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 Verdoodt, 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.

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

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 (HPV16/18/31/33/35/45/52/58). 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

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

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~~ 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 Bonferronni 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

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

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 UL1 RR024131, through the National Institutes of Health/National Center for Research Resources 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.

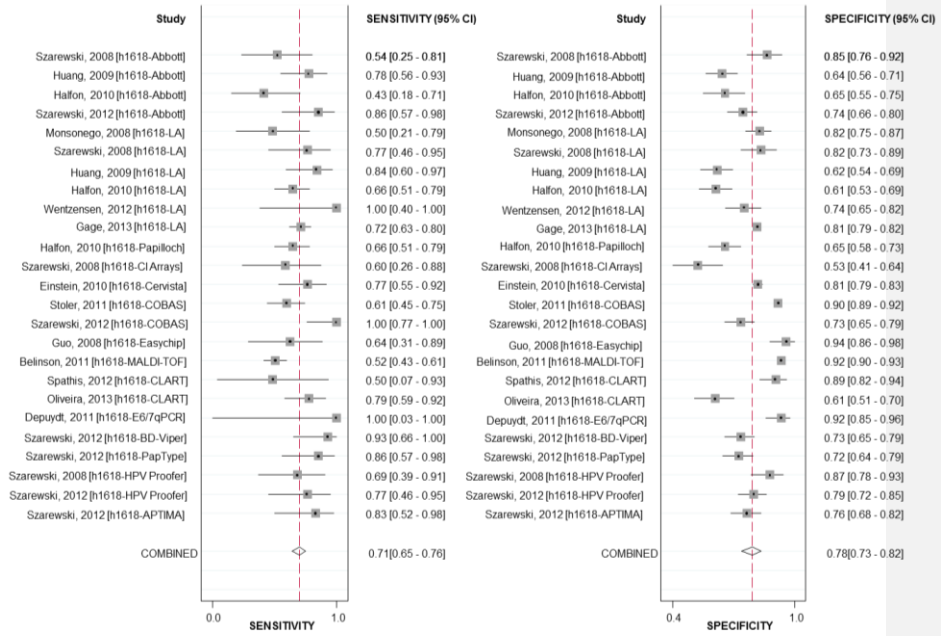

Figure 1. Meta-analysis of the sensitivity (left) and specificity (right) of genotyping for HPV16/18 to detect CIN3+ in women with ASC-US.
ASC-US: atypical squamous cells of undetermined significance; CIN: cervical intraepithelial lesion.
* HPV16/18 genotyping with APTMA included also HPV45.

