Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical cancer screening of women aged 25 years and older: a subanalysis of the ATHENA study

Philip E Castle, Mark H Stoler, Thomas C Wright Jr, Abha Sharma, Teresa L Wright, Catherine M Behrens

Lancet Oncol 2011; 12: 880–90

Published Online
August 23, 2011
DOI:10.1016/S1470-2045(11)70188-7

See Comment page 831

American Society for Clinical Pathology Institute, Washington, DC, USA (P E Castle PhD); **University of Virginia Health System, Charlottesville, VA, USA** (Prof M H Stoler MD); **Department of Pathology, Columbia University School of Medicine, New York, NY, USA** (Prof T C Wright Jr MD); **and Roche Molecular Systems, Pleasanton, CA, USA** (A Sharma PhD, T L Wright MD, C M Behrens MD)

Correspondence to:
Dr Philip E Castle, American Society for Clinical Pathology Institute, 1225 New York Avenue NW, Suite 250, Washington, DC 20005, USA
philip.castle@ascp.org

Summary

Background The ATHENA study was designed to assess the performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping compared with liquid-based cytology for cervical cancer screening in a large US population aged 21 years and older. We did a subanalysis of this population to compare the screening performance of the cobas HPV test versus liquid-based cytology in women aged 25 years and older, and assess management strategies for HPV-positive women.

Methods Women aged 25 years or older who were attending routine cervical screening were enrolled from 61 clinical centres in 23 US states. Cervical specimens were obtained for liquid-based cytology and HPV DNA testing with two first-generation assays (Amplicor HPV test and Linear Array HPV genotyping test) and the second-generation cobas HPV test (with individual HPV16 and HPV18 detection). Colposcopy and diagnostic biopsies were done on women with atypical squamous cells of undetermined significance (ASC-US) or worse cytology, those who tested positive with either first-generation HPV test, and a random sample of women who tested negative for HPV and cytology. All women not selected for colposcopy received their results and exited the study. Participants and colposcopists were masked to cytology and HPV test results until the colposcopy visit was completed. The primary endpoint for this substudy was histologically confirmed cervical intraepithelial neoplasia grade 3 (CIN3) or worse. This study is registered with ClinicalTrials.gov, number NCT00709891; the study is in the follow-up phase, which is due to be completed in December, 2012.

Findings From May 27, 2008, to Aug 27, 2009, 47 208 women were enrolled, of whom 41 955 met our eligibility criteria. Valid cobas HPV and liquid-based cytology test results were available for 40 901 women (97%), who were included in this analysis. Of these, 4275 women (10%) tested cobas HPV positive and 2617 (6%) had abnormal cytology. 431 women were diagnosed with CIN2 or worse and 274 with CIN3 or worse. In women who had colposcopy, the cobas HPV test was more sensitive than liquid-based cytology for detection of CIN3 or worse (252/274 [92·0%, 95% CI 88·1–94·6] *vs* 146/274 [53·3%, 95% CI 47·4–59·1]; difference 38·7%, 95% CI 31·9–45·5; p<0·0001). Addition of liquid-based cytology to HPV testing increased sensitivity for CIN3 or worse to 96·7% (265/274, 95% CI 93·9–98·3), but increased the number of screen positives by 35·2% (5783/40 901 *vs* 4275/40 901) compared with HPV testing alone. As a triage test to identify CIN3 or worse in HPV-positive women, detection of HPV16, HPV18, or both alone was equivalent to detection of ASC-US or worse alone in terms of sensitivity (150/252 [59·5%] *vs* 133/252 [52·8%]; p=0·11) and positive predictive value (PPV) (150/966 [15·5%] *vs* 133/940 [14·1%]; p=0·20). Among HPV-positive women, detection of HPV16, HPV18, or both or low-grade squamous intraepithelial lesion or worse cytology had better sensitivity (182/252 [72·2%]; p<0·0001) and similar PPV (182/1314 [13·9%]; p=0·70) for detection of CIN3 or worse than ASC-US or worse cytology alone. Furthermore, detection of HPV16, HPV18, or both or high-grade squamous intraepithelial lesion or worse cytology had higher sensitivity (165/252 [65·5%]; p=0·0011) and PPV (165/1013 [16·3%]; p=0·031) for detection of CIN3 or worse than ASC-US or worse cytology alone.

Interpretation HPV testing with separate HPV16 and HPV18 detection could provide an alternative, more sensitive, and efficient strategy for cervical cancer screening than do methods based solely on cytology.

Funding Roche Molecular Systems.

Introduction

Cervical cytology, first with the Papanicolaou (Pap) smear and now with liquid-based cytology, has been the traditional method for cervical cancer screening in developed countries. Since cytology-based programmes were introduced in the mid-20th century, rates of cervical cancer have decreased substantially wherever these screening programmes have been successfully implemented.[1] In the USA, cancer of the cervix is fairly uncommon, with an estimated 12 200 new cases and 4210 related deaths

occurring in 2010.[2] Despite their success, cytology-based screening programmes are now widely recognised as inefficient because of the low sensitivity of one screen and, therefore, many repeat screens are needed during a lifetime to achieve programmatic sensitivity. The US screening and prevention programme costs about US$4 billion annually.[3]

As a consequence, more efficient screening methods are desirable from a comparative performance and cost-effectiveness perspective. Human papillomavirus (HPV) DNA testing, because of its greater sensitivity for cervical precancer and cancer than cytology with one screen, provides lead-time detection of precancerous lesions. In turn, early detection reduces the future risk of cervical cancer[4–6] and related mortality,[7] thereby providing greater reassurance for screen-negative women than can be offered by cytology; as such, HPV testing permits a safe extension of screening intervals.[8–10] The increased sensitivity of HPV testing over cytology also applies to the detection of glandular cancer (adenocarcinoma) and its precursor (adenocarcinoma in situ),[11] which is increasingly important because of the rise in adenocarcinoma rates in the USA,[12] Canada,[13] and Europe.[14] Thus, a switch to HPV testing, either alone or in conjunction with cytology (co-testing), could provide a safer, more efficient screening programme.

A crucial consideration in the use of HPV testing, either as a co-test with cytology or as the primary screening test in cervical cancer screening, is the management of HPV-positive women. According to US guidelines, HPV-positive women who are aged 30 years or older and have normal cervical cytology (negative for intraepithelial lesion or malignancy [NILM]) should be rescreened in 1 year.[15] This approach might be less than optimum because some HPV-positive women with a negative cytology will be lost to follow-up[16] and a small fraction might develop invasive cancer during the rescreening interval.[17] To address these limitations of HPV testing, additional stratification of HPV-positive women with NILM cytology could be used to identify those at greatest risk for cervical precancerous lesions who warrant immediate colposcopy referral. Genotyping for HPV16, HPV18, or both has been proposed as a triage technique, because these two HPV genotypes are associated with about 70% of all invasive cervical carcinomas,[18] but few clinical trial data supporting the clinical performance of genotyping for HPV16, HPV18, or both are available.

Three DNA tests for carcinogenic or high-risk HPV have been approved by the US Food and Drug Administration (FDA): Hybrid Capture 2 (Qiagen, Gaithersburg, MD, USA; 2003), Cervista HPV HR (Hologic, Bedford, MA, USA; 2009), and the cobas HPV test (Roche Molecular Systems, Pleasanton, CA, USA; 2011). The cobas HPV test is a fully automated HPV DNA test that detects in three separate channels: HPV16 individually, HPV18 individually, and a pool of 12 other HPV genotypes (11 definite high-risk genotypes plus one possibly high-risk genotype).[19] Hereafter, HPV refers to carcinogenic or high-risk HPV genotypes.

The ATHENA (Addressing THE Need for Advanced HPV diagnostics) study is the largest clinical trial so far to assess HPV DNA tests and liquid-based cytology for cervical cancer screening in the USA. The study was done in three protocol-specified populations: (1) women aged 21 years or older with atypical squamous cells of undetermined significance (ASC-US) cytology to assess the performance of the assay for the intended use of ASC-US triage to colposcopy, which has been reported previously;[20] (2) women aged 30 years or older with NILM cytology to assess the performance of the assay for the intended use as an adjunctive test with cytology;[21] (3) women aged 25 years and older with any cytology result (ie, independent of cytology) to assess the performance of HPV testing with genotyping as a primary screening test. Here, we present data for the cobas HPV test in the third of these populations.

Methods

Study population

Women presenting for routine cervical cancer screening were enrolled into the ATHENA study at 61 clinical centres in 23 US states (webappendix p 1). Eligible women were aged 21 years or older and were not pregnant. Other study inclusion and exclusion criteria have been described previously[20]—briefly, eligible women had an intact uterus, had not received treatment for CIN with 12 months of enrolment, and had no present or planned participation in a clinical trial for HPV treatment. For this subanalysis, the population was restricted to women aged 25 years and older.

See Online for webappendix

The study protocol was approved by the institutional review boards of all study sites, and all women provided written informed consent before undergoing any study procedures. Institutional review board project number MWP-HPV-159 was obtained on March 13, 2008.

