of HPV 16/18 genotyping, cytology and p16/Ki-67 cyto-immunochemistry.

Authors: 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, Marc Arbyn⁶

Affiliations:

¹Dept. of Research and Development, Dumfries and Galloway Royal Infirmary, Bankend Rd, Dumfries, DG1 4 AP, United Kingdom.

²Dept. of Obstetrics and Gynaecology, Dumfries and Galloway Royal Infirmary, Dumfries, DG1 4 AP, United Kingdom,

³Scottish HPV Reference Laboratory, Royal Infirmary of Edinburgh, Edinburgh, EH16 4SA, United Kingdom.

⁴ Dept. of Pathology, Monklands Hospital, Monkscourt Avenuel, Airdrie, ML6 0JS, United Kingdom.

⁵Dept. of Pathology, NHS Highland, Old Perth Rd, Inverness, IV2 3UJ, United Kingdom.

⁶ Unit of Cancer Epidemiology/Belgian Cancer Centre, Scientific Institute of Public Health, J Wytsmanstreet 14, B1050 , Brussels, Belgium.

Running Title: Optimal triage of hrHPV positive women within PaVDaG study.

Keywords: HPV, cervical cancer screening, triage of hrHPV+, p16/Ki-67 cyto-immunochemistry, dual-stained

Corresponding author current contact: Dr Grazyna A. Stanczuk, Department of Obstetrics and Gynaecology, Western Isles Hospital, Mac Aulay Rd, Stornoway, SH1 2AF, tel: +44 1851 704704, fax: +44 1851 708235, e-mail: grazyna.stanczuk@nhs.net.

Funding: 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 Programme (2008-2013).

Conflicts of interest: GAS and GJB has received speaker's fees from Roche. KC has received project funding and/or consumables to carry out assay evaluations from HOLOGIC, Qiagen, Roche, NorChip, Cepheid, Becton Dickinson GSK & Abbott in the past. MA has received support from the European Commission, the Belgian Cancer Centre and the VALGENT projects (network for validation of HPV genotyping test, coordinated by the Scientific Institute of Public Health, Brussels, Belgium) and support from the seventh framework program of DG Research of the European Commission, through the COHEAHR network (grant No 603019), and from CanCon (funded by DG Sanco, Luxembourg, Grand-Duchy of Luxembourg).

This study collection kits and HPV testing was donated by Roche.

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 cyto-immunochemistry 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 cyto-immunochemistry 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 cyto-immunochemistry 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 cyto-immunochemistry 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 which incorporated HPV 16/18 typing prior to p16/ki-67 cyto-immunochemistry in HPV 16/18 negative women yielded a post test probability of CIN2+ of >20% while women who tested negative had a probability of CIN2+ of <2%.

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 to 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 roll out 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 organised settings the reliance on morphology and subjective skills may limit the overall effectiveness of this approach – particularly in immunised 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 binary scenario where negative women could be returned to routine recall and positive 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, Alameda, USA) (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 to other HR-HPV types and the observation that makers of abnormal cell proliferation, that are detectable through cyto-immunochemistry are indicative of potentially transforming infections (16-17). There is evidence to indicate that the specificity and positive predictive value 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 US demonstrated higher specificity of p16/Ki-67 dual-stained cytology compared to 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 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 over 5000 women recruited to the PaVdaG study (19). The present 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 et al, 2016. Briefly women aged 20-60, 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 which detects 14 hrHPV types and provides a read out which 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 histological 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 which incorporated 61 cases with histological 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 (three 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 above triage tests was also assessed for 57 CIN2+ cases and 335 controls HR-HPV+ 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 cyto-immunochemistry

New slides were made for the p16/Ki-67 dual-stained cytology analysis using the CINtec PLUS cytology kit (Roche MTM Laboratories, Heidelberg, Germany) 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 morphological 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. Positive predictive value (PPV) and complement 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%) (19). The number of tests used in a particular triage option, the number of colposcopies per 1000 hrHPV+ patients, as well as the number of referrals needed to find one CIN2+ case (NNR [number needed to refer= 1/PPV]) was 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 to 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 coloured 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 cytological 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 to traditional cytology at cut-off of borderline squamous changes (13). A matched design was used for sample size computation (25). We hypothesized non-inferior 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 *et al*, 2009. The number of non-diseased 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 a hrHPV positive self-sample were similar (see Supplementary Table 1).

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 to 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 vs. 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 Suppl. Table 2).

