Gynecologic Oncology 125 (2012) 175–180

Contents lists available at SciVerse ScienceDirect

Gynecologic Oncology

journal homepage: www.elsevier.com/locate/ygyno

Risk assessment and clinical impact of liquid-based cytology, oncogenic human papillomavirus (HPV) DNA and mRNA testing in primary cervical cancer screening (The FASE Study) ☆

Joseph Monsonego [a,*], Michael G. Hudgens [b], Laurent Zerat [c], Jean-Claude Zerat [c], Kari Syrjänen [d], Jennifer S. Smith [e]

[a] Institut of the Cervix, Paris, France
[b] Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA
[c] Laboratoire Lavergne, Paris, France
[d] Department of Oncology & Radiotherapy, Turku University Hospital, Turku, Finland
[e] Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

ARTICLE INFO

Article history:
Received 7 December 2011
Available online 9 January 2012

Keywords:
Cervical cancer screening
HPV RNA test
HPV DNA test
CIN gold standards
Absolute risk
Relative risk

ABSTRACT

Objective. New commercial HPV RNA assays require further validation studies in population-based cervical cancer screening settings.
To assess the performance of (FDA-approved) APTIMA® HPV Assay (AHPV), Hybrid Capture 2 (HC2), in-house PCR genotyping, and ThinPrep LBC in population-based screening, stratified by three histological gold standards.

Study design. A multi-center trial in 5006 women undergoing routine screening in France was designed to compare the absolute and relative risks of diagnosing CIN3+ and CIN2+ lesions by different diagnostic tests.

Results. Reproducibility between the primary and second pathology reading was excellent for CIN3+ and CIN2+ endpoints (Cohen's kappa 0.948 and 0.854). Absolute risks (PPV) of different tests (AHPV, HC2, PCR genotyping, LBC) in diagnosing CIN2+ (15–20%) and CIN3+ (4–6%) were similar for the first, second, and consensus pathology readings. The relative risks of diagnosing these lesions by the four tests were also similar when the first, second or third pathology readings were employed. AHPV had the highest absolute risk of both histological endpoints, and detects 5% to 15% more CIN3+ and CIN2+ lesions, respectively, than LBC. Compared with HC2 assay, the relative risk of AHPV is 24% to 29% higher, with a significant difference in CIN2+ detection. With LBC as reference, AHPV had the best sensitivity/specificity balance measured by AUC (area under ROC curve) comparison test (significant for CIN2+), and the colposcopy referral rate (9.2%) comparable to that of LBC (8.7%).

Conclusions. These data corroborate the suitability of AHPV for the primary cervical cancer screening.

Introduction

Organized cervical cancer (CC) screening based on conventional Pap smear has been effective in reducing this disease in countries where implemented [1–6]. Liquid-based cytology (LBC) seems to be more sensitive than conventional Pap for equivocal cytological abnormalities [7], but both tests have similar performance in detecting high-grade cervical intraepithelial neoplasia (CIN2+) [8]. Comparison between different studies is complex, however, due to the low reproducibility of histological diagnosis of CIN2 [6], to some extent amendable by consensus pathology reading.

High-risk human papillomaviruses (HR-HPV) are causally related to virtually all CC cases and their precursors [9]. Testing for HR-HPV DNA offers several potential advantages [10] in CC screening: i) detection of CIN earlier than by Pap smear [11]; ii) a negative HR-HPV DNA test predicts a low risk for incident CIN2+ which enables extended screening intervals [12–14]; and iii) HPV testing is more sensitive and has a higher negative predictive value (NPV) than cytology [15–18]. In the US, HPV DNA testing has been approved for the triage of ASC-US and for adjunct screening in combination with cytology for women older than 30 years.

Unfortunately, among women between 20 and 29 years, HPV DNA assays have a relatively low specificity for CIN2+ [19–22], which may be missed also by cytology. Therefore, cytological screening has little

☆ FASE (French Aptima Screening Evaluation).

* Corresponding author at: Institute of the Cervix, 174 rue de Courcelles, 75017 Paris, France. Fax: +33 1 47 66 74 70.
E-mail address: jmonsonego@wanadoo.fr (J. Monsonego).

0090-8258/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygyno.2012.01.002

impact on CC incidence among women under 30 years of age [23–25]. Given that a substantial proportion of women ≥30 years are Pap smear-negative [4,22] but HR-HPV DNA–positive (~5–10%), combining cytology and HPV DNA testing potentially results in unnecessary colposcopy referrals (with substantial increase of costs) [15–18].

Compared with commercial HPV DNA assays, significantly fewer HPV RNA tests are available, all based on detection of HR-HPV E6 and E7 mRNA [26,27]. The rational is that detection of E6/E7 mRNA transcripts may provide a higher specificity for CIN2+, because the oncogenic potential of HPV infection depends on expression of these two oncoproteins. The APTIMA® HPV Assay (AHPV; Gen-Probe Inc., San Diego, California) [28,29] detects HPV E6/E7 mRNA of 14 HR types, has been shown to favorably compare with HPV DNA tests (e.g. HC2 assay) in triage settings, and recently (October 2011) received an FDA approval [30–32].

