Genotyping for HPV16 and HPV18 in women with minor cervical lesions: a systematic review and meta-analysis

Marc Arbyn, MD, MSc, DrTMH; Lan Xu, MSc; Freija Verduodt, PhD; Jack Cuzick, PhD; Anne Szarewski, MD, PhD; Jerome L. Belinson, MD; Nicolas Wentzensen, MD, PhD; Julia C. Gage, PhD, MPH; Michelle J. Khan, MD, MPH

MA, LX, FV: Unit of Cancer Epidemiology/Belgian Cancer Centre, Scientific Institute of Public Health, J Wytsmanstreet 14, B1050 Brussels, Belgium
JC, AS: Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, Charterhouse Square, EC1M 6BQ, UK
JLB: Preventive Oncology International Inc. and The Cleveland Clinic, Women’s Health Institute, 9500 Euclid Ave. A81 Cleveland, Ohio 44118, USA
NW, JCG: Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS, 609 Medical Center Drive, Bethesda, Massachusetts 20892, USA
MJK: Department of Obstetrics and Gynecology, Kaiser Permanente Northern California, 2500 Merced St., San Leandro, CA 94577, USA

*Corresponding author:
Marc Arbyn (MD, MSc, DrTMH), Coordinator of the Unit of Cancer Epidemiology, J. Wytsmanstreet 14, B1050 Brussels, Belgium; tel: +32 2 642 50 21; fax: +32 2 642 54 10; email: marc.arbyn@wiv-isp.be

Key words
Cervical cancer screening, human papillomavirus, HPV, HPV16, HPV18, genotyping, triage, atypical squamous cells of undetermined significance, low-grade squamous intraepithelial lesions, diagnostic test accuracy, meta-analysis

Word count: Abstract (277 words); Manuscript without abstract (2974 words)

Abstract

Background: High-risk (hr) HPV (human papillomavirus) testing to triage women with minor cervical lesions generates many referrals.

Purpose: To evaluate the accuracy and utility of HPV16/18 genotyping and the utility of HPV16/18 genotyping as a second triage step after prior triage with hrHPV in women with minor cervical lesions.

Sources: Searches in four bibliographic databases, without language restrictions, from 01/01/1999 to 1/02/2016.

Study Selection: Studies involving women with atypical squamous cells of undetermined significance (ASC-US) or low-grade squamous intra-epithelial lesions (LSIL) who were triaged with tests for hrHPV and HPV16/18 to find cervical intra-epithelial neoplasia, grade 2 or 3 or worse (CIN2+/CIN3+).

Data Extraction: Independent study selection, extraction of data and quality assessment by two reviewers.

Results: We found 24 studies of moderate to good quality involving 8,587 ASC-US and 5,284 LSIL cases. The pooled sensitivity of HPV16/18 genotyping for CIN3+ was
around 70% in both ASC-US or LSIL. The pooled specificity (threshold <CIN2) was
83% (95% CI 80 to 86%) in ASC-US and 76% (95% CI 74 to 79%) in LSIL.
The average risk for CIN3+ was 17% and 19% in HPV16/18 positive women with ASC-
US and LSIL, respectively. The average risk for CIN3+ was 5% in hrHPV+ but
HPV16/18- women with either ASC-US or LSIL.

Limitations: Methodological and technical heterogeneity among studies, insufficient
data to assess accuracy of separate assays.

Conclusion: HPV16/18 testing as a sole triage test for women with minor abnormal
cytology is poorly sensitive, but may be useful as second triage after hrHPV testing, with
direct referral if HPV16/18+. Whether colposcopy or repeat testing is recommended for
hrHPV+/HPV1618- women depends on local decision thresholds that can be derived
from pretest-posttest probability plots.

Primary Funding Source: 7th Framework Programme of the European Commission.

Introduction
Several countries are switching to HPV-based screening, however, cytological
examination of a Pap smear is still the main form of cervical cancer screening in many
places around the world. Direct referral for diagnostic work-up by colposcopy and biopsy
usually is recommended for women with high-grade lesions. However, women with
minor cytological abnormalities, including ASC-US (atypical squamous cells of
undetermined significance) or LSIL (low-grade squamous intraepithelial lesions), have
only a modestly increased risk for developing cervical cancer (1). In the past, repeating
the Pap test (repeat cytology) was the recommended follow-up for women with ASC-US
or LSIL. Given the strong etiological link between high-risk (hr) human papillomavirus
(HPV) infection and cervical cancer, hrHPV testing has been proposed as an alternative
triage method for women with equivocal or mildly abnormal cytology.

