The APTIMA HPV assay *versus* the Hybrid Capture 2 test in triage of women with ASC-US or LSIL cervical cytology: A meta-analysis of the diagnostic accuracy

Marc Arbyn¹, Jolien Roelens¹, Kate Cuschieri², Jack Cuzick³, Ann Szarewski³, Sam Ratnam⁴, Miriam Reuschenbach⁵, Suzanne Belinson⁶, Jerome L. Belinson⁶ and Joseph Monsonego⁷

¹ Unit of Cancer Epidemiology, Scientific Institute of Public Health, Brussels, Belgium
² Scottish HPV Reference Laboratory, Royal Infirmary of Edinburgh, Edinburgh, Scotland, United Kingdom
³ Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, United Kingdom
⁴ Public Health Laboratory, Department of Health, Division of Community Medicine, Faculty of Medicine, Memorial University of Newfoundland, St. John’s, Newfoundland, Canada
⁵ Department of Applied Tumor Biology, Institute of Pathology, University of Heidelberg, Heidelberg, Germany
⁶ Preventive Oncology International, Inc., Cleveland, OH, USA
⁷ Institute of the Cervix, Paris, France

Testing for DNA of 13 high-risk HPV types with the Hybrid Capture 2 (HC2) test has consistently been shown to perform better in triage of women with cervical cytology results showing atypical squamous cells of undetermined significance (ASC-US) but often not in triage of low-grade squamous intraepithelial lesions (LSIL) detected in cervical cancer screening. In a meta-analysis, we compared the accuracy of the APTIMA HPV test, which identifies RNA of 14 high-risk HPV types, to HC2 for the triage of women with ASC-US or LSIL. Literature search-targeted studies where the accuracy of APTIMA HPV and HC2 for detection of underlying CIN2/3+ was assessed concomitantly including verification of all cases of ASC-US and LSIL. HSROC (Hierarchical Summary ROC) curve regression was used to compute the pooled absolute and relative sensitivity and specificity. Eight studies, comprising 1,839 ASC-US and 1,887 LSIL cases, were retrieved. The pooled sensitivity and specificity of APTIMA to triage ASC-US to detect underlying CIN3 or worse was 96.2% (95% CI = 91.7–98.3%) and 54.9%

**Key words:** cervical cancer screening, HPV, hybrid capture, APTIMA, triage, atypical squamous cells of undetermined significance, low-grade squamous intraepithelial lesions, diagnostic test accuracy, meta-analysis

**Grant sponsors:** European Commission, European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines (ECCG), Belgian Foundation Against Cancer (Brussels, Belgium), Gynaecological Cancer Cochrane Review Collaboration (Bath, United Kingdom)

**DOI:** 10.1002/ijc.27636

**History:** Received 17 Feb 2012; Accepted 8 May 2012; Online 18 May 2012

**Correspondence to:** Marc Arbyn, Unit of Cancer Epidemiology, J. Wytsmanstreet 14, B1050 Brussels, Belgium, Tel.: +32-2-642-50-21, Fax: +32-2-642-54-10; E-mail: marc.arbyn@wiv-isp.be

Conflict of Interest M. Arbyn and J. Roelens were supported by (i) The Belgian Foundation Against Cancer, Brussels, Belgium; (ii) the European Commission through the PREHDICT Network (grant No. 242061), coordinated by the Free University of Amsterdam (The Netherlands), funded by the 7th Framework program of DG Research (Brussels, Belgium) and through the ECCG (grant no. 2006322), European Cooperation on development and implementation of Cancer screening and prevention Guidelines, via IARC, Lyon, France), funded by Directorate of SANCO (Luxembourg, Grand-Duchy of Luxembourg) and (iii) the Gynaecological Cancer Cochrane Review Collaboration (Bath, United Kingdom).

K. Cuschieri has been on the advisory board of Gen-Probe and has received project funding from Gen-Probe. The same applies to other companies including Qiagen, Roche and Abbott.

J. Cuzick 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.

The study of Wu (Int J Gynecol Cancer, 2010) received support from Qiagen and Gen-Probe. S. Belinson has no further conflict of interest to declare.

J.R. Belinson 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.

A. Szarewski has no conflict of interest to declare.

The study of Ratnam (Cancer Epidemiol Biomarkers Prev, 2011) received support from Merck Frosst Canada Ltd. S. Ratnam has no further conflict of interest to declare.