Procedures

Two liquid-based cervical cytology samples were obtained from each participant at study visit one (enrolment visit), placed in PreservCyt (Hologic, Bedford, MA, USA), and processed with ThinPrep (Hologic). One sample was used for liquid-based cytology and for HPV testing with three assays (Roche Molecular Systems, Pleasanton, CA, USA): the first-generation Amplicor HPV test, the first-generation Linear Array HPV genotyping test, and the second-generation cobas HPV test. The second sample was reserved for additional testing. For the purposes of this analysis, a positive result for genotype 16, genotype 18, or both with the cobas HPV test was regarded as positive for HPV16, HPV18, or both, even if the sample was also positive for 12 other high-risk HPV types. Cytology was reported by use of the 2001 Bethesda System nomenclature.[22]

Sample processing and testing were done at four clinical laboratories in the USA, and a fifth laboratory, also in the USA, did cobas HPV testing only; testing was done as described previously.[20] The assignment of specimens to a clinical testing laboratory was made on the basis of the order and volume of patients enrolled at the clinical site, and the capacity of the laboratory; assignment was not based on geographical location.

A randomisation centre identified participants for colposcopy on the basis of the results of liquid-based cytology and first-generation HPV testing at study visit one. Women were eligible for colposcopy if they had ASC-US or worse (positive) cervical cytology, or tested HPV positive by any test and had NILM cervical cytology. A randomly selected subset of women who were HPV negative and had NILM cervical cytology was also referred to colposcopy; randomisation of this subset was done with a block size of 35 by use of SAS software (version 9.1.3). Women not selected for colposcopy received their results and exited the study.

Colposcopy with biopsy or endocervical curettage, or both, was done within 12 weeks of the enrolment visit (at study visit two), according to a standardised protocol.[20] Participants and colposcopists were masked to cytology and HPV test results until the colposcopy visit was completed. Biopsy and endocervical curettage results were reviewed by a panel of three pathologists (central pathology review) who were masked to cytology and HPV test results, and diagnosed according to standard criteria and cervical intraepithelial neoplasia (CIN) terminology.[23] The primary endpoint for the ATHENA study was biopsy-confirmed CIN grade 2 (CIN2) or worse.[24] Women reaching this endpoint exited the study after the colposcopy visit and were referred for treatment; those who did not reach CIN2 were eligible for the longitudinal follow-up phase of the trial, expected to be completed in December, 2012. However, for this subanalysis, we used the primary endpoint of CIN3 or worse because CIN3 is a more certain, rigorous histological diagnosis of precancer than CIN2 or worse. Data are also presented for CIN2 or worse because CIN2 is the standard threshold for treatment,[25] but this endpoint was judged to be secondary because CIN2 is an equivocal diagnosis of precancer, representing an admixture of CIN1 and CIN3.[26–28] CIN2 is poorly reproducible, can be caused by non-carcinogenic HPV genotypes, and is much more likely to regress than CIN3.

Statistical analysis

In accordance with the sample sizes in similar registration trials for the use of HPV testing as a reflex test for ASC-US (ie, an HPV test done on a sample of cervical cells after an ASC-US Pap test result),[29,30] we calculated that about 70 women with ASC-US cytology who had been diagnosed with histological CIN2 or worse would be needed to validate the cobas HPV test for this indication. Together with published rates of ASC-US cytology[31] and HPV positivity[30] in the overall population, we calculated that a sample size of about 45 000 women would be needed. The sample size of about 880 women in the NILM population aged 30 years and older was estimated a priori to have about 95% power to detect a difference in disease (CIN2 or worse) between women who were positive for HPV16, HPV18, or both and those who were HPV negative at baseline, and about 98% power to detect a difference in disease between women who were HPV positive and those who were HPV negative across 3 years.

Several characteristics were used to assess detection of cervical precancer and cancer by HPV testing and liquid-based cytology: sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), positive likelihood ratio (PLR) and negative likelihood ratio (NLR), and relative risk (RR) and, for comparison, cytology. All results are presented for the cobas HPV test, not the first-generation Amplicor or Linear Array HPV tests. Estimates of performance, both crude and those adjusted for sampling fractions to correct for possible verification bias, with 95% CIs, were calculated.[32] Crude estimates were based on results from women with valid results for disease verification. To adjust for verification bias, we calculated how many women had CIN3 or worse (or CIN2 or worse) of those who underwent verification by colposcopy and projected these estimates onto the corresponding women who had no verification. We tested differences in crude sensitivity and specificity by use of McNemar's χ^2 test, and differences in PPV and NPV with the method described by Leisenring and colleagues.[33] For estimates adjusted for verification bias, 95% CIs were estimated by bootstrapping (1000 times).[34] All analyses were done with SAS software (version 9.1.3).

Combinations of genotyping for HPV16, HPV18, or both provided by the cobas HPV test and different thresholds of cytological interpretations (ASC-US or worse, low-grade squamous intraepithelial lesion [LSIL] or worse, or high-grade squamous intraepithelial lesion [HSIL] or worse) were also assessed to identify potentially better algorithms for triage to immediate colposcopy, thereby avoiding problems created by deferred management. To address the main concerns of clinicians, we assessed sensitivity from the proportion of disease detected and PPV from the proportion of screen positives with disease. As a post-hoc analysis, we stratified these data by the four laboratories producing the HPV and liquid-based cytology results, excluding the one laboratory doing HPV testing alone, as a qualitative assessment of the triage method's reliability, which is another important feature of test performance.

The ATHENA study is registered with ClinicalTrials.gov, number NCT00709891; the study is in the follow-up phase, which is due to be completed in December, 2012.

Role of the funding source

The sponsor designed the study in consultation with the other investigators (PEC, MHS, and TCWJr) and was

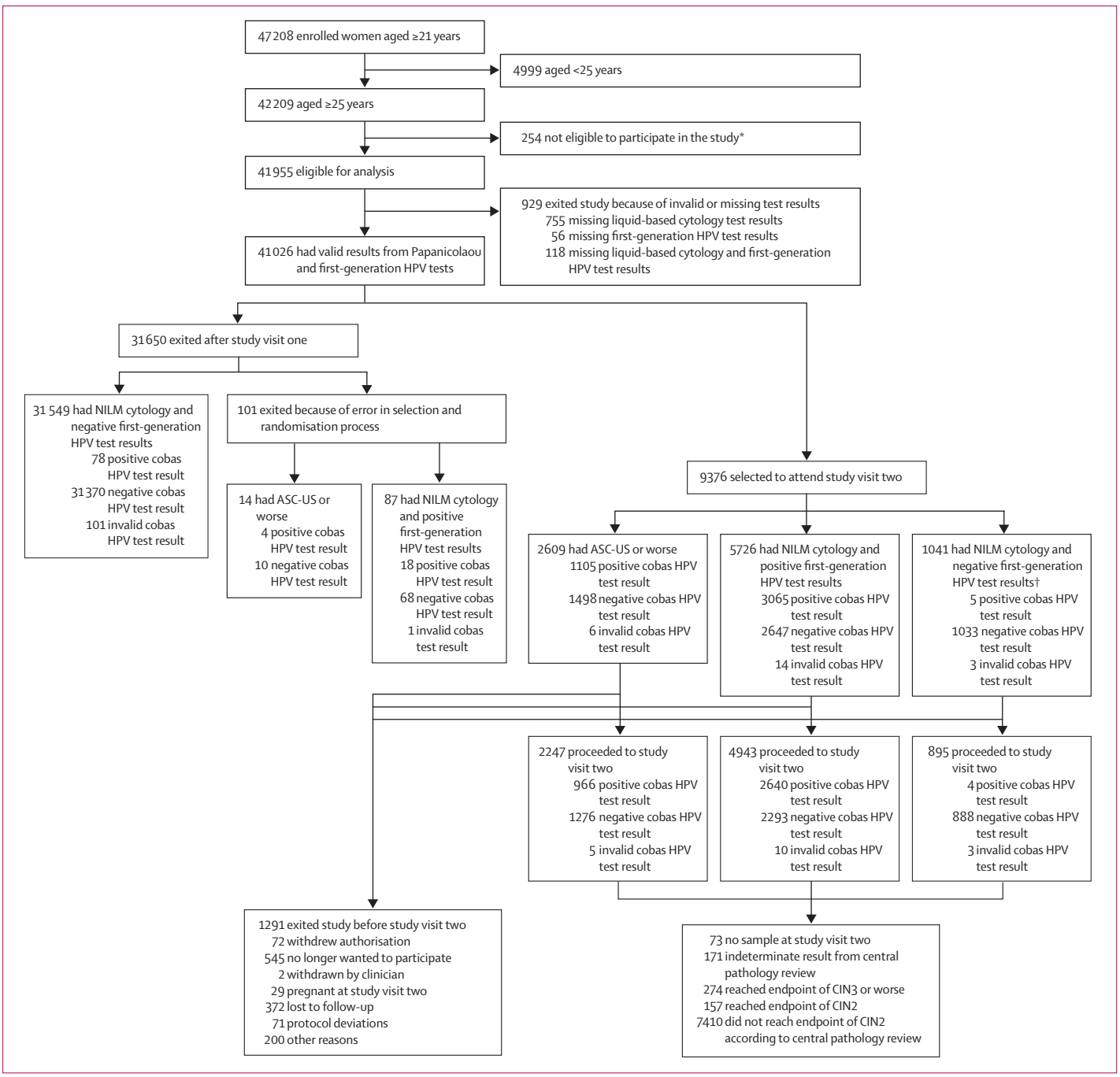


Figure: **Study profile**

First-generation HPV tests were Amplicor and Linear-Array HPV tests. NILM=negative for intraepithelial lesion or malignancy. ASC-US=atypical squamous cells of undetermined significance. CIN2 or CIN3=cervical intraepithelial neoplasia grade 2 or 3. *165 did not meet inclusion or exclusion criteria, 82 enrolled into the study more than once, and seven withdrew authorisation. †Women randomly selected to attend study visit two.

