RESEARCH ARTICLE

Accuracy of Combined Visual Inspection with Acetic Acid and Cervical Cytology Testing as a Primary Screening Tool for Cervical Cancer: a Systematic Review and Meta-Analysis

Phetsavanh Chanthavilay[1,2*], Mayfong Mayxay[1,3,4], Keokedthong Phongsavan[5], Donald E Marsden[5], Lisa J White[6,7], Lynne Moore[2], Daniel Reinharz[2,8]

Abstract

 Background: The performance of combined testing visual inspection with acetic acid (VIA) and cervical cytology tests might differ from one setting to another. The average estimate of the testing accuracy across studies is informative, but no meta-analysis has been carried out to assess this combined method. Objective: The objective of this study was to estimate the average sensitivity and specificity of the combined VIA and cervical cytology tests for the detection of cervical precancerous lesions. Materials and Methods: We conducted a systematic review and a meta-analysis, according to the Cochrane Handbook for Systematic Review of Diagnostic Test Accuracy. We considered two cases. In the either-positive result case, a positive result implies positivity in at least one of the tests. A negative result implies negativity in both tests. In the both-positive case, a positive result implies having both tests positive. Eligible studies were identified using Pubmed, Embase, Website of Science, CINHAL and COCRANE databases. True positive, false positive, false negative and true negative values were extracted. Estimates of sensitivity and specificity, positive and negative likelihood (LR) and diagnostic odds ratios (DOR) were pooled using a hierarchical random effect model. Hierarchical summary receiver operating characteristics (HSROC) were generated and heterogeneity was verified through covariates potentially influencing the diagnostic odds ratio. Findings: Nine studies fulfilled inclusion criteria and were included in the analysis. Pooled estimates of the sensitivities of the combined tests in either-positive and both-positive cases were 0.87 (95% CI: 0.83-0.90) and 0.38 (95% CI: 0.29-0.48), respectively. Corresponding specificities were 0.79 (95% CI: 0.63-0.89) and 0.98 (95% CI: 0.96-0.99) respectively. The DORs of the combined tests in either-positive or both-positive result cases were 27.7 (95% CI: 12.5-61.5) and 52 (95% CI: 22.1-122.2), respectively. When including only articles without partial verification bias and also a high-grade cervical intraepithelial neoplasia as a threshold of the disease, DOR of combined test in both-positive result cases remained the highest. However, DORs decreased to 12.1 (95% CI: 6.05-24.1) and 13.8 (95% CI: 7.92-23.9) in studies without partial verification bias for the combined tests in the either-positive and both-positive result cases, respectively. The screener, the place of study and the size of the population significantly influenced the DOR of combined tests in the both-positive result case in restriction analyses that considered only articles with CIN2+ as disease threshold. Conclusions: The combined test in the either-positive result case has a high sensitivity, but a low specificity. These results suggest that the combined test should be considered in developing countries as a primary screening test if facilities exist to confirm, through colposcopy and biopsy, a positive result.

Keywords: Screening test performance - combined test - cervical cytology - VIA - developing countries

Introduction

 Cervical cancer is the fourth most commonly diagnosed cancer and the fourth leading cause of cancer-related death in women worldwide, and is amenable to both primary and secondary preventative strategies (Ferlay et al., 2014). More than 85% of the cases and deaths occur in developing countries, where cervical cancers account for 12% of all female cancers (Ferlay et al., 2014). In Asia, 6.4 per 100 000 women die each year because of

[1]Faculty of Postgraduate Studies, University of Health Sciences, [3]Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit (LOMWRU), Microbiology Laboratory, Mahosot Hospital, Vientiane, [5]Gynecologic Oncology Unit, Setthathirath Hospital, [8]Institut de la Francophonie pour la Médecine tropicale, Vientiane, Lao PDR. [2]Department of Social and Preventive Medicine, Faculty of Medicine, Laval University, Quebec, Canada. [6]Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. [4]Centre for Tropical Medicine and Global Health, Churchill Hospital, [7]Nuffield Department of Medicine, University of Oxford, UK *For correspondence: phetsavanh456@gmail.com

the disease, but the rates vary largely among different sub-regions. The highest rate is found in South Asia (Bruni et al., 2015). Cervical cancer could be prevented through HPV vaccination and screening as primary and secondary prevention strategies, respectively (Jacob, 2009; Echelman and Feldman, 2012). Several approaches are available for the screening of precancerous cervical lesions. In developing countries, because of resources issues, the main options are cervical cytology and visual Inspection with Acetic Acid (VIA) (Sherris et al., 2009).