M. Reuschenbach has no conflict of interest to declare.

J. Monsonego has received research funding from Gen-Probe Inc. and has participated at Advisory Board of Sanofi Pasteur MSD, Gen-Probe, Qiagen and Roche Diagnostics.
Early Detection and Diagnosis

The purpose of the systematic review was to compare the accuracy of the HPV APTIMA assay and HC2 test to detect underlying cervical intraepithelial neoplasia grade 2 or worse (CIN2+) or CIN3+ in women with a cervical cytology specimen showing ASC-US or LSIL.

Literature search and eligibility criteria

PUBMED-Medline, EMBASE and CENTRAL were used for retrieving references. The search strategy used in PUBMED is provided in the Appendix. Furthermore, the abstract book from the 26th and 27th International Conferences of the Papillomavirus Society (IPV26, Montreal, 2010; IPV 27, Berlin, 2011) and the reference lists of retrieved reports were explored as well. The search started from 2007, the year of the first report on the APTIMA HPV test and included references published up to December 3, 2011. There was no language restriction. In addition, websites of the Food and Drugs Association (US FDA) and the European Medicine Agency were consulted to verify the presence of accuracy data of APTIMA submitted for authorization.

Two cytological groups were distinguished: (i) the ASC-US group, comprising ASCUS/ASC-US (defined according to the 1988/2001 editions of the Bethesda System23,24) or borderline dyskaryosis [defined according to the British Society of Clinical Cytology (BSCC) reporting guidelines] and (ii) the LSIL group, including LSIL24 and mild dyskaryosis. Authors were contacted to provide data separately for ASC-US or LSIL if the number of true and false positives and negatives could not be derived from the reports. The BSCC cytological categories25 borderline and mild dyskaryosis were considered as equivalent to ASC-US and LSIL, respectively. Atypical squamous cells where high-grade lesions cannot be excluded and atypical glandular cells were not targeted in the review.

Evaluated tests

The evaluated index test was the APTIMA HPV test, which detects mRNA coding for the E6 and E7 viral proteins of HPV types 16, 18, 31, 33, 35, 38, 39, 45, 51, 52, 56, 58, 59, 66 and 68, accepting the cutoff for a positive test as proposed by the manufacturer.26 The comparator test was the B-probe of HC2 test, which targets DNA sequences of the same HPV types with the exception of HPV 66 using the standard cutoff as positivity criterion (signal strength, relative light units ≥ 1, expressing semiquantitatively the viral load compared to a control sample with 1 pg of HPV DNA per milliliter).12

Reference standard

All women had to be submitted to verification with colposcopy, colposcopy-directed biopsies, possibly supplemented
with endocervical curettage when glandular disease was suspected. The histological interpretation of biopsies was considered as the outcome, accepting negative colposcopy as sufficient ascertainment for the absence of disease, when no biopsies were taken in the absence of colposcopic changes. Two levels of disease outcome were considered: CIN2+ and CIN3+. Adenocarcinoma in situ was included in the CIN3+ outcome.

Study selection, data extraction and statistical analysis
Two authors (M.A. and J.R.) independently checked eligibility of references and extracted the numbers of the true and false positives and negatives for each test, triage group and outcome. It was also noted whether tested women were attending routine screening or were referred to colposcopy for further assessment. The quality of the selected studies was evaluated using the QUADAS checklist.27

Accuracy parameters were pooled separately using a random effect model.28 The absolute specificity and specificity of the tests were estimated jointly using metandi, a procedure in STATA, based on a two-level mixed logistic regression model, with independent binomial distributions for the true positives and true negatives conditional on the sensitivity and specificity in each study, and a bivariate normal model for the logit transforms of sensitivity and specificity between studies.29,30 The relative sensitivity and specificity of APTIMA compared to HC2 were computed using metadas, a SAS macro for meta-analysis of diagnostic accuracy studies that allows the inclusion of test as a covariate making comparison of tests possible.31,32 The influence of study characteristics, including QUADAS items and the used cytological classification system, on test accuracy were assessed using the HSROC regression of metadas.