responsible for the study conduct and data collection. Data analysis was done by PEC with the assistance of the sponsor (AS, Guili Zhang, TLW, and CMB). The sponsor participated in the interpretation of the data. The sponsor (AS, TLW, and CMB) assisted in editing and review of the report. PEC had full access to all data in this subanalysis, designed the analytic approach, and had final responsibility for the decision to submit the report for publication.

Results

From May 27, 2008, to Aug 27, 2009, 47 208 women aged 21 years or older were enrolled in the ATHENA study.

42 209 were aged 25 years or older and were included in our subanalysis, of whom 41 955 (99%) met the eligibility criteria (figure). Table 1 shows characteristics of this population. After their enrolment visit, 9376 women were selected or randomly assigned for colposcopy (2609 had ASC-US or worse cervical cytology, 5726 had a positive HPV test by either first-generation HPV test and NILM cervical cytology, and 1041 [3% of 32 590] women were HPV negative and had NILM cervical cytology). Colposcopy was done on 2247 women with ASC-US or worse, 4943 who tested HPV positive, and 895 women with negative HPV and cytology results; 1291 exited before study visit two.

	Eligible women (n=41 955)*
Age (years)	
Mean (SD)	41·9 (11·3)
Median (range)	41 (25–93)
Ethnic origin	
White	34 975 (83%)
Black	5754 (14%)
Asian	657 (2%)
Other or data missing	569 (1%)
Hispanic or Latina ethnic origin†	7558 (18%)
Post-menopausal	13 387 (32%)
Received HPV vaccine	502 (1%)
Immunocompromised	250 (1%)
Smoking history‡	
Non-smoker	29 505 (70%)
Past smoker	6236 (15%)
Present smoker	6213 (15%)
Pap test in past 5 years	38 104 (91%)
HPV test in past 5 years	8378 (20%)
Colposcopy in past 5 years	2938 (7%)
Study cytology (Pap test results)	
NILM	38 433 (92%)
ASC-US or worse	2624 (6%)
ASC-US	1635 (4%)
LSIL	766 (2%)
ASC-H	53 (<1%)
HSIL	113 (<1%)
Squamous cell carcinoma	2 (<1%)
Atypical glandular cells	50 (<1%)
Atypical glandular cells, favour neoplastic	5 (<1%)
Other (endometrial cells)	30 (<1%)
Pap result not available (missing, not done, or invalid)	868 (2%)
Valid HPV test results	
Yes	41 687 (99·4%)
No	268 (0·6%)

Data are number (%) unless otherwise indicated. Pap=Papanicolaou smear. HPV=human papillomavirus. NILM=negative for intraepithelial lesion or malignancy. ASC-US=atypical squamous cells of undetermined significance. LSIL=low-grade squamous intraepithelial lesion. ASC-H=atypical squamous cells, cannot rule out HSIL. HSIL=high-grade squamous intraepithelial lesion. *Women met inclusion and exclusion criteria, had not been enrolled in the study previously, and had not withdrawn consent before undergoing any study procedures. Includes the test results of 929 participants who had exited the study. †Questionnaire design was based on US Food and Drug Administration standard, which regards race and ethnic origin (in this case Hispanic and Latina) as separate. ‡Data are missing for one patient, but the patient was not excluded from the denominator.

Table 1: **Demographic and clinical characteristics of ATHENA population aged 25 years and older at baseline**

Of the 41 955 eligible women, 1054 women had missing or invalid test results: 929 had missing liquid-based cytology or first-generation HPV test results, or both (143 of these women also had invalid cobas HPV test results), and 125 women had invalid cobas HPV test results only. Therefore, valid data from all tests were available for 40 901 women (97%).

According to crude data (uncorrected for verification bias) for outcomes, 10% (4275/40 901) of women tested HPV positive by cobas and 6% (2617/40 901) had abnormal cytology of ASC-US or worse (table 2; difference 4·1%, 95% CI 3·7–4·4; p<0·0001). These 2617 women included 2609 women selected for visit 2, minus six who had invalid cobas HPV test results, plus 14 who exited the study because of an error in randomisation. The crude prevalence of disease was low, with 1·1% (431/40 901) of women diagnosed with CIN2 or worse and 0·7% (274/40 901) of women diagnosed with CIN3 or worse. Notably, in women diagnosed with CIN3, 92% (234/254) tested HPV positive, whereas 52% (132/254) had ASC-US or worse liquid-based cytology (difference 40·2%, 95% CI 33·1–47·2; p<0·0001). Similarly, in women diagnosed with CIN2, 82% (128/157) tested HPV positive and 48% (76/157) had ASC-US or worse liquid-based cytology (difference 33·1%, 95% CI 24·0–42·2; p<0·0001).

In women who had colposcopy, HPV testing was more sensitive for CIN3 or worse than liquid-based cytology at a threshold of ASC-US or worse (table 3; difference 38·7%, 95% CI 31·9–45·5; p<0·0001), but was less specific than liquid-based cytology (difference 16·0%, 95% CI 14·5–17·6; p<0·0001). After adjustment for verification bias, sensitivity decreased substantially for both tests; HPV testing remained more sensitive but less specific for CIN3 or worse than did liquid-based cytology (table 3). RR data showed that testing HPV positive was more strongly associated with CIN3 or worse than was liquid-based cytology; accordingly, NLR was lower for HPV testing than for liquid-based cytology (table 3). These patterns were replicated for the endpoint of CIN2 or worse (table 3). Crude sensitivity for CIN3 or worse increased by only 4·7% (95% CI 2·8–8·0) to 96·7% (265/274, 95% CI 93·9–98·3) for combined liquid-based cytology and HPV testing compared with HPV testing alone (252/274 [92·0%, 95% CI 88·1–94·6]; table 3), but increased the number of screen positives by 35·2% (5783 *vs* 4275 of 40 901).

We examined different combinations of liquid-based cytology and genotyping for HPV16, HPV18, or both for triage to colposcopy for women who were HPV positive (table 4). By comparison with detection of ASC-US or

worse alone, the only available method for triage to colposcopy, detection of HPV16, HPV18, or both was more sensitive and had a greater PPV, but neither difference was significant (table 4). Sensitivity was further increased and PPV was decreased by use of detection of HPV16, HPV18, or both as an additional or complementary triage strategy to ASC-US or worse (table 4). Notably, testing positive for HPV16, HPV18, or both had a sensitivity of 53·8% (64/119, 95% CI 44·9–62·5) and PPV of 10·2% (64/629, 95% CI 8·1–12·8) for CIN3 or worse in women aged 25 years or older who were HPV positive and had NILM cytology. Use of a threshold of LSIL or worse with HPV16, HPV18, or both was more sensitive than detection of ASC-US or worse alone, with similar PPV (table 4). Furthermore, detection of HSIL or worse with HPV16, HPV18, or both had a higher sensitivity and PPV than did ASC-US or worse alone. Similar patterns were recorded for the endpoint of CIN2 or worse, with the exception that detection of HPV16, HPV18, or both did not have higher sensitivity or PPV than detection of ASC-US or worse alone, although these differences were not significant. Furthermore, although detection of HSIL or worse as an additional or alternative triage strategy to detection of HPV16, HPV18, or both increased sensitivity and PPV