Yet, the accuracy of both cervical cytology and VIA tests for detecting cervical precancerous lesions varies from one setting to another. According to a systematic review on 12 studies, cervical cytology sensitivity ranged from 30% to 87% and its specificity from 86% to 100% (Nanda et al., 2000). Meanwhile, sensitivity and specificity estimates for VIA were 72% to 80% and 79% to 92%, respectively (Sauvaget et al., 2011). In India, for instance, screening with VIA could prevent 22 000 deaths due to a cervical cancer each year (Kay, 2013). Nevertheless, VIA, besides its easiness of use and its low cost (Sherris et al., 2009), has interesting characteristics, particularly regarding its sensitivity and its negative predictive value compared to conventional cytology. The sensitivity of VIA is commonly higher than the sensitivity of Cervical cytology, but its specificity for the detection of precancerous cervical lesions is lower, leading to more false positive results (Consul et al., 2012).

There is evidence that in comparison with screening by cytology alone, double testing with HPV DNA and cervical cytology results in a 35% (95% CI = 15% to 60%) increase in sensitivity to detect high-grade cervical intraepithelial neoplasia (CIN) or a cancer, compared to testing with cervical cytology alone (Naucler et al., 2009). Co-testing with these screening techniques is now currently practiced in the USA (Saslow et al., 2012). However, HPV DNA testing is limited in low-resource settings. Another potential combined method for the detection of cervical precancerous lesions would be cervical cytology and VIA as the latter is readily available in low-income countries. A few studies have been published on the topic. However, results diverged. A systematic review and a meta-analysis are still required to evaluate the accuracy and the potential usefulness of this combined test.

Materials and Methods

Search strategy

We conducted a systematic review and meta-analysis in compliance with the guidelines of the Cochrane Handbook for Systematic Review of Diagnostic Test Accuracy (Deeks et al., 2010) and the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) guidelines (Liberati et al., 2009). Articles were searched up to June 2014 in Pubmed, Embase, Website of Science, CINHAL and COCRANE databases using the following key-words: cytology; VIA and sensitivity and their synonyms based on CisMef, without language or publication type restrictions. After removing duplicated records, all citations were included in the citation screening process using EndNote Software, version X6

(Thomson Reuters, 2012). Two reviewers independently screened titles, abstracts and full articles to establish eligibility and extract the data from included studies. A third reviewer was consulted in case of disagreement.

Eligibility criteria

To be eligible, articles had to report data on the sensitivity and specificity of combined VIA and cytology testing. Both VIA and cervical cytology had to be performed in the same women with asymptomatic or symptomatic conditions. Colposcopy and/or biopsy on at least a positive VIA or cervical cytology result had to be selected as a goal standard. Review articles were excluded.

Outcome of interest

The primary outcome was the sensitivity, specificity, positive and negative likelihood ratio (LR+ and LR-) and diagnostic odds ratio (DOR) of combined VIA and cytology testing. A secondary outcome was the difference in sensitivity and specificity ratios between the combined test and the single tests.

Two situations were examined: either-positive result cases and both-positive result cases. In the either-positive result case, a positive result implies positivity in at least one of the tests. A negative result implies negativity in both tests. In the both-positive case, a positive result implies having both tests positive. A negative result implies negativity in one of them.

The definition of a positive result on cervical cytology was low-grade squamous intraepithelial lesion (LSIL) or higher, according to the Bethesda System. The positive result of Visual Inspection with Acetic acid (VIA) was the color of the cervix turning to white when acetic acid is applied. These definitions were used in all included studies.

Quality assessment

Two authors independently examined the risk of bias and applicability using the Quality Assessment of Diagnostic Accuracy Study 2 (DUADAS-2) tool (Whiting et al., 2011). A third author was consulted to solve discrepancies. Items examined included: 1) patient selection, 2) index test, 3) reference standard and 4) flow and timing. Meanwhile, the items examining applicability concerns were 1) patient selection, 2) index test, and 3) reference standard. Each item was rated as high, low or unclear risk or concern.