Results
Selection of studies
The search in Medline yielded 225 articles, from which six relevant articles, evaluating the accuracy of both the APTIMA and HC2, could be identified.26,33–37 One additional reference could be retrieved from the IPV26 abstract book,38 and another was identified from the FDA web site http://www.accessdata.fda.gov/.39 No additional references were identified from EMBASE or CENTRAL. Two articles that addressed only virological issues,22,40 an article evaluating a prototype test different from the currently used APTIMA assay41 and two secondary publications containing no new data42,43 were excluded. In total, 1,839 women with ASC-US and 1,887 women with LSIL were enrolled in eight studies.26,33–38 Data were requested from the authors of studies who did not report the accuracy data for the outcomes CIN2+ and CIN3+, separately for the two triage groups.26,33–38 All of them provided the requested data and were offered coauthorship to this article. Data from the CLEAR trial, including triage of women aged 21 years or older with ASC-US, were downloaded from the file submitted by the manufacturer for FDA approval.39

Study characteristics
Three studies included women with ASC-US or LSIL identified through primary screening and four studies recruited subjects in colposcopy clinics where women were referred to because of prior cervical abnormalities.26,33,35,36 whereas one study used archived liquid samples reported as LSIL from a cytology biobank with the outcomes assessed through the follow-up pathology registry (see Table 1). Adjudicated cytology results were used to define the ASC-US and LSIL triage groups in the French FASE trial.34 From the CLEAR trial, we considered only the outcomes from colposcopy-targeted biopsies as this type of verification was generally used in the other included studies.39 All studies, collected cervical cells were transferred in PreservCyt, a transport medium developed for ThinPrep liquid-based cytology (Hologic Incorporated, Bedford, MA). By extracting only data from women with ASC-US or LSIL, the spectrum of disease should be considered representative for the study question, with some reservation for a German study where the cytology reporting system did not allow perfect translation into Bethesda terminology.33 Complete verification with an acceptably valid reference standard was systematically provided, and incorporation bias was avoided in all studies. The delay between triage testing and verification was not sufficiently documented in four studies, and independence between HPV testing and verification was unclear in two studies.34 As HC2 and APTIMA produce machine-based results, we can assume independent assessment even when not explicitly reported. Clinical data were not always provided, although this criterion can be considered as captured at least partially by the clinical setting. Equivocal results and withdrawal of cases were not optimally reported in most of the studies (see Table 1).

Accuracy in triage of ASC-US
The sensitivity of APTIMA to triage ASC-US varied between 75 and 100% for the outcome CIN2+ and between 93% and 100% for outcome CIN3+. The specificity ranged from 20 to 81% and from 38% to 81% for CIN2+ and CIN3+, respectively (see Table 2). The pooled sensitivity and specificity of APTIMA was 95.7% (95% CI = 91.5–97.2%) and 56.4% (95% CI = 44.7–67.5%), respectively, for CIN2+ and 96.2% (95% CI = 91.7–98.3%) and 54.9% (95% CI = 43.5–65.9%), respectively, for CIN3+ (see Table 2 and Fig. 1). APTIMA and HC2 showed similar sensitivity for CIN2+ and CIN3+ lesions (pooled ratios not significantly different from unity). However, the specificity of APTIMA was significantly higher (ratio: 1.19, 95% CI = 1.08–1.29 for outcome CIN3+ and ratio: 1.18, 95% CI = 1.08–1.29 for outcome CIN3+; see Fig. 2).

Accuracy in triage of LSIL
In triage of LSIL, the range of variation in sensitivity of the APTIMA test was 75–100% for outcome CIN2+ and 83–100% for outcome CIN3+, whereas for specificity, the ranges were 24–74% (CIN2+) and 23–73% (CIN3+). Considering CIN2+ as the outcome, the pooled sensitivity and specificity

Int. J. Cancer: 000, 000–000 (2012) © 2012 UICC
were 91.0% (95% CI = 85.2–94.7%) and 42.5% (95% CI = 33.3–52.3%), respectively, whereas for CIN3+, the estimates were 96.7% (95% CI = 91.4–98.9%) and 38.7% (95% CI = 30.5–47.6%), respectively (Table 2).

APTIMA was, on average, as sensitive as HC2 but substantially and significantly more specific for CIN2+ (ratio: 1.37; 95% CI = 1.22–1.54) as well as for CIN3+ (ratio: 1.35; 95% CI = 1.11–1.66) (Table 3 and Fig. 2).