Randomised trials and systematic reviews show that hrHPV testing is more sensitive and
similarly specific compared to repeat cytology to identify underlying or incipient cervical
precancer in women with ASC-US, (2-4) Accordingly, hrHPV triage has become
standard practice (5-8). LSIL is associated with a risk of precancer similar to hrHPV+
ASC-US (9). Since the large majority of LSIL cases test positive for hrHPV (10), triage
by hrHPV testing is inefficient (11,12). The widespread practice of referring all women
with hrHPV+ ASC-US or LSIL to colposcopy carries a considerable burden and cost. As
HPV types 16 and 18 cause around 70% of cervical cancers (8), genotyping for these
types has been proposed as an additional tool allowing more fine-tuned management.

In this paper, we present the results of a systematic review on the accuracy of genotyping
for HPV16/18 to triage women with ASC-US/LSIL, or to triage women with ASC-
US/LSIL who are hrHPV positive. We also present a framework to assess the clinical
utility of triage tests, based on the risk of cervical precancer before and after triage.
Methods
We developed a protocol (see data supplement); followed standard procedures for meta-analyses of diagnostic accuracy studies (12,13); and reported processes and results according to standard guidelines (14).

Data Sources and Searches
We searched, using no language restrictions and the search strategies given in the protocol, PUBMED-Medline, EMBASE, Scopus and CENTRAL from January 1, 1999 to February 1, 2016 and also culled reference lists of selected reports.

Study Selection
Two reviewers independently screened titles and abstracts to identify relevant studies. Studies had to involve 20 or more women with either ASC-US or LSIL who had cervical samples tested with an assay detecting hrHPV as well as HPV16 and HPV18, and with a reference test to verify presence or absence of CIN (cervical intraepithelial neoplasia) 2+ and/or CIN3+.

The group of women with atypical squamous cells of undetermined significance comprised ASCUS/ASC-US (defined according to the 1988/2001 editions of the Bethesda System) (15,16) or borderline dyskaryosis (17). The group with low-grade squamous intra-epithelial lesions included LSIL (16) and mild dyskaryosis (17). Authors were contacted if no separate accuracy data were reported for ASC-US and LSIL or when only the outcome CIN3+ was reported.

Tests and Reference Standards
The evaluated index tests were assays identifying DNA or RNA of HPV16 and HPV18, jointly or separately (HPV16/18). A positive HPV16/18 test was considered positive if HPV16 or HPV18 were present and negative when both types were absent. The comparator tests were hrHPV assays identifying at least eight hrHPV types (HPV16/18/31/33/35/52/58). The HC2 assay was chosen as hrHPV comparator test if present. In studies where HC2 was not applied, other hrHPV assays or genotyping tests identifying separate hrHPV types were accepted as the hrHPV comparator test. Details regarding test platforms and the panel of considered hrHPV types were noted. The cut-off proposed by the manufacturer of each assay was accepted as the positivity criterion.

In addition to the single triage strategy with HPV16/18 genotyping, a combined triage strategy was assessed, where HPV16/18 genotyping as a second triage step was restricted to women who were hrHPV-positive at a first triage step. All women underwent verification with colposcopy, colposcopy-directed biopsies (possibly supplemented with random biopsies) or endocervical curettage. The type of verification (reference standard) was recorded for each study. Two levels of disease outcome were considered: CIN2+ and CIN3+. Adenocarcinoma in situ was included in the CIN3+ outcome.

Data Extraction and Quality Assessment
Two authors (MA, MJK) independently checked the eligibility of references and extracted the numbers of true positives, true negatives, false positives, and false negatives
for each test, triage group and outcome. Information on the study design, the clinical setting where patients were enrolled, the HPV assays, and the verification procedures was condensed in comprehensive tables. The quality of the selected studies was evaluated independently by two co-authors (LX, FV or MK) using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) check list (18,19).

Data Synthesis and Analysis
The absolute sensitivity and specificity were pooled using a bivariate normal model for the logit transforms of sensitivity and specificity (20,21). Summary receiver-operating-characteristics (sROC) and forest plots were drawn to show the joint overall and study-specific sensitivity and specificity of triage of ASC-US or LSIL using HPV16/18 genotyping.

The relative sensitivity and specificity of the index tests versus comparator tests were computed by including the test as a covariate in the bivariate model (22,23). We used the Linear Array assay (Roche Molecular Systems, Pleasanton, CA, USA) as comparator to assess the relative accuracy of HPV16/18 typing with different assays. Sources of heterogeneity in accuracy were assessed by including a series of potentially influential covariates in the bivariate model: QUADAS items, type of gold standard, HPV test platform used for hrHPV testing or HPV16/18 genotyping. The Deeks' regression test, based on the regression of the log diagnostic odds ratio onto 1/(effective sample size), was used to assess small study effects (publication bias) (24). Statistical significance was defined at the level p<0.05. However, for the assessment of the variation of accuracy over multiple categories, we applied a Bonferroni correction (0.05/k, k being the number of categories) to adjust the significance level. We conducted the statistical analyses with STATA version 13 (StataCorp, College Station, TX, USA) and SAS Enterprise, version 5.1 (SAS Institute Inc, Cary, NC, USA).