A study was considered to be of appropriate quality in the following cases: it avoided a case-control study design, it used a randomized recruitment strategy and more than 80% of patients were included in the analysis, the reference standard was performed within two weeks of the combined test, the interpretation of cervical cytology was blinded to VIA result and all patients underwent the same reference standard test.

The study was considered of low quality when it referred to symptomatic patients, patients with high HPV prevalence such as HIV patients, patients with precancerous lesions and invasive cancer. Partial verification bias was considered possible if only some of the included patients underwent the reference standard

test.

Data collection

Two authors independently extracted the data from eligible studies. When results were discordant, a third author was consulted. We extracted information on the characteristics of the study; authors, year of publication, year the study was conducted, setting, study population and design, screener, threshold of cervical cytology positive results, and gold standard. The threshold for a positive result case of cervical cytology was either ASCUS or LSIL. When both ASCUS and LSIL thresholds were reported, we defined low-grade squamous intraepithelial lesions (LSIL) as a positive result because this was the threshold considered in most studies that were included in the analysis.

The true positive (TP), false positive (FP), true negative (TN) and false negative (FN) rates of both combined test and single tests were extracted from individual studies (Macaskill et al., 2010).

Data analysis

We used a bivariate hierarchical random-effects model, as recommended in Cochrane guidelines (Macaskill et al., 2010), using Stata program version 12 (StataCorp LP, College Station, TX, USA) with the metandi command (Harbord and Whiting, 2009). The meta-analytical random-effects model was used to pool and compare the relative ratios of sensitivity and specificity to detect precancerous lesions or cancers, using the combined test as numerator and single tests as denominators. A threshold of $p<0.05$ was used to establish statistical significance. Forest plots were produced to present pooled and individual estimates of sensitivity and specificity and their 95% confidence intervals using Cochrane Review Manager version 5.2 (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark, 2012).

Hierarchical summary receiver operating characteristics (HSROC) curves were generated. Heterogeneity was assessed by evaluating the influence of pre-established variables (site of study "*lower-middle-income countries or other*", the sample size "*more or less than 900*" and

the screener "*Physician or other*") on the DOR using a meta-regression model. The I2 statistic was calculated to quantify heterogeneity (Macaskill et al., 2010). Lower-middle-income countries were defined, according to the World Bank, as countries with a gross national income (GNI) per capita from $1,046 to $4,125 (World Bank, 2014). Statistically significance was set at $p<0.05$ (Macaskill et al., 2010).

Sensitivity analyses

Sensitivity analyses on verification bias and disease positivity criteria were conducted to evaluate the robustness of the results. We restricted the analyses to the five studies without partial verification bias and to five studies with only CIN2+ as a definition of positivity for the disease.

Results

Study characteristics

353 citations were identified based on article titles (Figure 1). After removing duplicates, 233 abstracts were examined. Forty-three were retained for full-text screening. Nine articles were retained. Among excluded articles, 29 did not provide data on the performance of combined VIA and cervical cytology testing and five were duplicates of the same study.

All included articles were based on cross-sectional studies (Table 1). Three were conducted in India and the others in Iran, Pakistan, Sudan, Brazil, Zimbabwe and Kenya. Five studies were conducted in asymptomatic healthy women; one in HIV-positive women, one in symptomatic women and two in women having an unknown clinical condition. The study with the highest sample size, 10,138 women, was a multiple setting study performed in Brazil and Argentina. Most screeners of VIA were trained nurses (55.6%). Most studies used LSIL as a cut-off point for a positive cervical cytology test (seven studies). Meanwhile, high-grade CIN was considered as a
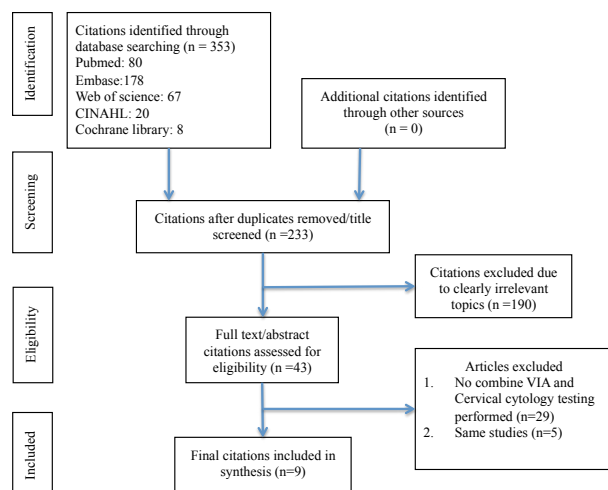