	LBC+	HPV+	LBC– and HPV–	LBC– and HPV+	LBC+ and HPV–	LBC+ and HPV+
No colposcopy (n=32 834)	375 (1%)	665 (2%)	31 937 (97%)	522 (2%)*	232 (1%)	143 (<1%)
Out of time window (n=60)†	18 (30%)	27 (45%)	23 (38%)	19 (32%)	10 (17%)	8 (13%)
Colposcopy or no biopsy (n=184)‡	38 (21%)	81 (44%)	83 (45%)	63 (34%)*	20 (11%)	18 (10%)
Negative biopsy (n=6802)	1704 (25%)	2756 (41%)	2922 (43%)	2176 (32%)*	1124 (17%)	580 (9%)
CIN1 (n=590)	260 (44%)	366 (62%)	124 (21%)	206 (35%)*	100 (17%)	160 (27%)
CIN2 (n=157)	76 (48%)	128 (82%)	20 (13%)	61 (39%)*	9 (6%)	67 (43%)
CIN3 (n=254)	132 (52%)	234 (92%)	9 (4%)	113 (44%)*	11 (4%)	121 (48%)
Adenocarcinoma in situ (n=16)	10 (63%)	14 (88%)	0	6 (38%)	2 (13%)	8 (50%)
Squamous cell carcinoma (n=3)	3 (100%)	3 (100%)	0	0	0	3 (100%)
ADC or ASC (n=1)	1 (100%)	1 (100%)	0	0	0	1 (100%)
Total (n=40 901)	2617§ (6%)	4275 (10%)	35 118 (86%)	3166 (8%)	1508 (4%)	1109 (3%)

LBC=liquid-based cytology. +=positive. HPV=human papillomavirus. –=negative. CIN1, CIN2, or CIN3=cervical intraepithelial neoplasia grade 1, 2, or 3. ADC=adenocarcinoma. ASC=adenosquamous carcinoma. *p<0·05 for LBC *vs* HPV with McNemar's χ^2 test. †Colposcopy visit happened after at least 12 months. ‡Includes inadequate biopsy for diagnosis and no biopsy taken for various reasons. §Includes 2609 women selected for visit 2, minus six who had invalid cobas HPV test results, plus 14 who exited the study because of an error in randomisation.

Table 2: **Diagnosis of biopsy by central pathology review panel, stratified by liquid-based cytology and HPV test results**

	Crude data			Adjusted data*	
	Liquid-based cytology	Cobas HPV test	p value	Liquid-based cytology	Cobas HPV test
CIN3 or worse					
Sensitivity	146/274 (53·3%, 47·4–59·1)	252/274 (92·0%, 88·1–94·6)	<0·0001	174/403 (43·2%, 32·1–55·9)	302/402 (75·1%, 58·4–94·1)
Specificity	5509/7549 (73·0%, 72·0–74·0)	4299/7549 (56·9%, 55·8–58·1)	<0·0001	38 055/40 498 (94·0%, 93·7–94·2)	36 526/40 499 (90·2%, 89·9–90·5)
PPV	146/2186 (6·7%, 6·0–7·4)	252/3502 (7·2%, 6·9–7·5)	0·20	174/2617 (6·6%, 5·5–7·6)	302/4275 (7·1%, 6·2–7·8)
NPV	5509/5637 (97·7%, 97·4–98·0)	4299/4321 (99·5%, 99·2–99·7)	<0·0001	38 055/38 284 (99·4%, 99·1–99·6)	36 526/36 626 (99·7%, 99·4–100·0)
PLR	1·97 (1·75–2·22)	2·14 (2·05–2·23)	..	7·16 (5·30–9·25)	7·66 (5·89–9·62)
NLR	0·64 (0·56–0·73)	0·14 (0·09–0·21)	..	0·60 (0·47–0·72)	0·28 (0·07–0·46)
RR	2·9 (2·3–3·7)	14·1 (9·2–21·8)	..	11·1 (9·2–13·5)	25·9 (20·7–32·4)
CIN2 or worse					
Sensitivity	222/431 (51·5%, 46·8–56·2)	380/431 (88·2%, 84·8–90·9)	<0·0001	265/738 (35·9%, 27·5–46·6)	455/738 (61·7%, 48·0–78·9)
Specificity	5428/7392 (73·4%, 72·4–74·4)	4270/7392 (57·8%, 56·6–58·9)	<0·0001	37 811/40 163 (94·1%, 93·9–94·4)	36 343/40 163 (90·5%, 90·2–90·8)
PPV	222/2186 (10·2%, 9·3–11·1)	380/3502 (10·9%, 10·4–11·3)	0·17	265/2617 (10·1%, 8·7–11·3)	455/4275 (10·6%, 9·6–11·6)
NPV	5428/5637 (96·3%, 95·9–96·6)	4270/4321 (98·8%, 98·5–99·1)	<0·0001	37 811/38 284 (98·8%, 98·3–99·2)	36 343/36 626 (99·2%, 98·7–99·7)
PLR	1·94 (1·76–2·14)	2·09 (2·00–2·18)	..	6·13 (4·64–8·02)	6·48 (5·00–8·38)
NLR	0·66 (0·60–0·73)	0·20 (0·16–0·27)	..	0·68 (0·57–0·77)	0·42 (0·23–0·58)
RR	2·7 (2·3–3·3)	9·2 (6·9–12·3)	..	8·2 (7·1–9·5)	13·8 (11·9–15·9)

Data are n/N (%, 95% CI) or point estimate (95% CI). Adjusted p values are two-sided; it is not possible to derive p values for verification bias adjusted data. HPV=human papillomavirus. CIN2 or CIN3=cervical intraepithelial neoplasia grade 2 or 3. PPV=positive predictive value. NPV=negative predictive value. PLR=positive likelihood ratio. NLR=negative likelihood ratio. RR=relative risk. Pap=Papanicolaou smear. *Data are adjusted for verification bias; fractions are projected numbers in the overall population, not actual numbers.

Table 3: **Clinical performance of Pap and HPV testing for identifying women with a biopsy diagnosis of CIN3 or CIN2, or worse**

compared with ASC-US or worse alone, these differences were not significant (table 4).

To assess test reliability as a post-hoc analysis, we aimed to qualitatively describe the variability in triage performance of ASC-US or worse, LSIL or worse, and HSIL or worse by testing centre and compare it with combinations of detection of HPV16, HPV18, or both with cytological thresholds. The variability in triage sensitivity for CIN3 or worse of any cytological threshold varied by about two times between testing centres (p values of 0·01–0·013; webappendix p 2). By comparison, a triage strategy that included detection of HPV16, HPV18, or both with any cytology cutoff was much less variable (p values of 0·30–0·72) and overall was equally or more sensitive than triage with ASC-US or worse. These findings were replicated for an endpoint of CIN2 or worse.