The present multi-center trial was designed to compare the test performance of AHPV, LBC and two HPV DNA tests (HC2, PCR genotyping) in a population-based CC screening in France [33]. In the present communication, we assessed the impact of different histology readings (i.e., changing gold standard) on i) the sensitivity/specificity balance of AHPV, HC2 assay, PCR genotyping assay, and LBC technique, ii) on the absolute risk (PPV) and relative risk of CIN2+ and CIN3+ detection by each assays, as well as iii) on their referral rates to colposcopy (an important measure of cost-effectiveness).

Material and methods

Study design

The design of the FASE (French APTIMA screening evaluation) study and the baseline data have been described in detail [33]. In brief, this population-based study recruited 5006 women (20–65 years of age; 1233 below 30 years of age) for CC screening from private practice gynecologists (17 centers) in Paris metropolitan area, from April 2008 to February 2009. Women were not eligible, if they had undergone total hysterectomy, were pregnant, or had an abnormal cytology in the past 6 months. The study protocol was conducted in accordance with the Declaration of Helsinki, and approved by an Independent Ethics Committee (Pitié-Salpétrière University Hospital).

LBC sample collection and analysis

Cervical samples were collected from the transformation zone (TZ) using a Cervex-Brush® (Rovers Medical Devices, Oss, The Netherlands), which was rinsed into the LBC medium (PreservCyt®, ThinPrep liquid PAP vial; Cytyc Corporation, Marlborough, Mass), and prepared for ThinPrep LBC specimens according to the manufacturer's instructions [33]. All LBC samples were classified according to the 2001 Bethesda System (TBS 2001), and final diagnoses were based on the consensus reading of all discrepant cases by an external reviewer and original cytopathologists.

HPV testing

The HC2 DNA assay is the first FDA-approved HPV test, based on qualitative detection of L1 in 13 HR-HPV types [34]. Specimens were tested according to the manufacturer's instructions, and considered positive using the relative light units to control cut-off (RLU/CO) of 1.0 pg/mL.

The (recently FDA-approved) AHPV RNA assay is based on the qualitative detection of E6/E7 viral mRNA of 14 HR-HPV genotypes (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, and 68) from cervical specimens collected into ThinPrep LBC vials [20,35]. All AHPV assays were performed according to the manufacturer's instructions [30], and considered positive when the signal to cut-off (S/CO) was ≥1.0 [35].

In LBC-, HC2- or AHPV-positive cases, HPV genotyping was performed by GP5/6 consensus-primer PCR [36–38]. HPV genotyping is based on the detection of the virus by MGPs, capable of identifying 14 HR types (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68) and 2 LR types (HPV6, 11) by the fluorescence-labeled probes, using real-time PCR [37,38].

Colposcopy and biopsies

Women positive for LBC, HC2 or AHPV tests were referred for colposcopy. The criteria for referral included ASC-US or greater, or a positive result with HC2 or AHPV test. Colposcopy was performed according to standard operating procedures, using the international (IFCPC) nomenclature [39], and a standardized practice of taking biopsies, as detailed recently [33].

The 3-tier CIN nomenclature was used for biopsy classification, and the most severe abnormality selected for final histological diagnosis. An independent reviewer (KS) re-examined all biopsies (the second reading). In all discrepant cases, the final diagnosis was based on consensus reached by a panel of three pathologists (study consensus reading).

Statistical analysis

Statistical analyses were done using two statistical packages: IBM SPSS 19.0.1 for Windows (IBM, NY, USA) and STATA/SE 11.2 software (STATA Corp., Texas, USA). The pair-wise reproducibility of the three different histology readings was analyzed using regular (Cohen's) kappa and weighted kappa; ICC (with 95%CI). We used the algorithm of Seed et al. [40] to calculate the conventional performance indicators and the area under ROC curve (AUC) (with 95%CI; ROC curves not presented), which characterises the SE/SP balance. The test for equality of AUC (STATA; roccomp test) was used to compare the four assays. In addition to the usual indicators, we focused on the absolute risks for detecting CIN2+ and CIN3+ endpoints by the four assays, as recently described by Zuna et al. [41]. The absolute risk is equivalent to PPV, indicating the detection of true (biopsy-confirmed) disease among all samples testing positive with that assay. The four assays were then compared by their relative risks (RR also called relative PPV; rPPV) for diagnosing CIN2+ and CIN3+, using the risk ratio statistics. In addition, referral rates (%) to colposcopy were calculated based on test positivity for each four assays. All statistical tests were 2-sided, and the values $p<0.05$ were regarded statistically significant.