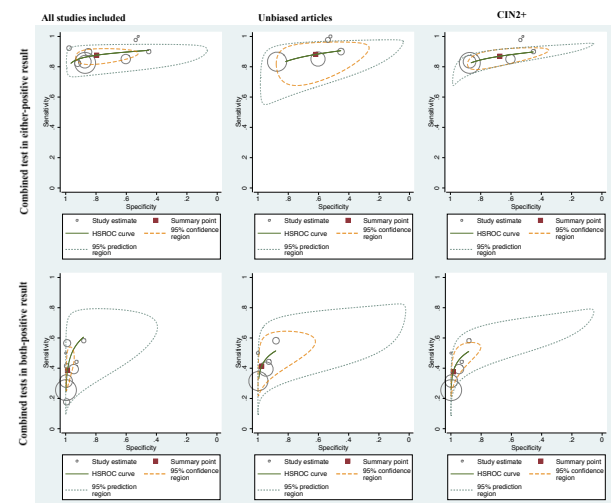**Figure 1. Flowchart of Procedure Performed in Systematic Review**

Figure 3. Hierarchical Summary Receiver Operating Characteristics (HSROC) Curves for the VIA and Cervical Cytology Combined Test in Either-Positive Result and in Both-Positive Result: Restriction and Non-Restriction Analyses

Table1. Characteristics of Included Articles in the Analysis

Authors	Study design	Study site	Clinical condition	Sample size	Screener	Threshold of cervical cytology ¶	Goal standard	Criteria for goal standard performance
Blumenthal et al., 2001.	Cross-sectional study from 1995-1997	Peri-urban primary care clinics in Zimbabwe	Unspecified clinical condition in women between 25 and 55years old	2073	Trained nurse-midwife	LSIL+	Colposcopy and biopsy in positive colposcopy (CIN2+)	All participants underwent colposcopy
Shastri et al., 2005.	Cross-sectional study from 2001-2003	Mobile field clinics in India	Asymptomatic women 30–65 years old	3749	Trained health worker	LSIL+	Colposcopy and biopsy in positive colposcopy (HG-CIN)	All participants underwent colposcopy
Bhatla et al., 2007.	Cross-sectional study in 2003	Gynecology out-patient department in India	Symptomatic women 30-74 year old #	100	Trained-nurse and gynecologist for cervical cytology	LSIL+	Colposcopy and biopsy in positive colposcopy (CIN2+)	All participants underwent colposcopy
Chung et al., 2013.	Cross-sectional study in 2009	OPD during clinical follow-up in Kenya	HIV-positive 18-55 years old	453	Trained-nurse and gynecologist for cervical cytology	LSIL+	Colposcopy and biopsy in positive colposcopy (HG-CIN)	All participants underwent colposcopy
Sahasrabuddhe et al., 2012.	Cross-sectional study in 2006-2007	Out-patient department in India	Non-pregnant and previously unscreened HPV-positive	266	Trained nurses	LSIL+	Colposcopy and biopsy in positive colposcopy (CIN2+)	All participants underwent colposcopy
Ibrahim et al., 2012.	Cross-sectional study in 2009-2010	Primary health care in Khartoum	Asymptomatic married women 25-50 years old	934	Trained physician	ASCUS+	Colposcopy and biopsy in positive colposcopy (CIN1+)	With at least one positive test underwent colposcopy
Longatto-Filho et al., 2012.	Cross-sectional study in 2002-2003	Clinics in Brazil and Argentina	Asymptomatic women 18–60 years old	10138	Trained nurse	LSIL+	Colposcopy and biopsy in positive colposcopy (CIN2+)	With at least one positive test underwent colposcopy
Mahmud et al., 2013.	Cross-sectional study in 2010	Out-patient department in Pakistan	Asymptomatic and symptomatic married women sexually active 19-51 years old	519	Unknown	LSIL+	Colposcopy and biopsy in positive colposcopy (CIN1+)	With at least one positive test underwent colposcopy
Ghaemmaghami et al., 2004.	Cross-sectional study in 1999-2001	Gynecology out-patient department in Iran	Unspecified, women 15-70 years old	1190	Trained midwife and gynecologist	ASCUS+	Colposcopy and biopsy in positive colposcopy (CIN1+)	All positive tests, and 25% randomly of negative results underwent colposcopy