Pretest posttest probability plots were constructed to help evaluate utility of tests and testing strategies. Posttest risks were computed from the average prevalence of precancer in the reviewed studies. Decision thresholds were based on benchmark risk levels of 1% and 10%, applied in Europe, and 2.6% and 5.2%, applied in the US (25-27).

Funding Source
The review was funded by the Directorate-General for Research & Innovation of the European Commission (7th Framework Program, grant No. 603019). The funders had no role in the design, conduct, or reporting of the review.

Results
Selection of studies, study characteristics
From 899 references, 24 studies were selected that met inclusion criteria (see PRISMA flow-chart in Supplementary Figure 1). Studies involved 8,587 women with ASC-US and 5,284 with LSIL (28-51). An overview of study design, population, and test characteristics is provided in the Supplementary Tables 1-2. Additional data was obtained from the authors of most of the studies, with the exception of eight papers that contained...
all the absolute numbers required for computation of sensitivity and specificity (29,30,37,47-51). Study settings included colposcopy clinics (30,32,34-36,38,40-42,44-46,49-51), primary screening settings (28,29,39,43,47), a maternity center (31), and pathology archives (33,37). Fifteen different HPV assays were evaluated.

Quality of included studies
Supplementary Table 3 summarizes the methodological quality of the included studies. Most studies were of moderate or good quality; three were scored as probably free of bias (29,32,43), while one was scored of poor quality (37). The most common potential sources of bias were failure to report on uninterpretable/equivocal test results (n=12), failure to account for withdrawals (n=9), and an unclear delay between tests (n=8).

Absolute sensitivity and specificity
The sROC (Supplementary Figure 2) and forest plots (Figure 1) display the variation among studies as well as the pooled values of the sensitivity and specificity of genotyping for HPV16/18 to detect CIN2+ or CIN3+ in women with ASC-US and LSIL. HPV16/18 genotyping identified, on average, 70.7% (95% CI: 64.9-76.0%) of CIN3+ in women with ASC-US and 70.0% (95% CI: 65.4-74.2%) in women with LSIL (Table 1). The sensitivity of HPV16/18 for CIN2+ was lower (difference of 12 to 14%) compared to CIN3+. The pooled specificity to exclude CIN2+ was 82.9% (95% CI: 79.6-85.7%) in ASC-US and 76.3% (95% CI: 73.5-78.9%) in LSIL.

Table 2 shows how the accuracy of genotyping for HPV16/18 varied by test system using the Linear Array assay as the comparator. No statistically significant differences in sensitivity were observed. However, a higher specificity (p<0.0036) was noted for Easy Chip in triage of ASC-US and for Pretect HPV Proofer and MALDI-TOF (matrix-assisted laser desorption-ionization time-of-flight) and E6/E7 qPCR in triage of LSIL, whereas a lower specificity was noted for Clinical Arrays in triage of ASC-US.

Relative accuracy of genotyping for HPV16/18 compared to detection of hrHPV
The relative sensitivity of genotyping for HPV16/18 compared to hrHPV testing for detecting CIN3+ was 0.75 (95% CI: 0.68-0.83) in women with ASC-US and 0.70 (0.63-0.77) in women with LSIL (Table 3). The specificity of HPV16/18 to exclude CIN2+ was substantially higher than hrHPV testing: 1.70 (95% CI: 1.51-1.90) in ASC-US and 3.14 (95% CI: 2.83-3.48) in LSIL. Results for genotyping for HPV16 compared to hrHPV testing and compared to genotyping for HPV16/18 are provided in the supplement.

Influence of study and test characteristics, and small study effects
Few study quality items influenced the accuracy of HPV16/18 to detect underlying precancer (Supplementary Table 6). In women with ASC-US, the specificity of HPV16/18 was higher when withdrawal of cases was unclear, and in cases of partial verification. In women with LSIL, the sensitivity was higher and the specificity lower when an inappropriate reference test was used or when withdrawal of cases was not explained.
Genotyping for HPV16/18 in triage of ASC-US was less sensitive when the reference standard involved additional random biopsies or a mixture of gold standard tests, and was more specific when only one biopsy from the most suspect area was taken compared to when multiple colposcopy-targeted biopsies were taken. In general, the choice of the platform used for hrHPV testing (HC2 or other platform, panel of targeted hrHPV types, DNA or RNA testing) did not influence the relative accuracy of HPV16/18 genotyping compared to hrHPV testing. However, in triage of LSIL the magnitude of the relative specificity was higher with HC2 than with other hrHPV test platforms as comparator (Supplementary Table 8). The relative accuracy of an RNA-based assay targeting HPV16/18-45 was not different from an assay targeting HPV16/18 (Supplementary Table 9).