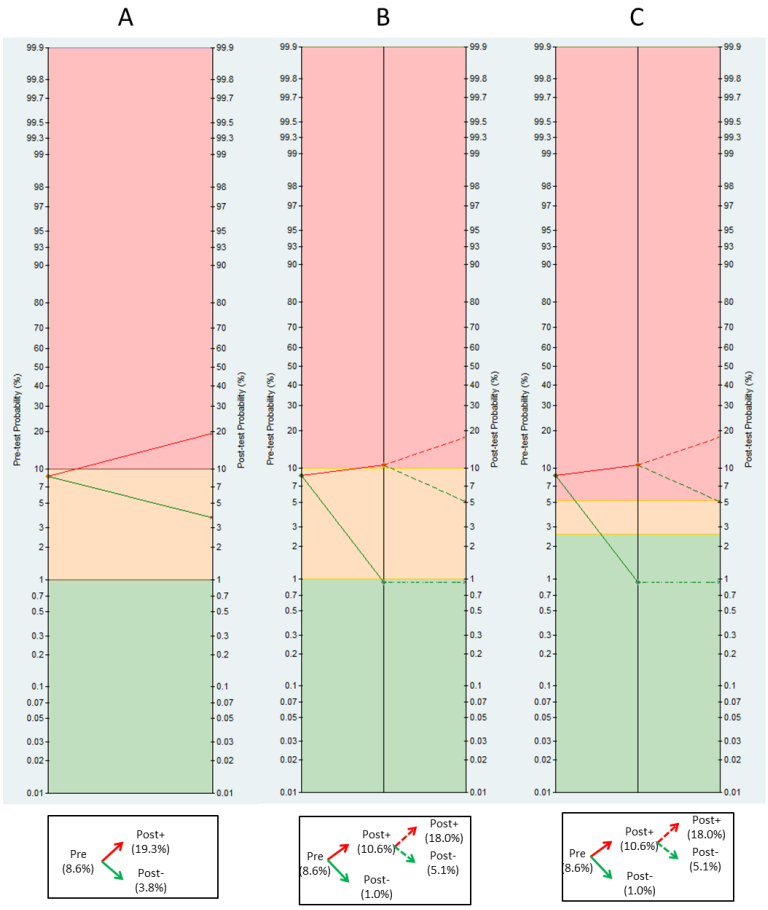

Figure 2. Pretest (left Y axis) and posttest probabilities (right Y axis) of CIN3+ after triage in women with LSIL using HPV16/18 genotyping as a single triage test (plot A) or using a two-step triage with hrHPV testing followed by HPV16/18 genotyping if hrHPV+ (plots B & C). Benchmarks are defined at 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 women with ASC-US or LSIL to detect underlying CIN2+ or CIN3+.

Genotyping	Triage group	Outcome	No of studies/tests	Pooled value, in % Sensitivity (95% CI)	Pooled value, in % Specificity (95% CI)
HPV16/18	ASC-US	CIN2+	32	58.8 (54.6-62.9)	82.9 (79.6-85.7)
		CIN3+	25	70.7 (64.9-76.0)	78.1 (73.3-82.3)
	LSIL	CIN2+	28	55.5 (52.4-58.5)	76.3 (73.5-78.9)
		CIN3+	24	70.0 (65.4-74.2)	72.5 (69.0-75.8)

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.

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).

Test	No. of studies	Triage of ASC-US Sensitivity ratio	p	Triage of ASC-US Specificity ratio	p	Triage of LSIL Sensitivity ratio	p	Triage of LSIL Specificity ratio	p
Linear array†	10	1.00		1.00		1.00		1.00	
Abbott RT PCR	5	0.85 (0.61-1.17)	0.3112	0.99 (0.93-1.05)	0.7274	0.88 (0.64-1.19)	0.4028	1.02 (0.97-1.07)	0.5137
PapilloCheck	1	0.967 (0.59-1.56)	0.8696	1.02 (0.85-1.22)	0.8461	1.03 (0.61-1.73)	0.9245	1.08 (0.92-1.27)	0.3546
Clinical Arrays	1	0.96 (0.52-1.79)	0.8972	0.63 (0.50-0.81)	0.0002*	1.01 (0.69-1.48)	0.9639	0.82 (0.71-0.95)	0.0081
Cervista	2	1.14 (0.78-1.65)	0.4973	1.00 (0.90-1.12)	0.9498	-		-	
COBAS 4800	3	1.13 (0.92-1.39)	0.2325	1.03 (0.97-1.08)	0.3432	0.92 (0.74-1.14)	0.4552	1.01 (0.96-1.07)	0.6530
Easy Chip	1	1.09 (0.64-1.86)	0.7414	1.16 (1.08-1.25)	0.0001*	0.78 (0.40-1.55)	0.4812	1.17 (1.02-1.36)	0.0304
MALDI-TOF	1	0.77 (0.46-1.30)	0.3329	1.12 (1.0-1.21)	0.0043	0.88 (0.64-1.21)	0.4266	1.19 (1.09-1.29)	<0.0001*
CLART	2	0.99 (0.64-1.54)	0.9608	0.99 (0.89-1.12)	0.9898	1.17 (0.89-1.54)	0.2655	0.87 (0.76-1.01)	0.0607
E6/E7 qPCR	1	1.39 (0.72-2.69)	0.3303	1.12 (1.0-1.22)	0.0043	0.60 (0.17-2.17)	0.4382	0.95 (0.77-1.18)	0.6644
BD Viper	1	0.98 (0.15-6.32)	0.9847	0.97 (0.73-1.29)	0.8148	0.91 (0.59-1.42)	0.6932	1.03 (0.91-1.16)	0.6498
PapType	1	0.89 (0.13-6.30)	0.9055	0.96 (0.72-1.29)	0.7965	0.94 (0.61-1.45)	0.7743	1.03 (0.91-1.17)	0.6498
HPV Proofer	2	0.96 (0.63-1.47)	0.8542	1.05 (0.9-1.13)	0.2679	0.86 (0.61-1.21)	0.3858	1.10 (1.03-1.18)	0.0031*
APTIMA	1	0.85 (0.11-6.60)	0.8750	1.00 (0.77-1.31)	0.9758	0.91 (0.57-1.45)	0.6798	1.00 (0.87-1.14)	0.9437

† 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, Bonferronni correction for multiple comparisons).