Results

Table 1 shows the results measuring the reproducibility of the three histology readings, assessed separately for i) all CIN categories, ii) CIN2+ and iii) CIN3+. The final pathology consensus reading was decided in favor of the original pathology reading, resulting in an almost perfect agreement between the first reading and the consensus classification. There was an almost perfect agreement for CIN3+ between the first and second readings, and slightly less agreement for CIN2+ (Cohen's kappa 0.948 and 0.854, respectively).

Table 2 shows the absolute and relative risks of the four tests, with the three pathology readings as the gold standard. At both CIN3+ and CIN2+ endpoints, the highest absolute risk was obtained with the AHPV assay, its relative risk exceeding that of LBC by 7% for CIN3+ and 15% for CIN2+. As compared with the HC2 test, the absolute risk was also consistently higher for the AHPV test, resulting in 24% to 29% higher relative risks. The relative risks of the three HPV tests did not significantly deviate from that of the reference LBC standard, but using HC2 as the reference, AHPV was superior in the detection of the CIN2+ endpoint (but not CIN3+), with RRs varying between 1.27 and 1.29 ($p=0.010$ to $p=0.002$) (Table 2).

Table 1
Reproducibility between the three histology readings of cervical biopsies.

Gold standard	Reproducibility between the gold standards (All CIN categories)[a]		
	Original reading	Re-reading by external pathology reviewer	Study consensus pathology reading
	Kappa (95%CI)	Kappa (95%CI)	Kappa (95%CI)
Original reading		0.847 (0.829–0.864)	0.996 (0.995–0.997)[b]
Re-reading by external pathology reviewer			0.855 (0.838–0.871)[b]
	Reproducibility between the gold standards (CIN3+ Category)[c]		
Original reading		0.948 (0.942–0.953)	0.983 (0.981–0.985)
Re-reading by external pathology reviewer			0.966 (0.962–0.969)
	Reproducibility between the gold standards (CIN2+ Category)[c]		
Original reading		0.854 (0.838–0.869)	0.971 (0.968–0.975)
Re-reading by external pathology reviewer			0.879 (0.866–0.891)

[a] Assessed by weighted kappa (ICC) for all CIN categories, including non-CIN (= flat HPV without CIN).
[b] The basically different practice of the two raters in classifying HPV and CIN1 lesions was decided in favor of the original readers, while making the study consensus diagnosis.
[c] Assessed by non-weighted (Cohen) kappa for two CIN cutoffs (yes/no).

Table 3 shows the AUC values and colposcopy referral rates for the four tests. The highest AUC values are obtained for the AHPV assay for both CIN3+ and CIN2+ endpoints, exceeding that of LBC significantly at the CIN2+ endpoint for all pathology readings. The AUC values of the AHPV assay were significantly higher than those of the HC2 assay for both CIN3+ and CIN2+ endpoints (Table 3; p values in the footnote). Because based on the numbers of positive screening tests, these referral rates remain the same for all pathology readings. The referral rate for colposcopy of AHPV (9.2%) only slightly exceeds that of LBC (8.7%) (p = 0.454). If ASC-US cases are left out from these calculations, the referral rates increase only by 1.2% and 1.3% for AHPV and HC2, respectively (data not shown).

Discussion

Randomized controlled trials (RCT) comparing HPV DNA testing and cytology [15–18] implicate that all HPV DNA techniques are limited by their lower specificity, which may lead to unnecessary

Table 2
Absolute and relative risks of diagnosing CIN2+ and CIN3+ lesions by LBC, APTIMA, HC2, and PCR genotyping, using different gold standards[a].