Deeks' regression test for funnel plot asymmetry did not reveal small study effects (Supplementary Table 10 and Supplementary Figure 6).

Meta-analysis of the accuracy of HPV16/18 genotyping in women with ASC-US or LSIL testing hrHPV positive

For women with ASC-US or LSIL who tested hrHPV positive in a first step triage, the pooled sensitivity was similar or slightly higher and the specificity was lower (Supplementary Table 11 and Figure 7) compared to all women with ASC-US or LSIL (table 1).

Pretest and posttest risk of cervical precancer

Table 4 displays accuracy measures and pre- and post-test probabilities of CIN2+ and CIN3+. Prior to triage testing, the average pre-test risk for CIN3+ is 6% among women with ASC-US. A positive hrHPV test raises the average risk to 10.1% and a negative hrHPV test decreases the risk to 0.5%. A positive HPV16/18 test in women with ASC-US raises the risk to 16.9% and a negative HPV16/18 test lowers it to 2.4%. In women with LSIL, the pretest probability of CIN3+ is 8.6%, whereas the post-test probabilities after triage are: 10.6% (if hrHPV+), 19.3% (if HPV16/18+), 1.0% (if hr HPV-) and 3.8% (if HPV16/18-). With the two-step triage, the risks of CIN3+ are 17.9% (if HPV16/18+) and 4.5% (if HPV16/18-) for women with ASC-US; and 18.0% (if HPV16/18+) and 5.1% (if HPV16/18-) for women with LSIL.

Pretest-posttest-probability plots

The utility of scenarios to triage women with LSIL is displayed in Figure 2, using European (at 1% and 10%) and US (at 2.6% and 5.2%) decision thresholds. A positive HPV16/18 result reclassifies 30% of patients with LSIL needing a colposcopy referral, whereas the other 70% need repeat testing (yellow zone in plot A). Plot 3B displays shifts in risk when hrHPV testing is followed by HPV16/18 genotyping. hrHPV testing minimally increases the risk of CIN3+ from 9% to 11% for hrHPV-positive LSIL but decreases the risk to 1% for hrHPV-negative LSIL. Adding HPV16/18 genotyping to hrHPV-positive LSIL women reclassifies 43% as needing immediate referral. Repeated surveillance testing is recommended for the remaining 57% who carry other hrHPV types.
In the US, the pretest risk of CIN3+ in women with LSIL is considered already sufficiently high to justify referral (Fig 3C). hrHPV-negative women could be followed conservatively (green zone in plots). HPV16/18-positive women need referral. Women who are hrHPV+ but HPV16/18- could be recommended to have either colposcopy or retesting, since their risk level is borderline (between the red and yellow risk zones). Additional plots, for diverse triage situations are shown in the Supplementary Figures 8-13.

Discussion

This meta-analysis found that genotyping for HPV16 and 18 detects approximately seven out of ten cases of CIN3+ and about six out of ten cases of CIN2+ in women with minor abnormal cytology. The pooled specificity (threshold < CIN2) is 83% in ASC-US and 76% in LSIL. HPV16/18 genotyping was substantially more specific but also less sensitive than testing for hrHPV. The average risk of underlying CIN3+ was 17% and 19% in HPV16/18-positive women with ASC-US or LSIL, respectively; 2% and 4% among women testing HPV16/18 negative, having ASC-US or LSIL, respectively; and 5% in women being hrHPV positive but HPV16/18 negative, having either ASC-US or LSIL.

These findings, in particular, the posttest risks of precancer, are useful for deciding how to incorporate HPV16/18 genotyping results into patient management. Up to recently, genotyping was done after hrHPV testing on hrHPV-positive women. Today, several genotyping platforms are available that allow inexpensive and high-throughput one-step genotyping. These platforms often give a readout of HPV16 and HPV18 genotyping separate from the other hrHPV testing, allowing the clinician immediate access to a secondary triage test. Our findings suggest that such partial genotyping tests can be used to risk stratify hrHPV-positive women to immediate colposcopy or to delayed follow-up. Local decision thresholds should help inform decisions about the clinical usefulness of the secondary triage strategy. In European guidelines, a risk for CIN3+ of >10% is considered the threshold for referring a woman to colposcopy. In the US, this decision threshold lies at >5.2%; an interval for surveillance testing 6-12 months later is proposed if the risk of CIN3+ is between 2.6-5.2%, and an interval for testing at longer intervals is proposed if the risk for CIN3+ is <2.6% (25,26). Women with ASC-US or LSIL who are hrHPV-positive but negative for HPV16/18 have a risk of underlying CIN3+ of around 5%. US guidelines would classify this risk as borderline. This means that both immediate referral to colposcopy or retesting would be plausible options without clear preference (Figure 2, plot C). The utility of genotyping is more obvious in a European setting, where delayed retesting could be proposed for women with minor cytological abnormalities who carry other high-risk types than HPV16/18.