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 and Table 3) and self samples (Suppl. Table 1 and Suppl. Table 3)

In clinician taken samples, application of two triage tests as co-tests [HPV 16/18 typing & 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 56-72%. 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 which included

p16/Ki-67 dual-stained cytology were significantly more sensitive than the comparator. HPV 16/18 typing with LBC at borderline (ASC-US) 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 \geq moderate was considered. Regarding specificity, all of the dual co-test triage approaches described above was significantly less specific than the comparator test.

Multi-step triage options

The lower part of Table 1 shows the accuracy estimates of three multi-step (sequential) triage strategies which 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/ASC-US+, 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.

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 heterogeneity of approaches considered across the various HR-HPV based implementation projects and screening programmes (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 to 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 Wentzsen *et al*, 2015 who showed similar sensitivity and a higher specificity of p16/Ki-67 triage compared to traditional cytology (13). However, it is well established that the performance of cytology varies between settings (8). and the cytology associated with present study was performed as part of an organised call-recall UK wide programme which is subject to nationally observed quality assurance and guidelines, whereas the US provides opportunistic screening. Importantly, unlike the US 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 $\geq 20\%$ 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. Supplementary figure 1 displays the sensitivity and specificity estimates of 27 triage strategies in ROC space.

Screening programmes 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 which showed a lower relative specificity compared to both LBC and the dual stain. The higher prevalence and diversity of hrHPV in the vagina compared to 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 bio specimens for host methylation biomarker detection (31-32, 19).

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 7-8 %). 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 which 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 immunisation this would be of significant value. A recent study which measured the prevalence of HPV 16/18, using clinically validated assays, in immunised young women attending for first screen showed that 16/18 prevalence was 75% lower than that of unimmunised cohorts. The changing pattern of 16/18 infection will clearly have implications for the predictive value of 16/18 typing in immunised women (37). A further report also showed significant reductions in the predictive value of cytology in immunised women, particularly of low-grade cytology for CIN2+ (9). Objective triage tests that are positive irrespective of HPV type would have greater durability for immunised 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. While 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 immunised population. While 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.

References

1. 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.
2. von Karsa L, Arbyn M, De Vuyst H, Joakim Dillner, Lena Dillner, Silvia Franceschi, 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.
3. 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: 516-520.
4. 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.
5. Luttmer R, De Strooper LM, Steenbergen RD, Berkhof J, Snijders PJ, Heideman DA, Meijer CJ. Management of high-risk HPV-positive women for detection of cervical (pre)cancer. *Expert Rev Mol Diagn*. 2016; 16: 961-74.
6. 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.
7. 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.
8. 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. Triage of 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, Meijer CJ. 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; NTCC working group. 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(12): djv257.

14. Bergeron C, Ronco G, Reuschenbach M, Wentzensen N, Arbyn M, Stoler M, von Knebel Doeberitz M. The clinical impact of using p16(INK4a) immunochemistry in cervical histopathology and cytology: an update of recent developments. *Int J Cancer.* 2015; 136: 2741-51.

15. Chatzistamatiou K, Moysiadis 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, Arbyn M. 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(4): e010660.

20. Smith JHF, Patnick J (ed). 2013. Achievable standards, benchmarks for reporting and criteria for conducting cervical cytopathology, 2nd ed. NHS Cancers Screening Programme, Sheffield, United Kingdom.
<http://www.cancerscreening.nhs.uk/cervical/publications/nhscsp01.html>.

21. Luesley D, Leeson S (ed). 2010. Colposcopy and programme management, 2nd ed. NHS Cancers Screening Programme, Sheffield, United Kingdom
<http://www.cancerscreening.nhs.uk/cervical/publications/nhscsp20.html>.

22. Hirschowitz L (ed). 2012. Histopathology reporting in cervical screening—an integrated approach, 2nd ed. NHS Cancers Screening Programme, Sheffield, United Kingdom.
<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: 118-127.

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.e1-7.
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, Poljak M. 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 and Wentzensen N. Use of HC2 to triage women with borderline and mild dyskaryosis in the UK. *Br J Cancer*. 2011; 105: 877- 880.
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 Aug 15; 135(4): 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.* 2016; Feb 17. [Epub ahead of print]

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.