ASSAY/endpoint	Performance indicators				
	Absolute risk (95%CI)	Relative risk (LBC reference)[b] (95%CI)	Relative risk (HC2 reference)[c] (95%CI)	Significance (Fisher's test)[d] LBC Ref3 /HC2 ref[c]	
First reading CIN3+:				LBC ref	HC2 ref
LBC consensus[e]	5.3 (3.3–7.9)	1.00	1.15 (0.86–1.64)		p = 0.543
APTIMA	6.1 (4.1–8.7)	1.07 (0.82–1.39)	1.24 (0.92–1.67)	p = 0.660	p = 0.165
HC2	4.2 (2.8–5.9)	0.91 (0.72–1.14)	1.00	p = 0.455	
PCR[f]	4.8 (3.1–6.9)	0.95 (0.73–1.24)	1.07 (0.79–1.46)	p = 0.762	p = 0.674
First reading CIN2+:					
LBC consensus[e]	17.6 (14.0–21.7)	1.00	1.08 (0.88–1.33)		p = 0.445
APTIMA	22.7 (18.9–26.8)	1.15 (0.99–1.33)	**1.29 (1.09–1.52)**	p = 0.074	**p = 0.004**
HC2	15.2 (12.6–18.1)	0.95 (0.84–1.08)	1.00	p = 0.455	
PCR[f]	18.4 (15.2–22.0)	1.02 (0.88–1.18)	1.10 (0.93–1.30)	p = 0.795	p = 0.246
2nd reading CIN3+:					
LBC consensus[e]	5.3 (3.3–7.9)	1.00	1.18 (0.83–1.68)		p = 0.365
APTIMA	5.9 (3.9–8.5)	1.05 (0.81–1.37)	1.24 (0.92–1.68)	p = 0.766	p = 0.160
HC2	4.1 (2.7–5.8)	0.89 (0.70–1.14)	1.00	p = 0.364	
PCR[f]	4.6 (2.9–6.8)	0.93 (0.71–1.23)	1.07 (0.78–1.47)	p = 0.646	p = 0.669
2nd reading CIN2+:					
LBC consensus[e]	14.8 (11.5–18.7)	1.00	1.08 (0.86–1.34)		p = 0.524
APTIMA	19.0 (15.5–22.9)	1.14 (0.98–1.32)	**1.27 (1.06–1.68)**	p = 0.121	**p = 0.010**
HC2	13.3 (10.9–16.1)	0.95 (0.83–1.09)	1.00	p = 0.524	
PCR[f]	15.5 (12.5–18.9)	1.02 (0.87–1.19)	1.10 (0.92–1.32)	p = 0.782	p = 0.282
Consensus reading CIN3+:					
LBC consensus[e]	5.5 (3.5–8.3)	1.00	1.16 (0.82–1.64)		p = 0.379
APTIMA	6.3 (4.3–8.9)	1.06 (0.83–1.38)	1.28 (0.95–1.72)	p = 0.665	p = 0.101
HC2	4.3 (2.9–6.1)	0.90 (0.72–1.13)	1.00	p = 0.379	
PCR[f]	5.0 (3.3–7.2)	0.95 (0.73–1.23)	1.08 (0.80–1.45)	p = 0.765	p = 0.679
Consensus reading CIN2+:					
LBC consensus[e]	19.1 (15.4–23.3)	1.00	1.10 (0.90–1.35)		p = 0.321
APTIMA	24.0 (20.2–28.2)	1.14 (0.98–1.32)	**1.29 (1.09–1.53)**	p = 0.096	**p = 0.002**
HC2	16.6 (13.9–19.6)	0.93 (0.83–1.06)	1.00	p = 0.321	
PCR[f]	19.5 (16.2–23.2)	1.01 (0.88–1.17)	1.11 (0.94–1.31)	p = 0.932	p = 0.199

Significant p values are bolded.
[a] First, second and consensus reading; Absolute risk: equivalent to PPV; Relative Risk: test-specific absolute risks compared by risk ratio statistics.
[b] LBC as the reference.
[c] HC2 as the reference.
[d] Risk ratio statistics, with two-sided Fisher's exact test.
[e] In all LBC cytology, the ASC-US cut-off is used.
[f] PCR, with GP5/GP6 primers, done for patients who were biopsied.

Table 3
Sensitivity/specificity balance and colposcopy referral rates of LBC, APTIMA, HC2, and PCR genotyping in detecting CIN3+ and CIN2+ lesions, using different gold standards#.

Assay/endpoint	Performance indicators		
	ROC area AUC (95%CI)	Referral rate to colposcopy*	Significance (ROC comparison)**
First reading CIN3+:			
LBC consensus[1]	0.706 (0.623–0.788)	398/4.541 (8.7%)	Reference
APTIMA	0.781 (0.734–0.828)	458/4.976 (9.2%)	0.076*[1]
HC2	0.699 (0.663–0.735)	691/5.006 (13.8%)	0.996
PCR[2]	0.716 (0.656–0.777)	522/1.225 (42.6%)[3]	0.630
First reading CIN2+:			
LBC consensus[1]	0.689 (0.641–0.736)	398/4.541 (8.7%)	Reference
APTIMA	0.796 (0.767–0.824)	458/4.976 (9.2%)	**0.0001***[2]
HC2	0.691 (0.662–0.720)	691/5.006 (13.8%)	0.201
PCR[2]	0.736 (0.703–0.769)	522/1.225 (42.6%)[3]	0.038
Second reading CIN3+:			
LBC consensus[1]	0.706 (0.623–0.788)	398/4.541 (8.7%)	Reference
APTIMA	0.780 (0.731–0.829)	458/4.976 (9.2%)	0.076*[3]
HC2	0.698 (0.662–0.735)	691/5.006 (13.8%)	0.995
PCR[2]	0.714 (0.652–0.776)	522/1.225 (42.6%)[3]	0.631
Second reading CIN2+:			
LBC consensus[1]	0.687 (0.635–0.738)	398/4.541 (8.7%)	Reference
APTIMA	0.788 (0.757–0.820)	458/4.976 (9.2%)	**0.0007***[4]
HC2	0.712 (0.689–0.735)	691/5.006 (13.8%)	0.194
PCR[2]	0.739 (0.705–0.772)	522/1.225 (42.6%)[3]	0.025
Consensus reading CIN3+:			
LBC consensus[1]	0.710 (0.630–0.791)	398/4.541 (8.7%)	Reference
APTIMA	0.782 (0.736–0.828)	458/4.976 (9.2%)	0.078*[5]
HC2	0.700 (0.665–0.734)	691/5.006 (13.8%)	0.919
PCR[2]	0.718 (0.660–0.777)	522/1.225 (42.6%)[3]	0.653
Consensus reading CIN2+:			
LBC consensus[1]	0.699 (0.654–0.745)	398/4.541 (8.7%)	Reference
APTIMA	0.799 (0.772–0.827)	458/4.976 (9.2%)	**0.0002***[6]
HC2	0.715 (0.693–0.737)	691/5.006 (13.8%)	0.315
PCR[2]	0.740 (0.709–0.772)	522/1.225 (42.6%)[3]	0.050