Strengths of this meta-analysis include a large number of patients with ASC-US and LSIL from 24 studies enrolling more than 8,000 women with ASC-US and more than 5,000 women with LSIL. Our group has previously reviewed the utility of hrHPV testing in triage of borderline and low-grade cytologic abnormalities (3,4,12,52), but no group has previously performed a systematic review of HPV16/18 genotyping as a primary or
secondary triage test. This study helps clinicians to understand the underlying risks associated with HPV16/18 positivity which is now routinely reported in many of the newer HPV testing platforms.

In the background of our meta-analysis, we also assessed the accuracy of triage using genotyping for only HPV16, the most carcinogenic type. Genotyping of HPV16/18, was 8% more sensitive for CIN3+ in both ASC-US and LSIL, but 5% and 8% less specific, in ASC-US and LSIL, respectively, compared to genotyping for HPV16 (Appendix chapters 5 & 6).

Only eight of the twenty-four included studies contained all the required data in the published reports, but the required data from the other studies was obtained directly from the contacted authors. The studies were of moderate to good methodological quality giving us confidence in the reliability of our sensitivity and specificity estimates. There was no evidence of publication bias or small study effects. We found consistent and precise estimates of all relative accuracy estimates of HPV16/18 genotyping compared with hrHPV testing.

Our review identified several limitations in available data, including the lack of age-stratified data and too little data to precisely assess the accuracy of separate assays. Test accuracy studies are observational in design with short term outcomes that do not provide evidence on effectiveness with respect to prevention of cancer (53). Randomised trials assessing cumulative incidence of CIN3+ or cancer among triage-negative women are needed to provide high-quality evidence on the efficacy of alternative management options. Studies had methodological and technical heterogeneity although the influence of study quality and test characteristics on estimates of test accuracy appeared limited.

**Conclusion**

Triage of women with minor cytological abnormalities with partial genotyping identifying HPV16/HPV18 increases efficiency compared to hrHPV but at the expense of loss in sensitivity. Whether a triage test has good triaging capacity can be demonstrated by plotting risks of precancer on pre- and post-test probability plots. Women testing positive for HPV16/18 are at high risk and should be referred to colposcopy. Women carrying other hrHPV types but not HPV16/18 cannot be released to routine screening. Whether the risk is sufficiently low in these women to avoid referral to colposcopy or to propose repeat testing depends on local decision thresholds.
Acknowledgements

The following researchers are gratefully acknowledged for having contributed requested non-published data: M. Benevolo (Pathology Department, Regina Elena National Cancer Institute, Rome, Italy) (41); S. Andersson (Department of Women’s and Children’s Health, Division of Obstetrics and Gynecology, Karolinska Institutet, Stockholm, Sweden) (31); M. Guo (Department of Pathology, University of Texas MD Anderson Cancer Center, Houston, TX, USA) (33); P. Halfon (Virological Laboratory and Infectious Diseases, Laboratory Alphabio, Marseille, France) (34); L.W. Huang (Department of Obstetrics and Gynecology, Shin Kong Wu Ho-Su Memorial Hospital, Taipei, Taiwan) (54); S. Huang (Abbott Molecular Inc., Des Plaines, Illinois 60018-3315, USA) (35); F. Coutlée, Département de Microbiologie et Infectiologie and de Gynécologie-Obstétrique et Pathologie, Centre Hospitalier de l’Université de Montréal, Montréal, Québec, Canada) (36); J. Monsonego (Institute of the Cervix, Paris, France) (38); A. Pista (National Institute of Health, Lisbon, Portugal) (39); A. Spathis (Department of Cytopathology, University General Hospital ATTIKON, School of Medicine, National and Kapodistrian University of Athens, Chaidari, Greece) (42); M.H. Stoler (University of Virginia Health System, Charlottesville, USA) (43).

Additional non-published data were provided by the following co-authors. JP Belinson, Preventive Oncology International Inc. Cleveland, Ohio, USA (28); J. Gage (Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD) (32); J. Cuzick (Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, United Kingdom) (44); N. Wentzensen (Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS, Bethesda, MD, USA) (46).

All acknowledged authors provided written permissions verified by the Editor of Ann Intern Med.

Funding:

M. Arbyn, L.Xu and F. Verdoodt were supported by (i) the Belgian Foundation Against Cancer, Brussels, Belgium; (ii) the European Commission through the COHEAHR Network (grant No. 603019), coordinated by the Free University of Amsterdam (The Netherlands), funded by the 7th Framework program of DG Research (Brussels, Belgium); (iii) the Joint Action CANCON which has received funding from the European Union in the framework of the Health Programme (2008-13); and (iv) the Medizinische Hochschule of Hannover (Hannover, Germany). MJ Khan was supported by (i) training grant National Institutes of Health–National Institute of Allergy Infectious Diseases 5T32AI065388-05; (ii) institutional grant U1L RR024131, through the National Institutes of Health/National Center for Research Resources; (iii) University of California San Francisco–Clinical and Translational Science Institute Strategic Opportunities Support program; and (iii) mentored career development award K12 HS023009 through the US Agency for Healthcare Research and Quality.