ASC-US: atypical squamous cells of undermined significance; CIN: cervical intra-epithelial neoplasia; HPV: human papillomavirus; LSIL: low-grade squamous intraepithelial lesion.

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

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

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.

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).

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

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.

Reference List

(1) **Holowaty P, Miller AB, Rohan T, To T**. Natural History of Dysplasia of the Uterine Cervix. J Natl Cancer Inst 1999; 91:252-258.

(2) **ASCUS-LSIL Triage Study Group**. Results of a randomized trial on the management of cytology interpretations of atypical squamous cells of undetermined significance. Am J Obstet Gynecol 2003; 188:1383-1392.

(3) **Arbyn M, Buntinx F, Van Ranst M, Paraskevaidis E, Martin-Hirsch P, Dillner J**. Virologic versus cytologic triage of women with equivocal Pap smears: a meta-analysis of the accuracy to detect high-grade intraepithelial neoplasia. J Natl Cancer Inst 2004; 96:280-293.

(4) **Arbyn M, Ronco G, Anttila A, Meijer CJLM, Poljak M, Ogilvie G et al.** Evidence regarding HPV testing in secondary prevention of cervical cancer. Vaccine 2012; 30 Suppl 5:F88-F99.

(5) **Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J et al.** American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. CA Cancer J Clin 2012; 62:147-172.

(6) **Provencher DM, Murphy KJ**. Canadian Consensus Guidelines on Human Papillomavirus: Chapter 3: The role of HPV testing. J Obstet Gynaecol Can 2008; 8:15-23.

(7) **MSAC**. Human papillomavirus triage test for women with possible or definite low-grade squamous intraepithelial lesions. MSAC reference 39, 1-200. 2009. Canberra, Medical Services Advisory Committee, Department of Health and Ageing.

(8) **Arbyn M, Anttila A, Jordan J, Ronco G, Schenck U, Segnan N et al.** European Guidelines for Quality Assurance in Cervical Cancer Screening. Second Edition - Summary Document. Ann Oncol 2010; 21:448-458.

(9) **ASCUS-LSIL Triage Study Group**. A randomized trial on the management of low-grade squamous intraepithelial lesion cytology interpretations. Am J Obstet Gynecol 2003; 188:1393-1400.

(10) **Zuna RE, Wang SS, Rosenthal DL, Jeronimo J, Schiffman M, Solomon D**. Determinants of human papillomavirus-negative, low-grade squamous intraepithelial lesions in the atypical squamous cells of undetermined significance/low-grade squamous intraepithelial lesions triage study (ALTS). Cancer 2005; 105:253-262.

(11) **Arbyn M, Martin-Hirsch P, Buntinx F, Van Ranst M, Paraskevaidis E, Dillner J**. Triage of women with equivocal or low-grade cervical cytology results. A meta-analysis of the HPV test positivity rate. J Cell Mol Med 2009; 13:648-659.

(12) **Arbyn M, Roelens J, Simoens C, Buntinx F, Paraskevaidis E, Martin-Hirsch PPL et al.** Human papillomavirus testing versus repeat cytology for triage of minor cytological cervical lesions. Cochrane Database Syst Rev 2013; 3:1-201.

(13) **Leeflang MM, Deeks JJ, Takwoingi Y, Macaskill P**. Cochrane diagnostic test accuracy reviews. Syst Rev 2013; 2:82.

(14) **Moher D, Liberati A, Tetzlaff J, Altman DG**. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009; 151:264-9, W64.

(15) **Lundberg GD, National Cancer Institute**. The 1988 Bethesda System for Reporting Cervical/Vaginal Cytologic Diagnoses. JAMA 1989; 262:931-934.

(16) **Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M et al.** The 2001 Bethesda System: terminology for reporting results of cervical cytology. JAMA 2002; 287:2114-2119.

(17) **Evans DMD, Hudson EA, Brown CL, Boddington MM, Hughes HE, Mackenzie EFD et al.** Terminology in gynaecological cytopathology : report of the working party of the BSCC. J Clin Pathol 1986; 39:933-944.

(18) **Whiting P, Rutjes AWS, Reitsma JB, Bossuyt PM, Kleijnen J**. The development of QUADAS : a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 2003; 3:1-13.

(19) **Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB et al.** QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; 155:529-536.