	Sensitivity		PPV		RR (95% CI)*	PLR (95% CI)
	n/N (%, 95% CI)	p value	n/N (%, 95% CI)	p value		
CIN3 or worse						
None (all HPV+ to colposcopy)	252/252 (100·0%, 98·5–100·0)	..	252/3502 (7·2%, 7·2–7·2)	..	..	..
ASC-US or worse	133/252 (52·8%, 46·6–58·9)	..	133/940 (14·1%, 12·6–15·8)	..	3·0 (2·4–3·9)	2·1 (1·9–2·4)
LSIL or worse	101/252 (40·1%, 34·2–46·2)	<0·0001	101/564 (17·9%, 15·5–20·6)	<0·0001	3·5 (2·8–4·4)	2·8 (2·4–3·3)
HSIL or worse	66/252 (26·2%, 21·1–31·9)	<0·0001	66/131 (50·4%, 42·5–58·2)	<0·0001	9·1 (7·3–11·4)	13·1 (9·5–18·0)
HPV16+	127/252 (50·4%, 44·3–56·5)	0·58	127/693 (18·3%, 16·3–20·6)	0·00016	4·1 (3·3–5·2)	2·9 (2·5–3·3)
HPV16+ or ASC-US or worse	188/252 (74·6%, 68·9–79·6)	<0·0001	188/1380 (13·6%, 12·7–14·7)	0·40	4·5 (3·4–5·9)	2·0 (1·9–2·2)
HPV16+ and ASC-US or worse	72/252 (28·6%, 23·3–34·4)	<0·0001	72/253 (28·5%, 23·8–33·6)	<0·0001	5·1 (4·0–6·5)	5·1 (4·0–6·5)
HPV16+ or LSIL or worse	172/252 (68·3%, 62·3–73·7)	<0·0001	172/1089 (15·8%, 14·5–17·2)	0·038	4·8 (3·7–6·2)	2·4 (2·2–2·7)
HPV16+ and LSIL or worse	56/252 (22·2%, 17·5–27·8)	<0·0001	56/168 (33·3%, 27·1–40·2)	<0·0001	5·7 (4·4–7·3)	6·4 (4·8–8·7)
HPV16+ or HSIL or worse	152/252 (60·3%, 54·2–66·2)	0·046	152/755 (20·1%, 18·2–22·2)	<0·0001	5·5 (4·4–7·0)	3·3 (2·9–3·7)
HPV16+ and HSIL or worse	41/252 (16·3%, 12·2–21·3)	<0·0001	41/69 (59·4%, 48·0–69·9)	<0·0001	9·7 (7·6–12·2)	18·9 (11·9–30·0)
HPV16+ or HPV18+	150/252 (59·5%, 53·4–65·4)	0·11	150/966 (15·5%, 14·0–17·1)	0·20	3·9 (3·0–4·9)	2·4 (2·1–2·7)
(HPV16+ or HPV18+) or ASC-US or worse	197/252 (78·2%, 72·7–82·8)	<0·0001	197/1569 (12·6%, 11·7–13·4)	0·017	4·4 (3·3–5·9)	1·9 (1·7–2·0)
(HPV16+ or HPV18+) and ASC-US or worse	86/252 (34·1%, 28·16–40·2)	<0·0001	86/337 (25·5%, 21·8–29·7)	<0·0001	4·9 (3·8–6·2)	4·4 (3·6–5·4)
(HPV16+ or HPV18+) or LSIL or worse	182/252 (72·2%, 66·4–77·4)	<0·0001	182/1314 (13·9%, 12·8–15·0)	0·70	4·3 (3·3–5·7)	2·1 (1·9–2·3)
(HPV16+ or HPV18+) and LSIL or worse	69/252 (27·4%, 22·2–33·2)	<0·0001	69/216 (31·9%, 26·7–37·7)	<0·0001	5·7 (4·5–7·3)	6·1 (4·7–7·8)
(HPV16+ or HPV18+) or HSIL or worse	165/252 (65·5%, 59·4–71·1)	0·0011	165/1013 (16·3%, 14·9–17·8)	0·031	4·7 (3·6–6·0)	2·5 (2·3–2·8)
(HPV16+ or HPV18+) and HSIL or worse	51/252 (20·2%, 15·7–25·6)	<0·0001	51/84 (60·7%, 50·4–70·1)	<0·0001	10·3 (8·3–12·8)	19·9 (13·1–30·3)
CIN2 or worse						
None (all HPV+ to colposcopy)	380/380 (100·0%, 99·0–100·0)	..	380/3502 (10·9%, 10·8–10·9)	..	..	..
ASC-US or worse	200/380 (52·6%, 47·6–57·6)	..	200/940 (21·3%, 19·4–23·3)	..	3·0 (2·5–3·7)	2·2 (2·0–2·5)
LSIL or worse	149/380 (39·2%, 34·4–44·2)	<0·0001	149/564 (26·4%, 23·5–29·5)	<0·0001	3·4 (2·8–4·0)	2·9 (2·5–3·4)
HSIL or worse	77/380 (20·3%, 16·5–24·6)	<0·0001	77/131 (58·8%, 50·6–66·5)	<0·0001	6·5 (5·5–7·8)	11·7 (8·4–16·3)
HPV16+	168/380 (44·2%, 39·3–49·2)	0·015	168/693 (24·2%, 21·8–26·9)	0·055	3·2 (2·7–3·9)	2·6 (2·3–3·0)
HPV16+ or ASC-US or worse	270/380 (71·1%, 66·3–75·4)	<0·0001	270/1380 (19·6%, 18·3–20·8)	0·015	3·8 (3·1–4·7)	2·0 (1·8–2·2)
HPV16+ and ASC-US or worse	98/380 (25·8%, 21·6–30·4)	<0·0001	98/253 (38·7%, 33·5–44·3)	<0·0001	4·5 (3·7–5·4)	5·2 (4·1–6·5)
HPV16+ or LSIL or worse	244/380 (64·2%, 59·3–68·9)	<0·0001	244/1089 (22·4%, 20·8–24·1)	0·22	4·0 (3·3–4·8)	2·4 (2·2–2·6)
HPV16+ and LSIL or worse	73/380 (19·2%, 15·6–23·5)	<0·0001	73/168 (43·5%, 36·6–50·6)	<0·0001	4·7 (3·9–5·8)	6·3 (4·7–8·4)
HPV16+ or HSIL or worse	199/380 (52·4%, 47·3–57·3)	0·93	199/753 (26·4%, 24·1–28·8)	0·0003	4·0 (3·3–4·8)	2·9 (2·6–3·3)
HPV16+ and HSIL or worse	46/380 (12·1%, 9·2–15·8)	<0·0001	46/69 (66·7%, 55·1–76·5)	<0·0001	6·9 (5·6–8·3)	16·4 (10·1–26·8)
HPV16+ or HPV18+	197/380 (51·8%, 46·8–56·8)	0·82	197/966 (20·4%, 18·6–22·3)	0·50	2·8 (2·3–3·4)	2·1 (1·9–2·4)
(HPV16+ or HPV18+) or ASC-US or worse	283/380 (74·5%, 69·9–78·6)	<0·0001	283/1569 (18·0%, 17·0–19·1)	<0·0001	3·6 (2·9–4·5)	1·8 (1·7–1·9)
(HPV16+ or HPV18+) and ASC-US or worse	114/380 (30·0%, 25·6–34·8)	<0·0001	114/337 (33·8%, 29·5–38·4)	<0·0001	4·0 (3·3–4·9)	4·2 (3·4–5·1)
(HPV16+ or HPV18+) or LSIL or worse	258/380 (67·9%, 63·0–72·4)	<0·0001	258/1314 (19·6%, 18·3–21·0)	0·074	3·5 (2·9–4·3)	2·0 (1·8–2·2)
(HPV16+ or HPV18+) and LSIL or worse	88/380 (23·2%, 19·2–27·7)	<0·0001	88/216 (40·7%, 34·9–46·9)	<0·0001	4·6 (3·8–5·6)	5·6 (4·4–7·3)
(HPV16+ or HPV18+) or HSIL or worse	217/380 (57·1%, 52·1–62·0)	0·16	217/1013 (21·4%, 19·7–23·3)	0·90	3·3 (2·7–4·0)	2·2 (2·0–2·5)
(HPV16+ or HPV18+) and HSIL or worse	57/380 (15·0%, 11·8–18·9)	<0·0001	57/84 (67·9%, 57·5–76·7)	<0·0001	7·2 (6·0–8·6)	17·3 (11·1–27·1)

All data are crude estimates. p values represent comparison of various triage strategies to ASC-US or worse. ASC-US or worse includes cytological interpretations of HSIL, ASC-H, AGC, and LSIL. LSIL or worse includes HSIL, ASC-H, and AGC. PPV=positive predictive value. RR=relative risk. PLR=positive likelihood ratio. CIN2 or CIN3=cervical intraepithelial neoplasia grade 2 or 3. HPV=human papillomavirus. +=positive. ASC-US=atypical squamous cells of undetermined significance. LSIL=low-grade squamous intraepithelial lesion. HSIL=high-grade squamous intraepithelial lesion. AGC=atypical glandular cell. ASC-H=atypical squamous cells, cannot rule out HSIL. *Calculated as the risk (or probability) of CIN3 or worse in test positives divided by the risk of CIN3 or worse in test negatives.

Table 4: **Strategies for triage to colposcopy for women who test HPV positive for the immediate identification of women with CIN3 or CIN2, or worse**

Discussion

In our comparison of HPV testing with liquid-based cytology, we showed that HPV testing was more sensitive, albeit less specific, than liquid-based cytology for identification of women with CIN3 or worse, confirming the findings of previous clinical trials in Europe[4–6,35] and Canada[36] (panel).

Combination of liquid-based cytology and HPV testing for CIN3 or worse increased sensitivity by less than 5% and increased the number of screen positives by more than a third compared with HPV testing alone. By contrast, use of HPV16 or HPV18 detection as an additional or alternative triage strategy to reproducible cytological abnormalities (LSIL or worse, or HSIL or worse) resulted in increased, more reliable (inter-laboratory) performance for identification of women with CIN3 or worse compared with the use of ASC-US or worse cytology alone.

In view of our observations, use of HPV testing as the primary screening test to rule out disease, and use of a specific test, like liquid-based cytology, to decide which women need immediate colposcopy, seems to be a rational approach. Use of either HPV testing or cytology in HPV-positive women sends almost the same group of women to immediate colposcopy as would the existing strategies of HPV and cytology co-testing and cytology-based screening—ie, women who are HPV positive and have ASC-US cytology or those with cytology of LSIL or worse. Most women with LSIL or worse are HPV positive except for those with the rare atypical glandular cell diagnosis, which, when HPV negative, is related mostly to endometrial rather than cervical abnormalities.[39]

Our findings support the premise that co-testing has little benefit over HPV testing alone for clinical performance. Addition of cytology to HPV testing increased safety (NPV) slightly compared with HPV testing alone for identification of CIN3 or worse (3066/3075, 99·7%, 95% CI 99·4–99·9 *vs* 4299/4321, 99·5%, 95% CI 99·2–99·7; p=0·0137). However, until clinicians become comfortable with use of HPV as a first-line test, they might initially favour co-testing, and so co-testing could have an underlying merit that is difficult to quantify.[40] Since none of the methods will be perfect for prevention of cervical cancer, the decision to switch from co-testing to HPV testing alone, and the intervals between screenings, will ultimately depend on clinicians' perceptions of acceptable risks.[41,42]

The management of HPV-positive women with negative cytology results remains a clinical dilemma.[43] Although some guidelines recommend rescreening of such women in 1 year,[15,44] this strategy has substantial drawbacks. First, some women will already have CIN3 or worse, which includes a small but appreciable number of women with invasive cervical cancer.[17] Second, loss to follow-up in this population can be high (about 50%)[16,45] and, as such, can offset the gain in sensitivity (but not NPV) of use of HPV testing for primary screening. In view of these issues, an immediate triage strategy is needed for this group.