Conflict of Interest:

MA, LX, FV: see funding.

AS, NW, JCG, MJK: no conflicts of interest declared.

JC: is an occasional consultant for Gen-Probe, and his institution has research funding from Gen-Probe.

The same applies to other HPV diagnostic companies: Abbott, BD, Qiagen and Roche.

JRB has received support (reagents, testing, funds or direct support and research) under the auspices of Preventive Oncology International Inc. from Hologic Inc., Qiagen, Gen-Probe Inc. and BGI Shenzhen.
Figure 1. Meta-analysis of the sensitivity (left) and specificity (right) of genotyping for HPV16/18 to detect CIN3+ in women with ASC-US.


* HPV16/18 genotyping with APTMA included also HPV45.
Figure 2. Pretest (left Y axis) and posttest probabilities (right Y axis) of CIN3+ after triage in women with LSIL using HPV16/18 genotyping as a single triage test (plot A) or using a two-step triage with hrHPV testing followed by HPV16/18 genotyping if hrHPV+ (plots B & C). Benchmarks are defined at risk levels 1% and 10%, often applied in Europe (Plots A & B), and at risk levels 2.6% and 5.2%, applied in the US (Plot C). CIN: cervical intra-epithelial neoplasia; HPV: human papillomavirus; hr: high-risk; LSIL: low-grade squamous intraepithelial lesion.
Table 1. Pooled absolute sensitivity and specificity of genotyping for HPV16/18 in triage of women with ASC-US or LSIL to detect underlying CIN2+ or CIN3+.

<table>
<thead>
<tr>
<th>Genotyping group</th>
<th>Outcome</th>
<th>No of studies/tests</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV16/18 ASC-US</td>
<td>CIN2+</td>
<td>32</td>
<td>58.8 (54.6-62.9)</td>
<td>82.9 (79.6-85.7)</td>
</tr>
<tr>
<td>HPV16/18 LSIL</td>
<td>CIN3+</td>
<td>25</td>
<td>70.7 (64.9-76.0)</td>
<td>78.1 (73.3-82.3)</td>
</tr>
<tr>
<td>HPV16/18 LSIL</td>
<td>CIN2+</td>
<td>28</td>
<td>55.5 (52.4-58.5)</td>
<td>76.3 (73.5-78.9)</td>
</tr>
</tbody>
</table>


Table 2. Variation in sensitivity and specificity of genotyping for HPV16/18 with a particular test system compared to genotyping for HPV16/18 with the Linear Array assay to detect CIN2+ in women with ASC-US (left) or LSIL (right).