### Influence of study characteristics

None of the study characteristics or QUADAS quality items influenced significantly the relative accuracy parameters, with the exception of one situation. In LSIL triage, the relative specificity was not significantly different from unity in the Scottish study where women had repeated LSIL and where samples were selected from a cytological archive.38

Variation of the accuracy by age group was documented in only one study. In the CLEAR trial, specificity of ASC-US triage increased by age for both tests. However, the relative specificity of APTIMA compared to HC2 to exclude CIN2+ decreased by age: 1.27 (1.05–1.53) in age group 21–29 years, 1.14 (0.97–1.27) in age group 30–39 years and 1.04 (0.95–1.13) in women aged 40 years or older.

### Discussion

Our meta-analysis showed greater specificity without loss in sensitivity for the new APTIMA HPV assay in triage of ASC-US.
US and LSIL with respect to the detection of high-grade CIN compared to HC2. HC2 is currently considered as a validated test recommended internationally discriminating between women with ASC-US needing further diagnostic work-up and those who can be called back for routine screening.\textsuperscript{2,44} However, findings on the performance of HPV testing in triage of LSIL, in particular regarding its specificity, are divergent, which are mainly due to local differences in cytological interpretation as reflected by the variation of HPV-positivity rate.\textsuperscript{45} Reflex HPV testing with HC2 in case of LSIL is therefore not generally recommended.\textsuperscript{2,5} An assay with as good sensitivity but with higher specificity than HC2 is required that could be proposed in triage of LSIL. The APTIMA test seems to be such a test with a specificity for excluding high-grade CIN that is 37% (95% CI = 22–54%) higher than HC2 and a sensitivity that is not lower.

Only eight studies have been conducted up to now comparing HC2 and APTIMA and providing separate data for triage of ASC-US or LSIL. However, one study evaluated the accuracy of APTIMA in LSIL, but did not include a comparator test.\textsuperscript{46} Adding this report, comprising 405 women with LSIL, to the meta-analysis changed the pooled accuracy parameters by less than 0.3%. Two other studies assessed the diagnostic performance in women with abnormal cytology without separation by cytological category and using other comparator assays. Ovestad \textit{et al.}\textsuperscript{47} compared APTIMA with two other high-risk HPV DNA tests (Amplicor, Roche Diagnostics GmbH, Mannheim, Germany and COBAS-4800, Roche Molecular Diagnostics, Pleasanton, CA) and with one mRNA test (PreTect HPV Proofer, NorChip AS, Klokkarstua, Oslo, Norway) in a series of women with ASC-US or LSIL without separation. The sensitivity of APTIMA for CIN2+...
was similar to that of the other two HPV DNA tests; however, APTIMA was substantially more specific than Amplicor (ratio: 2.14; 95% CI = 1.23–2.73). Compared to PreTect, a test that detects mRNA of five HPV types, APTIMA was substantially more sensitive (ratio: 1.91; 95% CI = 1.43–2.56) but less specific (ratio: 0.47; 95% CI = 0.34–0.63). Castle et al.\textsuperscript{41} compared a prototype version of APTIMA at two different cutoffs for test positivity with detection of 14 high-risk HPV genotypes with the Linear Array test (Roche Diagnostics GmbH, Mannheim, Germany) in women with diverse abnormal cytology. Although these studies did not contribute to our meta-analysis, they seem to corroborate the statement that APTIMA is as sensitive as other currently used HPV DNA test systems for the detection of CIN2\textsuperscript{+} but often more specific. Moreover, the FASE and CLEAR trials, contributing data to our meta-analysis, also demonstrated similar specificity but higher specificity of APTIMA compared to HC2 in a primary screening setting.\textsuperscript{34,39}

Because APTIMA detects mRNA, which is more fragile than DNA, its accuracy could be compromised by conditions of sample transport and storage. However, APTIMA’s sensitivity did not seem to be affected in the Scottish study where stored cellular material from a biobank was used\textsuperscript{38}; however, in contrast to the other studies, the specificity here was not higher than HC2. Probably the lack of improved specificity in the Scottish study can be ascribed to the fact that all LSIL