Symptomatic consisted of persistent vaginal discharge, intermenstrual bleeding, post coital bleeding, unhealthy cervix on examination; LSIL+ consisted of low grade squamous intraepithelial lesion or worse; ASCUS+ consisted of atypical Squamous Cells of Undetermined Significance and worse; ¶ Cervical cytology was Conventional cytology with Ayre's spatula and cytobrush; HG-CIN consisted of high-grade cervical intraepithelial neoplasia only; CIN2+ consisted of high-grade cervical intraepithelial neoplasia and invasive cervical cancer;CIN1+ consisted of low-grade and high-grade cervical intraepithelial neoplasia and invasive cervical cancer

Combined VIA and Cervical Cytology in Primary Screening for Cervical Cancer: a Systematic Review and Meta-Analysis

threshold for the disease in six studies. The gold standard test for confirming the cervical precancerous lesions was a colposcopy/direct biopsy (Table 1).

Quality assessment of studies

Overall, two of nine studies met the criteria of high quality according to the QUADAS-2 tool. First, there was no risk of bias in terms of patient selection as all studies were cross-sectional, and all subjects were included in the analysis. However, there were application concerns as nearly half of the included studies did not clearly specify whether participants were asymptomatic or not. The risk of bias in terms of the index test was low; all studies had a clear definition of a positive result for VIA and cervical cytology tests (low risk of bias in terms of index test). Only one study did not specify the occupation of the screeners. Some studies did not specify whether the histology interpretation was blind from the result of the cervical cytology test, leading to a potential concern on risk of bias in terms of the reference standard. Among the nine studies, four had a high risk of partial verification biases, because only some positive results were referred to a reference standard examination (data not shown).

Summary estimates of test performance

Figure 2 presents the summary estimates of the sensitivities and specificities of the combined VIA and cervical cytology tests and of the single tests in detecting cervical precancerous lesions in each study included in the analysis. The range of sensitivity and specificity was large for all tests.

The pooled estimates of sensitivity and specificity of the combined test in the either-positive result case for detecting cervical precancerous lesions were 0.87 (95% confidence interval: 0.83-0.90) and 0.79 (95% CI: 0.63-0.89), respectively. The corresponding values for the combined test in the both-positive result case were 0.38 (95% CI: 0.29-0.48) and 0.98 (95% CI: 0.96-0.99), respectively. The pooled estimates of the positive and negative likelihood ratio and diagnostic odds ratio (DOR) of the combined tests were lower in the either-positive cases compared to the both-positive result cases in all included studies. Details are presented in Table 2.

There was a significant difference in performance between the combined test and the single tests. Compared to the combined test in the both-positive result case, the combined test in the either-positive result case had a significantly higher pooled estimated relative sensitivity, even in the sensitivity analyses restricted to studies without partial verification bias and in the CIN2+ study. Compared to the VIA and cervical cytology tests alone, the combined test in the either-positive result case also had a higher sensitivity. However, its pooled estimated relative specificity was significantly lower than that of the combined test in the both-positive result case or the VIA and cervical cytology tests alone. Meanwhile, the combined test in the both-positive result case had a significant higher pooled estimated relative specificity than the VIA and cervical cytology tests alone in both non-restriction and restriction analyses (results not shown).

Figure 3 shows the hierarchical summary receiver operating characteristics (HSROC) curves of the combined test in the either-positive result case and in the both-positive result case under different scenarios i.e. all included studies, articles without partial verification bias and CIN2+ disease positive threshold analyses. The curves display the joint sensitivity and specificity in each study, showing the individual estimates, the summary estimates, their 95% confidence and the prediction region. Compared to the combined test in the both-positive result case, the summary point of the combined test in the either-positive result case was on the upper-right side, indicating a higher sensitivity and a lower specificity. Additionally, the 95% prediction region for the combined test in the either-positive result case was larger than the combined test in the both-positive result case.