<table>
<thead>
<tr>
<th>Test</th>
<th>No. of studies</th>
<th>Sensitivity</th>
<th>p</th>
<th>Specificity</th>
<th>p</th>
<th>Sensitivity</th>
<th>p</th>
<th>Specificity</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Triage of ASC-US</td>
<td></td>
<td>Triage of LSIL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear array†</td>
<td>10</td>
<td>0.85 (0.61-1.17)</td>
<td>0.3112</td>
<td>0.99 (0.93-1.05)</td>
<td>0.7274</td>
<td>0.88 (0.64-1.19)</td>
<td>0.4028</td>
<td>1.02 (0.97-1.07)</td>
<td>0.5137</td>
</tr>
<tr>
<td>Abbott RT PCR</td>
<td>5</td>
<td>0.96 (0.59-1.56)</td>
<td>0.8696</td>
<td>1.02 (0.85-1.22)</td>
<td>0.8461</td>
<td>1.03 (0.61-1.73)</td>
<td>0.9245</td>
<td>1.08 (0.92-1.27)</td>
<td>0.3546</td>
</tr>
<tr>
<td>PapilloCheck</td>
<td>1</td>
<td>0.96 (0.52-1.79)</td>
<td>0.8972</td>
<td>0.63 (0.50-0.81)</td>
<td>0.0002*</td>
<td>1.01 (0.69-1.48)</td>
<td>0.9639</td>
<td>0.82 (0.71-0.95)</td>
<td>0.0081</td>
</tr>
<tr>
<td>Clinical Arrays</td>
<td>1</td>
<td>1.14 (0.78-1.65)</td>
<td>0.4973</td>
<td>1.00 (0.90-1.12)</td>
<td>0.9498</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cervista</td>
<td>2</td>
<td>1.39 (0.72-2.69)</td>
<td>0.3303</td>
<td>1.12 (1.0-1.22)</td>
<td>0.00043</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>COBAS 4800</td>
<td>3</td>
<td>1.13 (0.92-1.39)</td>
<td>0.2325</td>
<td>1.03 (0.97-1.08)</td>
<td>0.3432</td>
<td>0.92 (0.74-1.14)</td>
<td>0.4552</td>
<td>1.01 (0.96-1.07)</td>
<td>0.6530</td>
</tr>
<tr>
<td>Easy Chip</td>
<td>1</td>
<td>0.99 (0.64-1.54)</td>
<td>0.3284</td>
<td>1.16 (1.08-1.25)</td>
<td>0.0001*</td>
<td>0.78 (0.40-1.55)</td>
<td>0.4812</td>
<td>1.17 (1.02-1.36)</td>
<td>0.0304</td>
</tr>
<tr>
<td>MALDI-TOF</td>
<td>2</td>
<td>0.77 (0.46-1.30)</td>
<td>0.3239</td>
<td>1.12 (1.0-1.21)</td>
<td>0.0033</td>
<td>0.88 (0.64-1.21)</td>
<td>0.4266</td>
<td>1.19 (1.09-1.29)</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>CLART</td>
<td>2</td>
<td>0.99 (0.64-1.54)</td>
<td>0.9608</td>
<td>0.99 (0.89-1.12)</td>
<td>0.9899</td>
<td>1.17 (0.89-1.54)</td>
<td>0.2635</td>
<td>0.87 (0.76-1.01)</td>
<td>0.0067</td>
</tr>
<tr>
<td>Es/E7 qPCR</td>
<td>1</td>
<td>0.96 (0.32-2.14)</td>
<td>0.3986</td>
<td>0.96 (0.69-1.34)</td>
<td>0.0001*</td>
<td>0.95 (0.77-1.18)</td>
<td>0.6644</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BD Viper</td>
<td>1</td>
<td>0.98 (0.15-6.32)</td>
<td>0.9847</td>
<td>0.97 (0.73-1.29)</td>
<td>0.8148</td>
<td>0.91 (0.59-1.42)</td>
<td>0.6932</td>
<td>1.03 (0.91-1.16)</td>
<td>0.6498</td>
</tr>
<tr>
<td>PapType</td>
<td>1</td>
<td>0.89 (0.13-6.30)</td>
<td>0.9055</td>
<td>0.96 (0.72-1.29)</td>
<td>0.7965</td>
<td>0.94 (0.61-1.45)</td>
<td>0.7743</td>
<td>0.93 (0.91-1.17)</td>
<td>0.6498</td>
</tr>
<tr>
<td>HPV Proofer</td>
<td>2</td>
<td>0.96 (0.63-1.47)</td>
<td>0.8542</td>
<td>1.05 (0.9-1.13)</td>
<td>0.2853</td>
<td>0.86 (0.61-1.21)</td>
<td>0.3585</td>
<td>1.10 (1.03-1.18)</td>
<td>0.0043*</td>
</tr>
<tr>
<td>APTIMA</td>
<td>1</td>
<td>0.85 (0.11-6.60)</td>
<td>0.8750</td>
<td>1.00 (0.77-1.31)</td>
<td>0.9758</td>
<td>0.91 (0.57-1.45)</td>
<td>0.6798</td>
<td>1.00 (0.87-1.14)</td>
<td>0.9437</td>
</tr>
</tbody>
</table>

† comparator test; *significant likelihood ratio test which assess whether the relative accuracy is statistically different from unity with significance level defined at 0.05/k (k=14, being the number of compared assays, Bonferroni correction for multiple comparisons).

Table 3. Meta-analysis of the relative sensitivity and relative specificity of genotyping for HPV16/18 compared to testing for high-risk HPV.

<table>
<thead>
<tr>
<th>Triage Group</th>
<th>Outcome</th>
<th>Number of comparisons</th>
<th>Relative sensitivity</th>
<th>Relative specificity</th>
<th>P</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASC-US</td>
<td>CIN2+</td>
<td>29</td>
<td>0.59 (0.54-0.65)</td>
<td>1.70 (1.51-1.90)</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>CIN3+</td>
<td>15</td>
<td>0.75 (0.68-0.83)</td>
<td>1.87 (1.64-2.12)</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>LSIL</td>
<td>CIN2+</td>
<td>19</td>
<td>0.56 (0.51-0.62)</td>
<td>3.14 (2.83-3.48)</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>CIN3+</td>
<td>15</td>
<td>0.70 (0.63-0.77)</td>
<td>3.49 (3.01-4.05)</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>


Table 4. Accuracy (sensitivity, specificity, likelihood ratios, test positivity rate), pre-test and post-test probabilities of CIN2+ and CIN3+ of triage with hrHPV testing or HPV16/18 genotyping among women with ASC-US or LSIL (white rows). Data for triage with HPV16/18 genotyping among women hrHPV-positive ASC-US or hrHPV-positive LSIL are shown in the grey rows (two-step triage).