Findings from an epidemiological study[46] indicate that detection of HPV16, HPV18, or both might be used in HPV-positive women to identify those at increased risk of CIN3 or worse who need immediate colposcopy.[15,47] In our study, we showed that detection of HPV16, HPV18, or both provided at least equivalent and more reliable clinical performance than did detection of ASC-US or worse with liquid-based cytology for triage to immediate colposcopy for all HPV-positive women. ASC-US is the most common abnormal cytological interpretation[15,48] and, as an equivocal result, the least reproducible. Thus, use of liquid-based cytology to detect ASC-US or worse (most cases of which are ASC-US in general) as a triage strategy for HPV-positive women could result in substantial variation in performance because of well-documented inter-laboratory differences in diagnosing ASC-US.[49] The variability in the performance of cytology might,

Panel: Research in context

Systematic review

The comparative performance of HPV DNA testing versus conventional cytology in primary cervical cancer screening has been documented,[5,6,36–38] but no studies have reported the performance of HPV DNA testing with individual genotyping for HPV16, HPV18, or both in a large screening population. In a review of North American (Canada and USA) and European screening studies, HPV DNA testing was more sensitive than cytology for detection of CIN2 or worse (96·1% *vs* 53·0%) but was less specific (90·7% *vs* 96·3%).[37] The screening studies were done in populations that were similar in age to the ATHENA population, and HPV DNA testing was generally done with Hybrid Capture 2.[37] Similarly, in a Canadian study of more than 10 000 women, HPV DNA testing was more sensitive than cytology for detection of CIN2 or worse (94·6% *vs* 55·4%) but was less specific (94·1% *vs* 96·8%).[36] As in the ATHENA trial, all women with ASC-US cytology or worse, all women who had normal cytology but were high-risk HPV positive, and a subset of women negative for both, were referred to colposcopy, and similar adjustments were made for verification bias.[36] In studies of algorithms incorporating HPV DNA testing for screening of cervical cancer, primary HPV DNA testing detected disease earlier than did conventional cytology.[5,6] Furthermore, triage with primary HPV DNA testing and cytology, and repeat HPV DNA testing of women with negative cytology, was shown to be a realistic strategy for cervical cancer screening.[38]

Interpretation

ATHENA is the largest US registration trial to assess the performance of HPV DNA testing with individual genotyping for HPV16, HPV18, or both compared with liquid-based cytology for cervical cancer screening. Our findings show that HPV DNA testing has higher sensitivity than cytology; detection of HPV16, HPV18, or both alone has similar sensitivity to ASC-US or worse cytology; and detection of HPV16, HPV18, or both in combination with low-grade squamous intraepithelial lesion or worse cytology has better sensitivity than ASC-US or worse cytology. HPV testing with individual genotyping for HPV16, HPV18, or both could provide a more efficient strategy for cervical cancer screening than do existing programmes based on cytology.

HPV=human papillomavirus. CIN2 or CIN3=cervical intraepithelial neoplasia grade 2 or 3. ASC-US=atypical squamous cells of undetermined significance.

however, be lower in places such as Europe, where testing is more centralised and might have better quality control than in the USA. Nevertheless, detection of HPV16 and HPV18 provides an objective measurement of risk and is already provided by the cobas HPV test, making this indicator easy to integrate into management algorithms.

One such algorithm, combining ASC-US or worse with detection of HPV16, HPV18, or both to triage women to colposcopy, would use the same criteria for referral as would use of HPV16 or HPV18 detection to triage HPV-positive women and negative cytology,[15,47] and would result in higher, more reliable sensitivity but with a lower PPV than ASC-US or worse alone. Furthermore, use of combinations of the most specific and severe cytological interpretations of abnormality (LSIL or worse, or HSIL or worse) with detection of HPV16, HPV18, or both provides an even more sensitive method to triage HPV-positive women than does either HPV16 or HPV18 detection, or ASC-US or worse alone, while retaining good PPV. The tradeoffs in sensitivity for immediate detection of CIN3 or worse versus the over-referral and poor PPV, together with the added costs to maintain cytology as part of this algorithm, will have to be considered by professional medical groups, who will ultimately decide the preferred or acceptable methods to triage women who are HPV positive. Nevertheless, on the basis of our findings, we suggest that detection of HPV16, HPV18, or both combined with a raised threshold of abnormal cervical cytology (LSIL or worse) might be preferable to the existing recommendations for management of HPV-positive women.

We acknowledge that the imperfect sensitivity of any triage strategies will result in some women with CIN3 or worse failing to get immediate colposcopy and being deferred to a 1-year follow-up. For example, if HPV16 or HPV18 detection or LSIL or worse was the strategy used for triage, 27·8% of HPV-positive women with 13·9% risk of CIN3 or worse would be deferred to a 1-year follow-up, compared with 47·2% of HPV-positive women with 14·1% risk of CIN3 or worse if ASC-US or worse was used for triage (table 4). Although sending all HPV-positive women to immediate colposcopy would obviate this concern, doing so comes at a cost of excessive referral to colposcopy, and detection and treatment of some CIN2 that might have otherwise regressed.[6] Each strategy has a tradeoff between programmatic sensitivity and specificity.

We also acknowledge that this cross-sectional analysis did not allow measurement of any benefit of lead-time detection by HPV testing of CIN3 or adenocarcinoma in situ for reduction of subsequent cancer incidence. However, because immediate detection and treatment of all HPV-positive precancerous lesions reduces the subsequent incidence of cancer[6] and cancer-related mortality,[7] it seems logical that increasing the proportion of CIN3 or adenocarcinoma in situ that is immediately detected and treated will ultimately reduce cancer risk. However, in view of the low incidence of CIN3 or adenocarcinoma in situ, such a benefit might only be observable in large, organised programmes.[11]

We noted that the clinical performance of HPV testing and liquid-based cytology was reduced after correction for verification bias due to identification of some CIN3 in the HPV-negative, cytology-negative subgroup. In the clinical trial of HPV and Pap testing in Canada, a similar reduction in the estimated overall performance occurred after adjustment for verification bias.[36] The most likely explanation is that these rare but true cases of CIN3 or worse are missed by both HPV and cytology testing, resulting in double false negatives, because of poor sampling.

Alternatively, at least some of these cases might have morphological changes that mimic the appearance of precancerous lesions, resulting in misclassification as CIN3 or worse.[50] These lookalike lesions are not related to cervical cancer risk, but have to be included in our endpoints because we cannot accurately differentiate them from true precancerous lesions. Of the nine cases of CIN3 diagnosed in women who were negative by both cobas HPV test and liquid-based cytology, four were negative for p16 immunohistochemistry (mtm laboratories, Heidelberg, Germany),[51] two of whom were also negative by Linear Array and Amplicor HPV tests. One p16-positive case tested positive for HPV82 by Linear Array HPV test, an HPV genotype that rarely if ever causes cervical cancer[52] and therefore is not targeted by any FDA-approved tests. Thus, about half of these cases were possibly falsely diagnosed or caused by an HPV genotype that is not targeted by the cobas HPV test. Cases that are negative for both HPV and CIN3 or worse according to liquid-based cytology do not present a substantial risk of cervical cancer, because well screened populations are at a very low risk of cervical cancer. Even with the inclusion of the falsely diagnosed cases, the sensitivity of HPV testing was still better than liquid-based cytology.

In summary, on the basis of our findings and other published data, we propose that rational use of HPV testing (and genotyping for HPV16, HPV18, or both) with or without liquid-based cytology can provide potentially cost-effective[53] and safe cervical cancer screening. Because HPV16 and HPV18 readouts for the cobas HPV test are provided concurrently with the pooled detection of other carcinogenic HPV genotypes, testing for HPV16 and HPV18 to triage HPV-positive women could be very efficient and reduce manpower requirements in clinical laboratories compared with cytology. We have also shown that cytology could be applied reflexively to women who are HPV positive without the HPV16 or HPV18 genotype, with women referred to colposcopy only if they have LSIL or HSIL, or worse. This strategy would increase the sensitivity for detection of CIN3 or worse in HPV-positive women above that provided by detection of HPV16,

HPV18, or both, while maintaining good PPV. Future studies will need to assess the comparative performance and cost-effectiveness of the different cervical cancer strategies to identify best practices.

Contributors

PEC designed and did data analyses and wrote the report. MHS participated in the study analysis, data interpretation, review of pathological slides, and editing of the report. TCWJr participated in study design and analysis, data interpretation, review of pathological slides, and writing and reviewing the report. AS directed data analysis. TLW designed the study, supervised study execution, supervised data analyses, reviewed the report and made comments, and reviewed the final draft. CMB participated in the study execution, data analysis, and writing and editing of the report.