#First, second and consensus reading; *All test positive cases referred to colposcopy in each setting; ** Test for equality of ROC areas (STATA, roccomp test); 1In all LBC cytology, the ASC-US cut-off is used; 2PCR, with GP5/GP6 primers, only done for patients who were biopsied (n = 1.225); 3If adjusted for all samples, the referral rate of PCR would be 10.4%; *1–6AUC (Area Under ROC curve) comparisons between APTIMA and HC2: *1p = 0.0054; *2p = 0.0001; *3p = 0.0075; *4p = 0.0001; *5p = 0.0038; *6p = 0.0001; significant p values are bolded.

referrals for colposcopy in a screening setting. In this respect, detection of E6/E7 mRNA transcripts e.g. by AHPV assay might provide a more specific test in the primary screening [26–32]. Because of this, AHPV recently (October 2011) received FDA approval for testing women over 21 years with ASC-US cytology and also to screen women >30 years of age as an adjunct to Pap test. The FASE is the first study designed to assess the performance of AHPV in a population-based screening setting [33]. In our previous analysis of a sub-cohort of 4429 of these women, we reported that both AHPV and HC2 are more sensitive than LBC in detecting CIN2+ and CIN3+ [33]. Specificity of AHPV was higher than that of HC2 but similar to that of LBC in all age groups and separately for women less and above 30 years of age [33]. These data substantiate the previous experience in triage settings where APTIMA® HPV Assay was equally sensitive but more specific than HC2 in detecting CIN2+/CIN3+ endpoints [29,31]. Because extensively discussed in our previous communication [33], the stratification by age was not repeated in the present analysis.

Instead of conventional performance indicators, the present analysis focused on calculating the absolute risks of diagnosing CIN3+ or CIN2+ by each assay. The absolute risk is equivalent to PPV, indicating the proportion of true clinical lesions (CIN2+, CIN3+) diagnosed among all samples testing positive with each assay [41]. This is of particular interest in a screening setting, because the best screening test should be the one with the highest PPV, as previously discussed [22]. A test with a high PPV only detects the true lesions without the need to unnecessarily spend resources for adjunct (triage) tests to confirm the diagnosis [22,42]. Absolute risk is also an appropriate means to estimate the risk of developing incident CIN in a longitudinal setting [41]. Estimating the absolute risks for each assay also enables their comparison by calculating the relative risks (or relative PPV, rPPV) using risk ratio statistics (Table 2). This is an alternative to comparison of the diagnostic tests by their AUC values, which reflect the differences in their SE/SP balance [22]. In the present study, we present AUC values only as indicators of the SE/SP balance of each test (without actual ROC curves) and used the test for equality of AUC (roccomp) to compare the different assays [22,43] (Table 3).

Another focus of this study was to determine the impact of different histology readings used as the gold standard. The agreement between the first and second readings (and both of those with the study consensus) was excellent (Table 1). In evaluating these exceptionally high kappa values, two issues need to be considered. First, the basically different practice of the two raters in classifying HPV (with no CIN) and CIN1 lesions was decided in favor of the original reader while making the study consensus. This results in almost perfect agreement between the first reading and the study consensus. Second, because weighted kappa (used for calculating the agreement in classifying all CIN categories) is more forgiving for one-grade discrepancies, the agreement between the second reading and the study consensus reading is almost perfect. This reproducibility assessment clearly indicates an extremely high reproducibility of CIN grading between expert European pathologists. The discrepancies mostly affected classification of low-grade lesions (HPV lesions vs. CIN1), and as such do not have any impact on the performance indicators of the diagnostic tests calculated for the CIN2+ and CIN3+ endpoints.

There were not any dramatic changes in the test performance indicators when any of these three pathology readings was used as the gold standard. In all three settings, the absolute risk of all four tests falls within the range of 4–6% for CIN3+ and within 15–24% for CIN2+. Similarly, the mutual comparison of the four tests by their relative risks gives a similar profile in all three pathology readings; compared with LBC as the reference, the relative risk is always highest for AHPV, remaining practically unchanged (1.05–1.07) for both CIN3+, and (1.14–1.15) for CIN2+. This is regularly followed by PCR genotyping test, which has the second highest relative risk for both CIN3+ (0.93–0.95) and CIN2+ (1.01–1.02). Using the risk ratio statistics, none of the three HPV assays deviate significantly from the LBC test. However, when HC2 is used as the reference, AHPV has a significantly higher relative risk at CIN2+ endpoint of all three histological readings (Table 2). In this comparison, AHPV has relative risk ranging from 1.24 to 1.29, implicating 24% and 29%, respectively, higher absolute risk of detecting CIN2+ as compared with the HC2 assay.