Table 1. Accuracy and diagnostic efficiency of triage tests to detect CIN2+ among women with a hrHPV positive clinician-taken cervical sample. This includes the absolute sensitivity and the specificity for CIN2+ (and 95% CI), the computed PPV (positive predictive value) and cNPV (complement of the negative predictive value 1-NPV), the number of tests used in the triage strategy including 1000 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 liquid-base cytology (LBC) triage considering three possible cut offs for LBC positivity (LSIL-low grade and HSIL- high grade dyskaryosis). Several other triage test combinations include co-testing with 2 or 3 tests or multi step triage when the subsequent test is invoked after a negative result by the previous test.

* HPV16/18 genotyping considered as a separate triage test; ** HPV16/18 genotyping not counted as a separate triage test since integrated in the screening test; *** ASC-US corresponds to borderline cytology

Triage tests	Sensitivity (95% CI)	Specificity (95% CI)	PPV	cNPV	Triage tests		Colpo- scopies /1000	NNR
					(1)* /1000	(2)**		
One single triage test								
Dual-stained cytology	85.0% (73.4-92.9%)	76.7% (71.1-81.8%)	41.4%	3.6%	1000	1000	333	2.4
HPV16/18	61.7% (48.2-73.9%)	70.5% (64.6-76.0%)	0.288	0.095	1000	0	347	3.5
LBC (ASC-US+)**	68.3% (55.0-79.7%)	89.1% (84.7-92.7%)	0.549	0.064	1000	1000	202	1.8
LBC (LSIL+)	56.7% (43.2-69.4%)	95.0% (91.5-97.3%)	0.685	0.081	1000	1000	134	1.5
LBC (HSIL+)	28.3% (17.5-41.4%)	100.0% (98.6-100.0%)	1.000	0.122	1000	1000	46	1.0
Co testing with 2 triage tests								
HPV16/18 & Dual-stained cytology	95.0% (86.1-99.0%)	56.2% (49.9-62.3%)	0.295	0.017	2000	1000	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	2000	2000	376	2.6
HPV16/18 & LBC (ASC-US+)	93.3% (83.8-98.2%)	62.8% (56.6-68.7%)	0.327	0.020	2000	1000	463	3.1
Co testing with 3 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	3000	2000	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	1653	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	1653	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	2175	1175	556	3.5

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 a hrHPV positive clinician-taken sample.

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

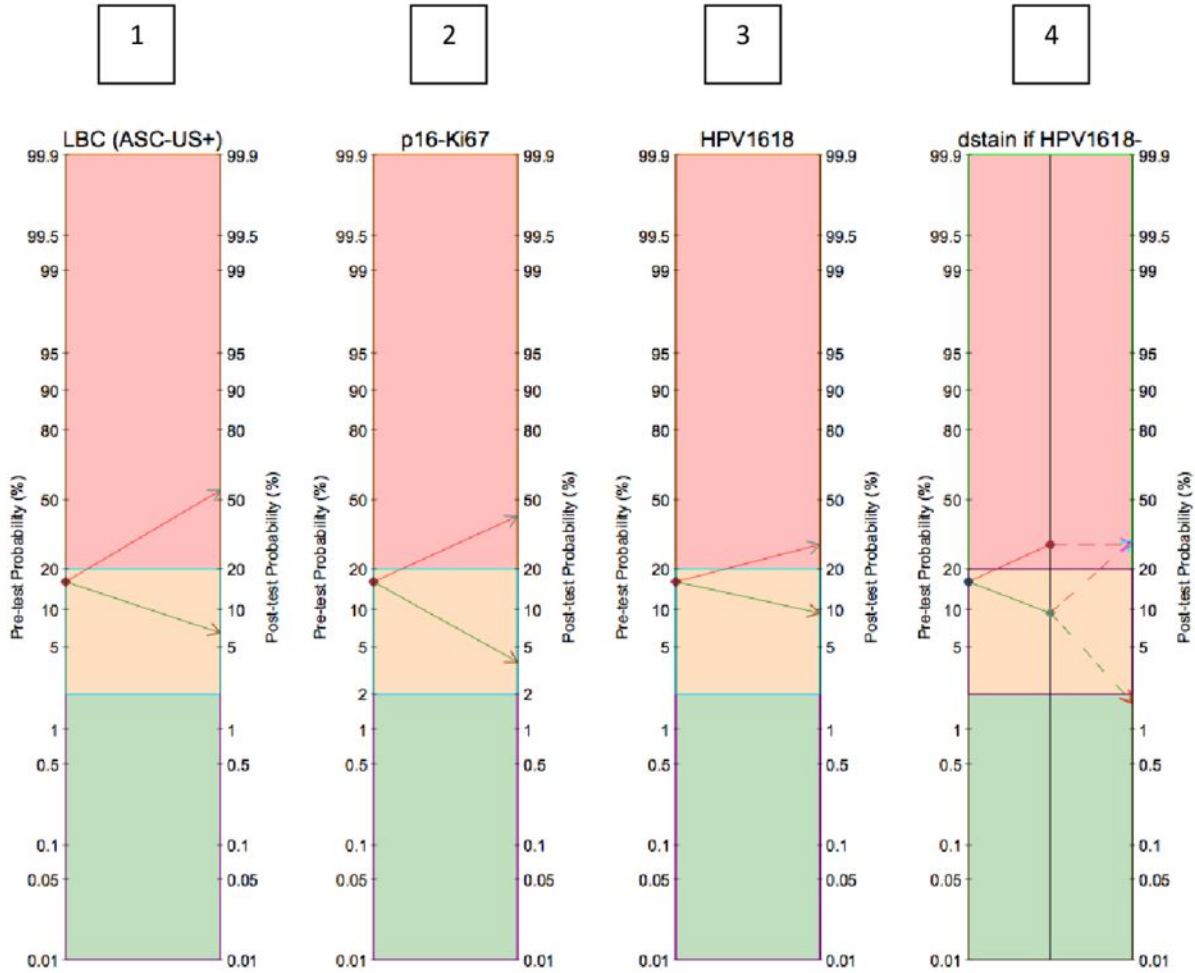
Table 3. Relative sensitivity and specificity for cervical intraepithelial neoplasia of grade 2 or worse of combinations of triage tests compared to liquid-base cytology (LBC)-based triage (at cut-off ASC-US+/borderline+) used to triage women with a hrHPV positive result in clinician taken samples.