<table>
<thead>
<tr>
<th>Triage Group</th>
<th>Test</th>
<th>Outcome</th>
<th>Pre-test risk*</th>
<th>Pooled sensitivity</th>
<th>Pooled specificity</th>
<th>PLR</th>
<th>NLR</th>
<th>test+</th>
<th>Post-test risk if test+</th>
<th>Post-test risk if test-</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASC-US</td>
<td>hrHPV</td>
<td>CIN2+</td>
<td>10.1%</td>
<td>95.0%</td>
<td>48.6%</td>
<td>1.85</td>
<td>0.10</td>
<td>55.8%</td>
<td>17.2%</td>
<td>1.1%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CIN3+</td>
<td>6.0%</td>
<td>96.6%</td>
<td>45.0%</td>
<td>1.76</td>
<td>0.08</td>
<td>57.5%</td>
<td>10.1%</td>
<td>0.5%</td>
</tr>
<tr>
<td>HPV16/18</td>
<td></td>
<td>CIN2+</td>
<td>10.1%</td>
<td>58.8%</td>
<td>82.9%</td>
<td>3.43</td>
<td>0.50</td>
<td>21.3%</td>
<td>27.7%</td>
<td>5.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CIN3+</td>
<td>6.0%</td>
<td>70.7%</td>
<td>78.1%</td>
<td>3.23</td>
<td>0.38</td>
<td>24.8%</td>
<td>16.9%</td>
<td>2.4%</td>
</tr>
<tr>
<td>ASC-US</td>
<td>HPV16/18</td>
<td>CIN2+</td>
<td>17.2%</td>
<td>60.1%</td>
<td>67.3%</td>
<td>1.84</td>
<td>0.59</td>
<td>37.4%</td>
<td>27.5%</td>
<td>11.0%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CIN3+</td>
<td>10.1%</td>
<td>73.8%</td>
<td>61.7%</td>
<td>1.93</td>
<td>0.42</td>
<td>41.9%</td>
<td>17.9%</td>
<td>4.5%</td>
</tr>
<tr>
<td>LSIL</td>
<td>hrHPV</td>
<td>CIN2+</td>
<td>21.1%</td>
<td>96.9%</td>
<td>24.8%</td>
<td>1.29</td>
<td>0.12</td>
<td>79.7%</td>
<td>25.6%</td>
<td>3.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CIN3+</td>
<td>8.6%</td>
<td>97.7%</td>
<td>22.4%</td>
<td>1.26</td>
<td>0.10</td>
<td>79.3%</td>
<td>10.6%</td>
<td>1.0%</td>
</tr>
<tr>
<td>HPV16/18</td>
<td></td>
<td>CIN2+</td>
<td>21.1%</td>
<td>55.5%</td>
<td>76.3%</td>
<td>2.34</td>
<td>0.58</td>
<td>30.4%</td>
<td>38.5%</td>
<td>13.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CIN3+</td>
<td>8.6%</td>
<td>70.0%</td>
<td>72.5%</td>
<td>2.55</td>
<td>0.41</td>
<td>31.1%</td>
<td>19.3%</td>
<td>3.8%</td>
</tr>
<tr>
<td>LSIL</td>
<td>HPV16/18</td>
<td>CIN2+</td>
<td>25.6%</td>
<td>60.2%</td>
<td>64.6%</td>
<td>1.70</td>
<td>0.62</td>
<td>41.7%</td>
<td>36.9%</td>
<td>17.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CIN3+</td>
<td>10.6%</td>
<td>72.6%</td>
<td>60.7%</td>
<td>1.85</td>
<td>0.45</td>
<td>42.8%</td>
<td>18.0%</td>
<td>5.1%</td>
</tr>
</tbody>
</table>

ASC-US: atypical squamous cells of undermined significance; CIN: cervical intra-epithelial neoplasia; HPV: human papillomavirus; hr: high-risk; LSIL: low-grade squamous intraepithelial lesion; PLR: positive likelihood ratio; NLR: negative likelihood ratio; PPV: positive predictive value; cNPV: complement of the negative predictive value (cNPV=1-NPV).

*Pretest risk: average of the prevalence or short term cumulative incidence of CIN2+ or CIN3+ pooled from the studies included in the meta-analysis. For 2-step triage, the pre-test risk corresponds with the post-test risk after hrHPV testing.

Reproducible Research Statement
Study Protocol: see Data Supplement
Statistical Code: see Methods and Study Protocol in Data Supplement
Data Set: available from the first author upon request.

HPV1618triage_27clean.docx
Reference List


HPV1618triage_27cleanHPV1618triage_27clean.docx