AHPV had the best balance between sensitivity and specificity, as measured by the AUC values (Table 3), in addition to having the highest absolute and relative risks of diagnosing CIN3+ and CIN2+ lesions (up to 15% more than with LBC). The difference between AHPV and LBC was significant at CIN2+ endpoint, irrespective the gold standard, while at the CIN3+ endpoint, the difference did not reach statistical significance ($p=0.076$ to $p=0.078$). However, when AHPV was compared with HC2 assay, the difference is significant for both CIN2+ and CIN3+ endpoints of all gold standards (Table 3; p values in the footnote).

As to the colposcopy referral rates based on test positivity, AHPV has only a slightly higher (9.2%) referral rate as LBC (8.7%) ($p=0.454$). Noteworthy, however, is the fact that the colposcopy referral rate of PCR genotyping (42.6%) (Table 3) is not comparable with the others, because the test was performed only for patients who had any of the other screening tests positive ($n=1,225$). If calculated similarly as for the other tests, the true referral rate after PCR genotyping test would be 10.4%, i.e., not markedly different from that of LBC and AHPV. The referral rate is highest for HC2 assay, 13.8%.

Despite its merits, our study also has some limitations. First, it is a cross-sectional study, precluding the determination of the tests' longitudinal NPV (to substantiate extended screening intervals) or longitudinal PPV (to estimate the absolute risk of developing incident CIN2+). Based on the cross-sectional equivalence of AHPV to HC2 in this and other studies, however, AHPV screening intervals would likely be the same as for HC2 [33]. Second, the cohort also includes women younger than 30 years, among whom both LBC and all HPV tests are known to perform less optimally than among older women [22,33]. Third, histological results were obtained for all women who tested HPV-positive or had an abnormal cytology, but only for a small proportion (14%) of those who were HPV-negative and had normal cytology. Thus, the estimates of SE/SP balance (AUC) are less precise than in setting where histological results are available from all women.

This study has several important implications. First, the inter-rater agreement between experienced European gynecological pathologists in grading cervical cancer precursor lesions in a screening setting in France was excellent. Second, absolute and relative risks of different tests in diagnosing CIN2+ and CIN3+ lesions are practically unaffected by the three gold standards used. This implicates that the three HPV tests (AHPV, HC2, PCR genotyping) compared here by their relative risks (with LBC as the reference) of diagnosing CIN3+ and CIN2+ lesions did perform consistently across the changing histological gold standard. In all three settings, AHPV had the highest absolute risk, detecting 5% to 15% more CIN3+ and CIN2+ lesions, respectively, as LBC in this trial. When HC2 was used as a reference, relative risk of AHPV was even higher; 1.24–1.29. This situation did not change if the four tests are compared using their SE/SP balance by the AUC equality test; AHPV is clearly superior to LBC at CIN2+ endpoint and significantly better than HC2 at both CIN2+ and CIN3+ endpoints.

Conflict of interest statement

J. Monsonego has received research funding from Gen-Probe Inc. and has participated to Steering Committees at Merck, and to the Advisory Board of Sanofi Pasteur MSD, Gen-Probe, Qiagen, and Roche Diagnostics. J. Smith has received research grants or contracts, honoraria, and consulting fees during the last four years from GlaxoSmithKline, Hologic, Gen-Probe, Merck, and Qiagen. K. Syrjänen, M. Hudgens, L. Zerat, and J.C. Zerat have no conflicts of interest to report.

Acknowledgments

The authors would like to thanks the gynecologists who recruited patients for this study: Drs. Jocelyne Brun, Marie-Christine Wind Mazel, Michèle Alia, Claudine Armand, Alice Bonnier-Garnier, Sylvie Holcman, Laurence Avril, Véronique Dapsance, Orly Amar, Céline Waserman, Lisette Pleskof, Myrtille Riera-Ponge, Pascale Sabban, Sylvie Nguyen, Fabienne Wicart-Poque, and Christine Vahdat. The authors would like to acknowledge the collaborators for their contribution to this study: N. Aït Saïd MSc and F. Beghdad MSc (data collection); H. Khiri PhD, A. Martineau MSc, S. Ravet PhD, A. Raymondo PhD (performed assays and collected and analyzed data); M. Ricard (cytology) as well as Dr. P. Halfon and Dr. F. Ruiz (laboratory work). All persons acknowledged herein were compensated for their work. All doctors' visits and procedures that were considered standard screening procedures were reimbursed by health care insurances. Gen-Probe provided financial support for the assays and logistic conduct of the study, but was not involved in the study design, data collection, data analysis and interpretation, or in writing the manuscript. EUROGIN provided research funding for data analysis and secretary assistance.