Index test	Relative sensitivity	p _{McN}	Relative specificity	p _{McN}
<i>Co-testing with 2 tests</i>				
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
<i>Co-testing with 3 tests</i>				
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
<i>2 step triage</i>				
HPV16/18 → LBC (ASC-US+) if HPV16/18-	1.37 (1.17-1.60)	0.0001	0.70 (0.65-0.77)	<0.0001
HPV16/18 → LBC (LSIL+) if HPV16/18-	1.27 (1.05-1.53)	0.0192	0.75 (0.69-0.82)	<0.0001
HPV16/18 → LBC (HSIL+) if HPV16/18-	1.10 (0.87-1.39)	0.4328	0.79 (0.72-0.87)	<0.0001
HPV16/18 → 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+) → HPV16/18 if <ASC-US	1.37 (1.17-1.60)	0.0001	0.70 (0.65-0.77)	<0.0001
LBC (LSIL+) → HPV16/18 if <LSIL	1.27 (1.05-1.53)	0.0192	0.75 (0.69-0.82)	<0.0001
LBC (HSIL+) → HPV16/18 if <HSIL	1.10 (0.87-1.39)	0.5572	0.79 (0.72-0.87)	<0.0001
LBC (ASC-US+) → dual-stained cytology if <ASC-US	1.32 (1.13-1.53)	0.0002	0.81 (0.76-0.87)	<0.0001
LBC (LSIL+) → dual-stained cytology if <LSIL	1.32 (1.13-1.53)	0.0002	0.83 (0.78-0.89)	<0.0001
LBC (HSIL+) → dual-stained cytology <HSIL	1.24 (1.05-1.48)	0.0213	0.86 (0.80-0.92)	<0.0001
dual-stained cytology → 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 HR-HPV positive women (in cervical samples) using (1) cytological triage using ASC-US (borderline) or worse as cut off; (2) p16/Ki-67 dual-stained cytology; (3) HPV16/18 genotyping; or (4) a double step triage with HPV16/18 genotyping followed by p16/Ki-67 dual-stained cytology of HPV16/18 negative women. 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.

Figure 1.



Cancer Epidemiology, Biomarkers & Prevention

Defining optimal triage strategies for hrHPV screen positive women - an evaluation of HPV 16/18 genotyping, cytology and p16/Ki-67 cyto-immunochemistry.

Grazyna A. Stanczuk, Gwen J. Baxter, Heather Currie, et al.

Cancer Epidemiol Biomarkers Prev Published OnlineFirst September 8, 2017.

Updated version	Access the most recent version of this article at: doi: 10.1158/1055-9965.EPI-17-0534
Supplementary Material	Access the most recent supplemental material at: http://cebp.aacrjournals.org/content/suppl/2017/09/08/1055-9965.EPI-17-0534.DC1
Author Manuscript	Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org .
Permissions	To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org .