Conflicts of interest

PEC has a non-disclosure agreement to work with Roche on the analysis of their clinical trial but receives no financial compensation. MHS has received consulting fees and honoraria from Roche Molecular Systems; and has been a paid consultant for BD, Qiagen, Gen-Probe, Ventana, and Merck. TCWJr has received consultancy fees and honoraria, received payment for lectures, and financial support for travel to meetings by Roche Molecular Systems, mtm laboratories, BD Diagnostics, Gen-Probe, and Merck. AS, TLW, and CMB are employees of Roche Molecular Systems; and TLW has stock and stock options in Roche Molecular Systems.

Acknowledgments

This research was funded by Roche Molecular Systems, Pleasanton, CA, USA. Guili Zhang assisted in data analysis. Victoria Tomlinson (Health Interactions, London, UK) provided editorial assistance for formatting and proofing of the final draft of the report.

References

1 International Agency for Research on Cancer. Cervical cancer screening. Lyon: International Agency for Research on Cancer Press, 2005.

2 Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA Cancer J Clin* 2010; **60:** 277–300.

3 Insinga RP, Dasbach EJ, Elbasha EH. Assessing the annual economic burden of preventing and treating anogenital human papillomavirus-related disease in the US: analytic framework and review of the literature. *Pharmacoeconomics* 2005; **23:** 1107–22.

4 Anttila A, Kotaniemi-Talonen L, Leinonen M, et al. Rate of cervical cancer, severe intraepithelial neoplasia, and adenocarcinoma in situ in primary HPV DNA screening with cytology triage: randomised study within organised screening programme. *BMJ* 2010; **340:** c1804.

5 Bulkmans NWJ, Berkhof J, Rozendaal L, et al. 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* 2007; **370:** 1764–72.

6 Ronco G, Giorgi-Rossi P, Carozzi F, et al, and the New Technologies for Cervical Cancer screening (NTCC) Working Group. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. *Lancet Oncol* 2010; **11:** 249–57.

7 Sankaranarayanan R, Nene BM, Shastri SS, et al. HPV screening for cervical cancer in rural India. *N Engl J Med* 2009; **360:** 1385–94.

8 Dillner J, Rebolj M, Birembaut P, et al. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. *BMJ* 2008; **337:** a1754.

9 Kjaer SK, Frederiksen K, Munk C, Iftner T. Long-term absolute risk of cervical intraepithelial neoplasia grade 3 or worse following human papillomavirus infection: role of persistence. *J Natl Cancer Inst* 2010; **102:** 1478–88.

10 Sherman ME, Lorincz AT, Scott DR, et al. Baseline cytology, human papillomavirus testing, and risk for cervical neoplasia: a 10-year cohort analysis. *J Natl Cancer Inst* 2003; **95:** 46–52.

11 Katki HA, Kinney WK, Fetterman B, et al. Cervical cancer risk for women undergoing concurrent testing for human papillomavirus and cervical cytology: a population-based study in routine clinical practice. *Lancet Oncol* 2011; **12:** 663–72.

12 Wang SS, Sherman ME, Hildesheim A, Lacey JV Jr, Devesa S. Cervical adenocarcinoma and squamous cell carcinoma incidence trends among white women and black women in the United States for 1976–2000. *Cancer* 2004; **100:** 1035–44.

13 Liu S, Semenciw R, Probert A, Mao Y. Cervical cancer in Canada: changing patterns in incidence and mortality. *Int J Gynecol Cancer* 2001; **11:** 24–31.

14 Bray F, Carstensen B, Moller H, et al. Incidence trends of adenocarcinoma of the cervix in 13 European countries. *Cancer Epidemiol Biomarkers Prev* 2005; **14:** 2191–99.

15 Wright TC Jr, Massad LS, Dunton CJ, Spitzer M, Wilkinson EJ, Solomon D. 2006 consensus guidelines for the management of women with abnormal cervical cancer screening tests. *Am J Obstet Gynecol* 2007; **197:** 346–55.

16 Kitchener HC, Almonte M, Thomson C, et al. HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. *Lancet Oncol* 2009; **10:** 672–82.

17 Kinney W, Fetterman B, Cox JT, Lorey T, Flanagan T, Castle PE. Characteristics of 44 cervical cancers diagnosed following Pap-negative, high risk HPV-positive screening in routine clinical practice. *Gynecol Oncol* 2011; **121:** 309–13.

18 de Sanjose S, Quint WGV, Alemany L, et al, on behalf of the Retrospective International Survey and HPV Time Trends Study Group. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol* 2010; **11:** 1048–56.

19 Castle PE, Sadorra M, Lau T, Aldrich C, Garcia FA, Kornegay J. Evaluation of a prototype real-time PCR assay for carcinogenic human papillomavirus (HPV) detection and simultaneous HPV genotype 16 (HPV16) and HPV18 genotyping. *J Clin Microbiol* 2009; **47:** 3344–47.

20 Stoler MH, Wright TC Jr, Sharma A, Apple R, Gutekunst K, Wright TL. High-risk human papillomavirus testing in women with ASC-US cytology: results from the ATHENA HPV study. *Am J Clin Pathol* 2011; **135:** 468–75.

21 Wright TC Jr, Stoler MH, Sharma A, Zhang G, Behrens CM, Wright TL. Evaluation of HPV-16 and HPV-18 genotyping for the triage of high-risk HPV-positive cytology-negative women. *Am J Clin Pathol* (in press).

22 Solomon D, Davey D, Kurman R, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA* 2002; **287:** 2114–19.

23 Kurman RJ, Ellenson LH, Ronnett BM. Blaustein's pathology of the female genital tract, 5th edn. New York, NY: Springer, 2002.

24 US Department of Health and Human Services, Food and Drug Administration, Center for Device and Radiological Health, 2009. Establishing the performance characteristics of in vitro diagnostic devices for the detection or detection and differentiation of human papillomaviruses (draft guidance). http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM181511.pdf (accessed Aug 4, 2011).

25 Wright TC Jr, Massad LS, Dunton CJ, Spitzer M, Wilkinson EJ, Solomon D. 2006 consensus guidelines for the management of women with cervical intraepithelial neoplasia or adenocarcinoma in situ. *J Low Genit Tract Dis* 2007; **11:** 223–39.

26 Castle PE, Stoler MH, Solomon D, Schiffman M. The relationship of community biopsy-diagnosed cervical intraepithelial neoplasia grade 2 to the quality control pathology-reviewed diagnoses: an ALTS report. *Am J Clin Pathol* 2007; **127:** 805–15.

27 Castle PE, Schiffman M, Wheeler CM, Solomon D. Evidence for frequent regression of cervical intraepithelial neoplasia-grade 2. *Obstet Gynecol* 2009; **113:** 18–25.

28 Carreon JD, Sherman ME, Guillen D, et al. CIN2 is a much less reproducible and less valid diagnosis than CIN3: results from a histological review of population-based cervical samples. *Int J Gynecol Pathol* 2007; **26:** 441–46.

29 Third Wave Technologies. Cervista HPV HR test instructions for use. Madison, WI: Third Wave Technologies, 2009.

30 Qiagen. HC2 high-risk HPV DNA test (package insert). Gaithersburg, MD: Qiagen, 2007.

31 ALTS Group. Results of a randomized trial on the management of cytology interpretations of atypical squamous cells of undetermined significance. *Am J Obstet Gynecol* 2003; **188:** 1383–92.

32 Zhou XH, Obuchowski NA, Obuchowski DM. Statistical methods in diagnostic medicine. New York, NY: John Wiley and Sons, 2002.

33 Leisenring W, Alonzo T, Pepe MS. Comparisons of predictive values of binary medical diagnostic tests for paired designs. *Biometrics* 2000; **56:** 345–51.

34 Chemick M. Bootstrapping methods: a practitioner's guide. New York, NY: Wiley Interscience, 1999.

35 Naucler P, Ryd W, Tornberg S, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *N Engl J Med* 2007; **357:** 1589–97.

36 Mayrand MH, Duarte-Franco E, Rodrigues I, et al. Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. *N Engl J Med* 2007; **357:** 1579–88.

37 Cuzick J, Clavel C, Petry KU, et al. Overview of the European and North American studies on HPV testing in primary cervical cancer screening. *Int J Cancer* 2006; **119:** 1095–101.

38 Naucler P, Ryd W, Tornberg S, et al. Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. *J Natl Cancer Inst* 2009; **101:** 88–99.

39 Castle PE, Fetterman B, Lorey T, Poitras N, Kinney W. Relationship of atypical glandular cell cytology, age, and human papillomavirus detection to cervical and endometrial cancer risks. *Obstet Gynecol* 2010; **115:** 243–48.

40 Cox JT. History of the use of HPV testing in cervical screening and in the management of abnormal cervical screening results. *J Clin Virol* 2009; **45** (suppl 1)**:** S3–12.

41 Castle PE, Sideri M, Jeronimo J, Solomon D, Schiffman M. Risk assessment to guide the prevention of cervical cancer. *Am J Obstet Gynecol* 2007; **197:** 356e1–6.