Heterogeneity of diagnostic performance

Heterogeneity between studies was tested with the $I2$ statistic in addition to the influence of covariates on DOR. Results show that the combined test in the either-positive result case and in the both-positive result case presented a large heterogeneity between studies, with an $I2$ statistic higher than 75% (Figure 2).

Table 3 shows that there was no significant association between any covariates and DOR for the combined test in the either-positive result or the both-positive result cases if all studies were included in the meta-regression model. When the analysis was restricted to include only studies with CIN2+ as a threshold of the disease, we found that the place of the study had a significant influence on the DOR of the combined test in the either-positive result as well as in the both-positive result cases. Additionally, other covariates, including the screener and the size of study had a significant influence on DOR of the combined test in the both-positive results case.

Sensitivity analyses

In analyses restricted to articles without partial verification bias and high-grade CIN or worse (CIN2+)
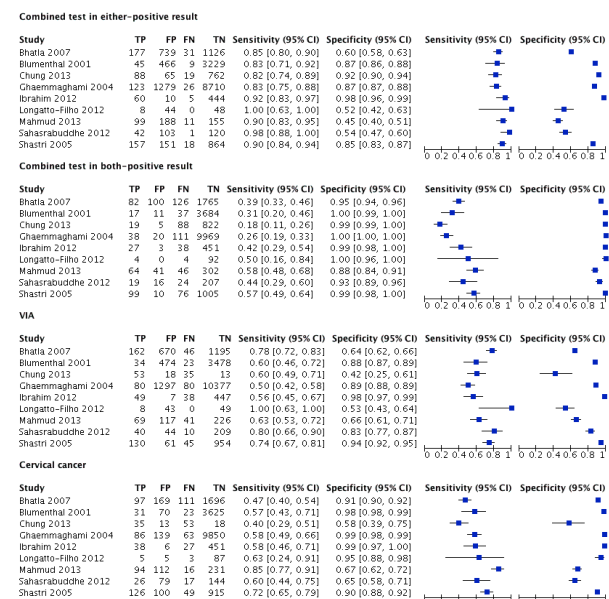

Figure 2. Forest Plot of the VIA and Cervical Cytology Combined test and Single Test

Table 2. Pooled estimates of combined VIA and cervical cytology testing: Meta-analysis results in all studies included, verification unbiased articles and CIN2+

Test/ Category	N	Sensitivity (95% CI)	Specificity (95% CI)	Positive likelihood ratio (95% CI)	Negative likelihood ratio (95% CI)	DOR
Either positive result						
All studies	9	0.87 (0.83-0.91)	0.79 (0.63-0.89)	4.29 (2.26-8.13)	0.15 (0.11-0.21)	27.7 (12.5-61.3)
Unbiased†	5	0.88 (0.80-0.92)	0.61 (0.45-0.76)	2.31 (1.56-3.43)	0.19 (0.12-0.31)	12.1 (6.05-24.1)
CIN2+ℐ	6	0.86 (0.82-0.90)	0.67 (0.50-0.80)	2.67 (1.70-4.17)	0.19 (0.15-0.24)	13.8 (7.92-23.9)
Both positive result						
All studies	9	0.38 (0.29-0.48)	0.98 (0.96-0.99)	32.32 (13.27-78.73)	0.62 (0.54-0.72)	52 (22.1-122.2)
Unbiased†	5	0.41 (0.31-0.52)	0.97 (0.90-0.99)	16.59 (4.85-56.83)	0.61 (0.52-0.7)	27.6 (8.54-89.2)
CIN2+ℐ	6	0.37 (0.28-0.47)	0.98 (0.93-0.99)	23.65 (7.11-78.67)	0.63 (0.55-0.73)	37.3 (12.3-113.1)
VIA alone						
All studies	9	0.67(0.59-0.74)	0.81 (0.66-0.91)	3.74 (1.92-7.32)	0.39 (0.31-0.51)	9.42 (4.17-21.3)
Unbiased†	5	0.73 (0.62-0.82)	0.72 (0.59-0.83)	2.71 (1.82-4.01)	0.37 (0.26-0.51)	7.36 (4.12-13.2)
CIN2+ℐ	6	0.69 (0.57-0.79)	0.76 (0.63-0.85)	2.94 (2.04-4.24)	0.39 (0.29-0.53)	7.41 (4.68-11.7)
Cervical cytology alone						
All studies	9	0.60 (0.50-0.70)	0.91 (0.80-0.96)	7.4 (3.03-18.08)	0.43 (0.33-0.55)	17.3 (6.31-47.7)
Unbiased†	5	0.63 (0.47-0.76)	0.88 (0.72-0.96)	5.73 (2.28-14.39)	0.41 (0.29-0.58)	13.9 (5.2-37.5)
CIN2+ℐ	5	0.62 (0.49-0.73)	0.92 (0.78-0.97)	7.95 (2.97-21.31)	0.41 (0.3-0.54)	19.6 (7.13-54.1)