(20) **Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA**. A unification of models for meta-analysis of diagnostic accuracy studies. Biostatistics 2007; 8:239-251.

(21) **Harbord RM, Whiting P**. metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. The Stata Journal 2009; 9:211-229.

(22) **Deeks JJ, Bossuyt PM, Gatsonis C**. **Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0** . The Cochrane Collaboration; 2013.

(23) **Takwoingi Y, Leeflang MM, Deeks JJ**. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. Ann Intern Med 2013; 158:544-554.

(24) **Deeks JJ, Macaskill P, Irwig L**. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol 2005; 58:882-893.

(25) **Katki HA, Schiffman M, Castle PE, Fetterman B, Poitras NE, Lorey T et al.** Benchmarking CIN 3+ Risk as the Basis for Incorporating HPV and Pap Cotesting into Cervical Screening and Management Guidelines. J Low Genit Tract Dis 2013; 17:S28-S35.

(26) **Wentzensen N, Fetterman B, Castle PE, Schiffman M, Wood SN, Stiemerling E et al.** p16/Ki-67 Dual Stain Cytology for Detection of Cervical Precancer in HPV-Positive Women. J Natl Cancer Inst 2015; 107.

(27) **Arbyn M, Roelens J, Martin-Hirsch P, Leeson S, Wentzensen N**. Use of HC2 to triage women with borderline and mild dyskaryosis in the UK. Br J Cancer 2011; 105:877-880.

(28) **Belinson JL, Wu R, Belinson SE, Qu X, Yang B, Du H et al.** A population-based clinical trial comparing endocervical high-risk HPV testing using hybrid capture 2 and Cervista from the SHENCCAST II Study. Am J Clin Pathol 2011; 135:790-795.

(29) **Cuzick J, Thomas CJ, Zhang G, Einstein MH, Stoler M, Trupin S et al.** Human papillomavirus testing for triage of women with low-grade squamous intraepithelial lesions. Int J Cancer 2013; 132:959-966.

(30) **Einstein MH, Martens MG, Garcia FAR, Ferris DG, Mitchell AL, Day SP et al.** Clinical validation of the Cervista« HPV HR and 16/18 genotyping tests for use in women with ASC-US cytology. Gynecol Oncol 2010; 118:116-122.

(31) **Fröberg M, Johansson B, Hjerpe A, Andersson S**. Human papillomavirus "reflex" testing as a screening method in cases of minor cytological abnormalities. Br J Cancer 2008;1-6.

(32) **Gage JC, Schiffman M, Solomon D, Wheeler CM, Gravitt PE, Castle PE et al.** Risk of precancer determined by HPV genotype combinations in women with minor cytologic abnormalities. Cancer Epidemiol Biomarkers Prev 2013;1095-1101.

(33) **Guo M, Lin CY, Gong Y, Cogdell DE, Zhang W, Lin E et al.** Human papillomavirus genotyping for the eight oncogenic types can improve specificity of HPV testing in women with mildly abnormal Pap results. Mod Pathol 2008; 21:1037-1043.

(34) **Halfon P, Benmoura D, Agostini A, Khiri H, Penaranda G, Martineau A et al.** Evaluation of the clinical performance of the Abbott RealTime High-Risk HPV for carcinogenic HPV detection. J Clin Virol 2010; 48:246-250.

(35) **Huang S, Erickson B, Tang N, Mak WB, Salituro J, Robinson J et al.** Clinical performance of Abbott RealTime High Risk HPV test for detection of high-grade cervical intraepithelial neoplasia in women with abnormal cytology. J Clin Virol 2009; 45 Suppl 1:S19-S23.

(36) **Lapierre SG, Sauthier P, Mayrand MH, Dufresne S, Petignat P, Provencher D et al.** Human Papillomavirus (HPV) DNA Triage of Women with Atypical Squamous Cells of Undetermined Significance with cobas 4800 HPV and Hybrid Capture 2 Tests for Detection of High-Grade Lesions of the Uterine Cervix. J Clin Microbiol 2012; 50:1240-1244.

(37) **Wong AK, Chan RC, Nichols WS, Bose S**. Invader human papillomavirus (HPV) type 16 and 18 assays as adjuncts to HPV screening of cervical papanicolaou smears with atypical squamous cells of undetermined significance. Cancer 2009; 115:823-832.