References

[1] Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. GLOBOCAN 2008, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 [Internet]. Lyon, France: International Agency for Research on Cancer; 2010. Available from http://globocan.iarc.fr.
[2] Miller AB, Nazeer S, Fonn S, Brandup-Lukanow A, Rehman R, Cronje H, et al. Report on consensus conference on cervical cancer screening and management. Int J Cancer 2000;86:440–7.
[3] Syrjänen KJ, Shabalova IP, Ivanchenko O, Kljukina LB, Grunberga V, Syrjänen SM. Reproducibility of classification and correction for verification bias as determinants of performance of Papanicolaou smear cytology in the screening setting: experience from the New Independent States of the former Soviet Union cohort study. Acta Cytol 2009;53:548–57.
[4] Haute Autorité de Santé. Cancer du col de l'utérus. Press release, November 15, 2010. Available from http://www.has-sante.fr/portail/jcms/c_998815/cancer-du-col-de-luterus.
[5] Solomon D, Schiffman M, Tarone R, ALTS Study group. Comparison of three management strategies for patients with atypical squamous cells of undetermined significance: baseline results from a randomized trial. J Natl Cancer Inst 2001;93:293–9.
[6] Stoler MH, Schiffman M. Atypical Squamous Cells of Undetermined Significance-Low-grade Squamous Intraepithelial Lesion Triage Study (ALTS) Group. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. JAMA 2001;285:1500–5.
[7] Monsonego J, Autillo-Touati A, Bergeron C, Dachez R, Liaras J, Saurel J, et al. Liquid-based cytology for primary cervical cancer screening: a multi-centre study. Br J Cancer 2001;84:360–6.
[8] Arbyn M, Bergeron C, Klinkhamer P, Martin-Hirsch P, Siebers AG, Bulten J. Liquid compared with conventional cervical cytology: a systematic review and meta-analysis. Obstet Gynecol 2008;111:167–77.
[9] zur Hausen H. Papillomaviruses in the causation of human cancers—a brief historical account. Virology 2009;384:260–5.
[10] 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;96:280–93.
[11] Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla Palma P, Del Mistro A, et al. New Technologies for Cervical Cancer screening (NTCC) Working Group. Efficacy of human papilloma testing for the detection of invasive cervical cancers and cervical epithelial neoplasia: a randomized controlled trial. Lancet Oncol 2010;11:249–57.
[12] Cuzick J, Szarewski A, Mesher D, Cadman L, Austin J, Perryman K, et al. Long-term follow-up of cervical abnormalities among women screened by HPV testing and cytology—results from the Hammersmith study. Int J Cancer 2008;122:2294–300.
[13] Dillner J, Rebolj M, Birembaut P, Petry KU, Szarewski A, Munk C, et al. Joint European Cohort Study. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. BMJ 2008;337:a1754.
[14] Shi JF, Belinson JL, Zhao FH, Pretorius RG, Li J, Ma JF, et al. Human papillomavirus testing for cervical cancer screening: results from a 6-year prospective study in rural China. Am J Epidemiol 2009;170:708–16.
[15] Mayrand MH, Duarte-Franco E, Rodrigues I, Walter SD, Hanley J, Ferenczy A, et al. Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. N Engl J Med 2007;357:1579–88.
[16] Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla PP, Del MA, et al. Results at recruitment from a randomized controlled trial comparing human papillomavirus testing alone with conventional cytology as the primary cervical cancer screening test. J Natl Cancer Inst 2008;100:492–501.
[17] Baseman JG, Kulasingam SL, Harris TG, Hughes JP, Kiviat NB, Mao C, et al. Evaluation of primary cervical cancer screening with an oncogenic human papillomavirus DNA test and cervical cytologic findings among women who attended family planning clinics in the United States. Am J Obstet Gynecol 2008;199:26–8.
[18] Kulasingam SL, Hughes JP, Kiviat NB, Mao C, Weiss NS, Kuypers JM, et al. Evaluation of human papillomavirus testing in primary screening for cervical abnormalities: comparison of sensitivity, specificity, and frequency of referral. JAMA 2002;288:1749–57.