42 Katki HA, Wacholder S, Solomon D, Castle PE, Schiffman M. Risk estimation for the next generation of prevention programmes for cervical cancer. *Lancet Oncol* 2009; **10:** 1022–23.

43 Schiffman M, Wentzensen N, Wacholder S, Kinney W, Gage JC, Castle PE. Human papillomavirus testing in the prevention of cervical cancer. *J Natl Cancer Inst* 2011; **103:** 368–83.

44 Wright TC Jr, Schiffman M, Solomon D, et al. Interim guidance for the use of human papillomavirus DNA testing as an adjunct to cervical cytology for screening. *Obstet Gynecol* 2004; **103:** 304–09.

45 Sasieni P, Castle PE, Cuzick J. Further analysis of the ARTISTIC trial. *Lancet Oncol* 2009; **10:** 841–42.

46 Khan MJ, Castle PE, Lorincz AT, et al. The elevated 10-year risk of cervical precancer and cancer in women with human papillomavirus (HPV) type 16 or 18 and the possible utility of type-specific HPV testing in clinical practice. *J Natl Cancer Inst* 2005; **97:** 1072–79.

47 American Society for Colposcopy and Cervical Pathology. HPV genotyping clinical update 2009. http://www.asccp.org/ConsensusGuidelines/HPVGenotypingClinicalUpdate/tabid/5963/Default.aspx (accessed Aug 4, 2011).

48 Castle PE, Fetterman B, Thomas Cox J, et al. The age-specific relationships of abnormal cytology and human papillomavirus DNA results to the risk of cervical precancer and cancer. *Obstet Gynecol* 2010; **116:** 76–84.

49 Stoler MH, Schiffman M. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. *JAMA* 2001; **285:** 1500–05.

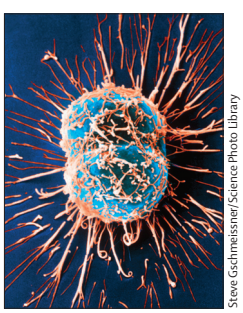
50 Castle PE, Cox JT, Jeronimo J, et al. An analysis of high-risk human papillomavirus DNA-negative cervical precancers in the ASCUS-LSIL Triage Study (ALTS). *Obstet Gynecol* 2008; **111:** 847–56.

51 Galgano MT, Castle PE, Atkins KA, Brix WK, Nassau SR, Stoler MH. Using biomarkers as objective standards in the diagnosis of cervical biopsies. *Am J Surg Pathol* 2010; **34:** 1077–87.

52 Bouvard V, Baan R, Straif K, et al, on behalf of the WHO International Agency for Research on Cancer Monograph Working Group. A review of human carcinogens—Part B: biological agents. *Lancet Oncol* 2009; **10:** 321–22.

53 Vijayaraghavan A, Efrusy MB, Goodman KA, Santas CC, Huh WK. Cost-effectiveness of using human papillomavirus 16/18 genotype triage in cervical cancer screening. *Gynecol Oncol* 2010; **119:** 237–42.

HPV16 and HPV18 genotyping in cervical cancer screening



Steve Gschmeissner/Science Photo Library

Published Online
August 23, 2011
DOI:10.1016/S1470-2045(11)70195-4

See Articles page 880

Cervical screening based on testing for high-risk human papillomavirus (HPV) is more effective than cytology-based primary cervical cancer screening,[1,2] but the best management strategy for women infected with high-risk HPV remains unclear. In *The Lancet Oncology*, Philip Castle and colleagues[3] report cross-sectional findings from a large study of the performance of genotyping for the two most dangerous HPV genotypes (HPV16 and HPV18) in the management of HPV-positive women. In women with high-risk HPV infection, HPV16 or HPV18 genotyping alone had sensitivity and positive predictive value (PPV) similar to abnormal cytology (atypical squamous cells of undetermined significance [ASC-US] or worse) for detection of cervical intraepithelial neoplasia grade 3 (CIN3) or worse. Additionally, the performance of HPV16 and HPV18 genotyping was much more consistent than cytology across assessment laboratories. Castle and colleagues also identified other potentially useful combinations of tests that could be used for triage. For example, to identify CIN3 or worse, detection of high-grade squamous intraepithelial lesion cytology or positive HPV16 or HPV18 was significantly more sensitivity and had a higher PPV than detection of ASC-US or worse.

In the USA, women infected with HPV who have positive cytology are immediately referred to colposcopy. The recommended policy for women infected with HPV who have normal cytology is for both tests to be repeated after 1 year. Colposcopy is then done if HPV infection is still present or cytology has become abnormal. The main drawback of this approach is the need to retest women after a fairly short interval, resulting in anxiety and loss to follow-up. Therefore, a strategy allowing immediate identification of all women with lesions needing treatment would be preferable. Unfortunately, all combinations of genotyping and cytology in Castle and colleagues' study had less than 80% sensitivity, leading the investigators to recommend test repetition after 1 year. Nevertheless, the increased sensitivity provided by the combined triage tests would allow some CIN3 or worse to be detected 1 year earlier, which should provide additional (though ill-defined) protection against cervical cancer. For example, immediate referral of women who were HPV16-positive, HPV18-positive, or had ASC-US or worse (instead of only those with ASC-US or worse), allowed detection of about 25% of HPV-positive CIN3 or worse a year earlier (corresponding with the difference in cross-sectional sensitivity between the two triage strategies). This increase in sensitivity was at the expense of a small decrease in cross-sectional PPV (from 14·1% to 12·6%). However, in a re-analysis of the Swedescreen study,[4] which additionally took into account the results of colposcopies from repeat testing in women not immediately referred, the loss in PPV between the same two triage strategies was larger (relative PPV for CIN3 or worse 0·61, 95% CI 0·41–0·89).

Strategies using other biomarkers to triage women with HPV infection are under assessment.[5,6] The cross-sectional sensitivity of immunochemistry for p16INK4a overexpression for CIN3 or worse is 91% (95% CI 77–97)

in women with HPV infection, suggesting that short-term retesting could be avoided in women who test negative for p16INK4a.[5] Women with HPV infection, however, are at strongly increased risk for development of new colposcopically detectable lesions, and premature reallocation of these women to screening intervals that are as long as those recommended in women without HPV infection might not be advisable, even after a negative colposcopy. Additional longitudinal data are needed to define the safest time interval before retesting in women with HPV infection who were negative for p16INK4a or any other triage test.

Follow-up data is expected from Castle and colleagues' study, but an important strength of their findings so far is the inclusion of scenarios that rely little, if at all, on cytology. Indeed, co-testing of HPV and cytology will probably be replaced by standalone HPV testing as primary screening test in high-income countries, because addition of cytology seems to provide little gain according to Castle and colleagues' findings (4·7% increase in sensitivity with 35·2% increase in screen positives) and the results of longitudinal studies.[2,7] The study by Castle and colleagues, although designed for high-income countries, can also provide useful information about triage strategies for countries where high-quality cytology has been difficult to implement and combinations of HPV tests might eventually offer a more sustainable option.

Guglielmo Ronco, Silvia Franceschi, Nereo Segnan
Centre of Cancer Prevention, Turin, Italy (GR, NS); and International Agency for Research on Cancer, Lyon, France (SF)
guglielmo.ronco@cpo.it

We declare no conflicts of interest.

1 Sankaranarayanan R, Nene BM, Shastri SS, et al. HPV screening for cervical cancer in rural India. *N Engl J Med* 2009; **360:** 1385–94.
2 Ronco G, Giorgi-Rossi P, Carozzi F, et al, and the New Technologies for Cervical Cancer screening (NTCC) Working Group. Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial. *Lancet Oncol* 2010; **11:** 249–57.
3 Castle PE, Stoler MH, Wright TC Jr, Sharma A, Wright TL, Behrens CM. Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical cancer screening of women aged 25 years and older: a subanalysis of the ATHENA study. *Lancet Oncol* 2011; published online Aug 23. DOI:10.1016/S1470-2045(11)70188-7.
4 Naucler P, Ryd W, Törnberg S, et al. Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. *J Natl Cancer Inst* 2009; **101:** 88–99.
5 Carozzi F, Confortini M, Dalla Palma P, et al, and the New Technologies for Cervical Cancer screening (NTCC) working group. Use of p16-INK4A overexpression to increase the specificity of human papillomavirus testing: a nested substudy of the NTCC randomised controlled trial. *Lancet Oncol* 2008; **9:** 937–45.
6 Cuschieri K, Wentzensen N. Human papillomavirus mRNA and p16 detection as biomarkers for the improved diagnosis of cervical neoplasia. *Cancer Epidemiol Biomarkers Prev* 2008; **17:** 2536–45.
7 Katki HA, Kinney WK, Fetterman B, et al. Cervical cancer risk for women undergoing concurrent testing for human papillomavirus and cervical cytology: a population-based study in routine clinical practice. *Lancet Oncol* 2011; **12:** 663–72.