(38) **Monsonego J, Pollini G, Evrard MJ, Sednaoui P, Monfort L, Quinzat D et al.** Linear array genotyping and hybrid capture II assay in detecting human papillomavirus genotypes in women referred for colposcopy due to abnormal Papanicolaou smear. Int J STD AIDS 2008; 19:385-392.

(39) **Oliveira A, Verdasca N, Pista A**. Use of the NucliSENS EasyQHPV assay in the management of cervical intraepithelial neoplasia. J Med Virol 2013; 85:1235-1241.

(40) **Depuydt CE, Makar AP, Ruymbeke MJ, Benoy IH, Vereecken AJ, Bogers JJ**. BD-ProExC as Adjunct Molecular Marker for Improved Detection of CIN2+ after HPV Primary Screening. Cancer Epidemiol Biomarkers Prev 2011; 20:628-637.

(41) **Dona MG, Vocaturo A, Giuliani M, Ronchetti L, Rollo F, Pescarmona E et al.** p16/Ki-67 dual staining in cervico-vaginal cytology: Correlation with histology, Human Papillomavirus detection and genotyping in women undergoing colposcopy. Gynecol Oncol 2012; 126:198-202.

(42) **Spathis A, Kottaridi C, Chranioti A, Meristoudis C, Chrelias C, Panayiotides IG et al.** mRNA and DNA detection of human papillomaviruses in women of all ages attending two colposcopy clinics. PLoS ONE 2012; 7:e49205-0.

(43) **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; 135:468-475.

(44) **Szarewski A, Ambroisine L, Cadman L, Austin J, Ho L, Terry G et al.** Comparison of Predictors for High-Grade Cervical Intraepithelial Neoplasia in Women with Abnormal Smears. Cancer Epidemiol Biomarkers Prev 2008; 17:3033-3043.

(45) **Szarewski A, Mesher D, Cadman L, Austin J, Ashdown-Barr L, Ho L et al.** A comparison of seven tests for high grade cervical intraepithelial neoplasia in women with abnormal smears: the Predictors 2 study. J Clin Microbiol 2012; 50:1867-1873.

(46) **Wentzensen N, Schwartz L, Zuna RE, Smith KM, Mathews C, Gold M et al.** Performance of p16/Ki-67 immunostaining to detect cervical cancer precursors in a colposcopy referral population. Clin Cancer Res 2012; 18:4154-4162.

(47) **Wong OG, Lo CK, Szeto E, Cheung AN**. Efficacy of Abbott RealTime High Risk HPV test in evaluation of atypical squamous cells of undetermined significance from an Asian screening population. J Clin Virol 2011; 51:136-138.

(48) **Lin CQ, Cui JF, Zhang X, Pan QJ, Chen W, Qiao YL**. Human Papillomavirus Genotyping to Predict the Risk of Cervical Precancerous Lesions or Cancer in Women with Minor Abnormal Cytology in China. Acta Cytol 2015; 59:405-411.

(49) **Iwata T, Hasegawa T, Ochiai K, Takizawa K, Umezawa S, Kuramoto H et al.** Human Papillomavirus Test for Triage of Japanese Women With Low-Grade Squamous Intraepithelial Lesions. Reprod Sci 2015; 22:1509-1515.

(50) **Nakamura Y, Matsumoto K, Satoh T, Nishide K, Nozue A, Shimabukuro K et al.** HPV genotyping for triage of women with abnormal cervical cancer screening results: a multicenter prospective study. Int J Clin Oncol 2015; 20:974-981.

(51) **Khunamornpong S, Settakorn J, Sukpan K, Srisomboon J, Intaraphet S, Siriaunkgul S**. High performance of combined HPV testing and genotyping for HPV16/18/52/58 in triaging women with minor cervical cytological abnormalities in northern Thailand. J Med Virol 2016; 88:135-143.

(52) **Arbyn M, Sasieni P, Meijer CJ, Clavel C, Koliopoulos G, Dillner J**. Chapter 9: Clinical applications of HPV testing: a summary of meta-analyses. Vaccine 2006; 24 (SUPPL. 3):S78-S89.

(53) **Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE et al.** GRADE: Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ 2008; 336:1106-1110.

(54) **Huang LW, Lin YH, Pan HS, Seow KM, Lin CY**. Human papillomavirus genotyping as a predictor of high-grade cervical dysplasia in women with mildly cytologic abnormalities: A two-year follow-up report. Diagn Cytopathol 2012; 40:673-677.