Table 3. Pooled relative sensitivity and specificity of APTIMA compared to HC2

<table>
<thead>
<tr>
<th>Triage group</th>
<th>Outcome</th>
<th>Parameter</th>
<th>Ratio (APTIMA/HC2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASC-US</td>
<td>CIN2+</td>
<td>Sensitivity</td>
<td>1.01 (0.97–1.06)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity</td>
<td>1.19 (1.08–1.31)</td>
</tr>
<tr>
<td>CIN3+</td>
<td>Sensitivity</td>
<td>1.01 (0.96–1.06)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity</td>
<td>1.18 (1.08–1.29)</td>
</tr>
<tr>
<td>LSIL</td>
<td>CIN2+</td>
<td>Sensitivity</td>
<td>0.96 (0.92–0.93)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity</td>
<td>1.37 (1.22–1.54)</td>
</tr>
<tr>
<td>CIN3+</td>
<td>Sensitivity</td>
<td>0.98 (0.91–1.06)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity</td>
<td>1.35 (1.11–1.66)</td>
</tr>
</tbody>
</table>

\textsuperscript{1}The metadas SAS macro failed to converge. Therefore, the pooled relative sensitivity and specificity were computed separately as ratios of two proportions using a random effect model.
cases showed persistent cytological abnormality. All studies used the PreservCyt liquid as transport medium. More data are needed to assess the possible influence of other transport media for liquid-based cytology.

The higher specificity of APTIMA may be ascribed to reduced cross-reaction with nontargeted low-risk or intermediate risk types (such as HPV53), which has been reported for HC2. However, we cannot exclude that the increased specificity was partly due to small missed CIN2+ cases, not picked up by colposcopy and colposcopy-targeted punch biopsy, which were HC2 positive but APTIMA negative.

**Clinical implications**

Given widely accepted evidence for the use of HC2 in triage of ASC-US and the similar sensitivity and higher specificity of APTIMA to detect high-grade CIN, we certainly can recommend APTIMA for ASC-US triage.

As HC2 is not generally recommended in triage of LSIL, a more formal judgment of the quality of evidence derivable from our systematic review is warranted based on the following five criteria: (i) consistency, (ii) large and precise effect size, (iii) directness, (iv) study design and (v) absence of publication bias. Although wide variability in absolute accuracy, in particular, in specificity, of both tests, the relative accuracy estimates were consistent (sensitivity ratio always >1 and significantly >1 in four studies; see Fig. 2). Moreover, the magnitude of the difference in specificity was substantial (pooled absolute specificity of 43% for APTIMA versus 29% for HC2, pooled ratio 1.37, with the confidence interval clearly excluding 1). Verification with a reference standard was performed in all cases, and the QUADAS evaluation was considered satisfactory. Moreover, the accuracy could be evaluated separately for ASC-US and LSIL cases making the evidence direct. Statistical tests for publication bias were insignificant (data not shown). Despite the limited number of studies and the low power of the statistical tests for publication bias, we could restrict the risk of publication bias by consulting other sources such as conference reports and the websites of regulatory agencies. Given these arguments, we conclude that APTIMA can also be recommended to triage women with LSIL.

**Conclusions**

The APTIMA HPV test is as sensitive as HC2 to detect CIN2+ or CIN3+ in women with ASC-US or LSIL. The specificity of APTIMA is on average 19% (95% CI = 8–29%) more specific in triage of ASC-US and 37% (95% CI = 22–54%) more specific in triage of LSIL. The APTIMA HPV test might be considered for use not only for triage of women with ASC-US but also for women with LSIL to identify those who need further diagnostic assessment with colposcopy and biopsy.

**References**

Early Detection and Diagnosis

Appendix: Search Strategy for Retrieval of Studies Allowing the Evaluation of APTIMA in Triage of ASC-US or LSIL

#1: Cervix OR cervic*
#3: #1 AND #2.
#4: diagnos* OR test OR assay

#5: HPV OR “human papillomavirus” OR papillomavirus
#6: APTIMA OR RNA OR “ribonucleic acid” OR transcript
#7: #4 AND #5 AND #6
#8: #3 AND #7
#9: screening OR triage OR management OR follow-up
#10: #8 AND #9
#11: Limits: Humans, Female
#12: Publication Date from January 1, 2007 to October 29, 2011
#13: #10 AND #11 AND #12