[19] Sasieni P, Castanon A, Parkin DM. How many cervical cancers are prevented by treatment of screen-detected disease in young women? Int J Cancer 2009;124:461–4.
[20] Chan PK, Chang AR, Yu MY, Li WH, Chan MY, Yeung AC, et al. Age distribution of human papillomavirus infection and cervical neoplasia reflects caveats of cervical screening policies. Int J Cancer 2010;126:297–301.
[21] Cuzick J, Arbyn M, Sankaranarayanan R, Tsu V, Ronco G, Mayrand MH, et al. Overview of human papillomavirus-based and other novel options for cervical cancer screening in developed and developing countries. Vaccine 2008;26:K29–41.
[22] Syrjänen K, Derchain S, Roteli-Martins C, Longatto-Filho A, Hammes LS, Sarian L. Value of conventional pap smear, liquid-based cytology, visual inspection and human papillomavirus testing as optional screening tools among Latin American women <35 and > or =35 years of age: experience from the Latin American Screening Study. Acta Cytol 2008;52:641–53.
[23] Quinn M, Babb P, Jones J, Allen E. Effect of screening on incidence of and mortality from cancer of cervix in England: evaluation based on routinely collected statistics. BMJ 1999;318:904–8.
[24] Sasieni P, Adams J, Cuzick J. Benefit of cervical screening at different ages: evidence from the UK audit of screening histories. Br J Cancer 2003;89:88–93.
[25] Arbyn M, Sasieni P, Meijer CJLM, Clavel C, Koliopoulos G, Dillner J. Chapter 9: clinical applications of HPV testing: a summary of meta-analyses. Vaccine 2006;24S3:78–89.
[26] Molden T, Kraus I, Skomedal H, Nordstrom T, Karlsen F. PreTect HPV-Proofer: real-time detection and typing of E6/E7 mRNA from carcinogenic human papillomaviruses. J Virol Methods 2007;142:204–12.
[27] Castle PE, Dockter J, Giachetti C, Garcia FA, McCormick MK, Mitchell AL, et al. A cross-sectional study of a prototype carcinogenic human papillomavirus E6/E7 messenger RNA assay for detection of cervical precancer and cancer. Clin Cancer Res 2007;13:2599–605.
[28] Dockter J, Schroder A, Eaton B, Wang A, Sikhamsay N, Morales L, et al. Analytical characterization of the APTIMA HPV assay. J Clin Virol 2009;45(Suppl 1):S39–47.
[29] Dockter J, Schroder A, Hill C, Guzenski L, Monsonego J, Giachetti C. Clinical performance of the APTIMA HPV assay for the detection of high-risk HPV and high-grade cervical lesions. J Clin Virol 2009;45(Suppl 1):S55–61.
[30] Waldstrom M, Ornskov D. Comparison of the clinical performance of an HPV mRNA test and an HPV DNA test in triage of atypical squamous cells of undetermined significance (ASC-US). Cytopathology 2011;19. doi:10.1111/j.1365-2303.2011.00923.x. [Epub ahead of print].
[31] 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;17:3033–42.
[32] Getman D, Aiyer A, Dockter J, Giachetti C, Zhang F, Ginocchio CC. Efficiency of the APTIMA HPV assay for detection of HPV RNA and DNA targets. J Clin Virol 2009;45(Suppl 1):S49–54.
[33] Monsonego J, Hudgens MG, Zerat L, Zerat J-C, Syrjänen K, Halfon P, et al. Evaluation of oncogenic human papillomavirus RNA and DNA tests with liquid-based cytology in primary cervical cancer screening: the FASE study. Int J Cancer 2011;129:691–701.
[34] Castle PE, Solomon D, Wheeler CM, Gravitt PE, Wacholder S, Schiffman M. Human papillomavirus genotype specificity of hybrid capture 2. J Clin Microbiol 2008;46:2595–604.
[35] Gen-Probe APTIMA® HPV Assay [package insert]. San Diego, Calif: Gen-Probe Inc; 2008.
[36] Söderlund-Strand A, Carlson J, Dillner J. Modified general primer PCR system for sensitive detection of multiple types of oncogenic human papillomavirus. J Clin Microbiol 2009;47:541–6.
[37] Flores-Munguia R, Siegel E, Klimecki WT, Giuliano AR. Performance assessment of eight high-throughput PCR assays for viral load quantitation of oncogenic HPV types. Mol Diagn 2004;6:115–24.
[38] Lindh M, Görander S, Andersson E, Horal P, Mattsby-Balzer I, Ryd W. (4) Real-time Taqman PCR targeting 14 human papilloma virus types. J Clin Virol 2007;40:321–4.
[39] Walker P, Dexeus S, De Palo G, Barrasso R, Campion M, Girardi F, et al. Nomenclature Committee of the International Federation for Cervical Pathology and Colposcopy. International terminology of colposcopy: an updated report from the International Federation for Cervical Pathology and Colposcopy. Obstet Gynecol 2003;101:175–7.
[40] Seed PT, Tobias A. Summary statistics for diagnostic tests. Stata Techn Bull 2001;59:9–12.
[41] Zuna R, Wang SS, Schiffman M, Solomon D, for the ALTS study group. Comparison of human papillomavirus distribution in cytological subgroups of low-grade squamous intraepithelial lesion. Cancer Cytopathol 2006;108:288–97.
[42] Koliopoulos G, Arbyn M, Martin-Hirsch P, Kyrgiou M, Prendiville W, Paraskevaidis E. Diagnostic accuracy of human papillomavirus testing in primary cervical screening: a systematic review and meta-analysis of non-randomized studies. Gynecol Oncol 2007;104:232–46.
[43] 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;19:385–92.