† Unbiased articles refer to studies that all women with positive result on index test were referred to colposcopy/direct biopsy; ℐ CIN2+: Cervical Intraepithelial Neoplasia or worse; Hierarchical bivariate random effect model was used to analyze the pooled estimate

Table 3. Sources of Heterogeneity Assessment Through the Analysis of Covariates Influencing DORs in All Included Studies, CIN2+ and Asymptomatic Women

Variable	Combined test in either-positive result				Combined test in both-positive result			
	All included studies		CIN2+ ℐ		All included studies		CIN2+ ℐ	
	Coefficient	P	Coefficient	P	Coefficient	P	Coefficient	P
Place of study:								
Lower-middle income countries	89.8	0.32	35.7	0.001	15.1	0.69	146.6	<0.0001
Otherwise	(Ref)		(Ref)		(Ref)		(Ref)	
Size of population								
≥ 900	-124.7	0.14	5.3	0.51	59.7	0.09	152.1	<0.0001
< 900	(Ref)		(Ref)		(Ref)		(Ref)	
Screener								
Physician	-44.2	0.63	13.8	0.19	8.86	0.82	154.9	<0.0001
Otherwise	(Ref)		(Ref)		(Ref)		(Ref)	

ℐ CIN2+: Cervical Intraepithelial Neoplasia or worse; The meta-regression was used to assess the heterogeneity. The influence of covariate on DOR could not done in articles with verification bias through this analysis due to limited number of included studies

as a threshold for the diagnosis of the disease, the same pattern was produced. DORs rank did not change; the DOR of the combined test in the both-positive results case remained the highest. However, the DORs in the restricted analyses were lower than those calculated on all studies. In addition, the specificity of the combined test in the either-positive result case was lower when analyses were restricted to studies without partial verification bias and high-grade CIN as a threshold of positive disease (Table 2).

Discussion

To the best of our knowledge, this is the first meta-analysis aiming to determine the accuracy of combined VIA and cervical cytology testing in detecting cervical precancerous and cancerous lesions. The major findings in this meta-analysis are: 1) under the either-positive result case the combined VIA and cervical cytology test has a higher sensitivity but a lower specificity than under the both-positive result case for detecting cervical precancerous lesions; 2) the sensitivity of the combined test in the either-positive result case was significantly higher than the sensitivities of the VIA or cervical cytology tests alone; 3) specificity of the combined test in the either-positive result case decreased in analyses restricted to articles without partial verification bias and CIN2+ disease positive threshold; and 4) restriction analyses showed that the screener, the place of study and the size of the population are covariates that significantly influence the diagnostic accuracy of the combined test in the both-positive result case.

The low specificity of the combined test in the either-positive result case, compared to VIA or cervical cytology tests alone, is probably due to the fact that a true negative result required negativity of both VIA and cervical cytology. Similarly, low sensitivity of the combined test in the both-positive result case, compared to VIA or cervical cytology tests alone, required positivity of both VIA and cervical cytology. In contrast, the combination of HPV DNA and cervical cytology increases test sensitivity and maintains an adequate specificity (Vesco et al., 2011). Effectively, maintaining the performance of the test requires a high consistence of diagnostic accuracy in both tests to detect and rule out the disease. This might not be the case of VIA and cervical cytology. Result interpretation of these tests is subjective. VIA commonly has a high sensitivity, but a low specificity compared to cervical cytology (Arbyn et al., 2008; Consul et al., 2012). The positive result of VIA could be related not just only to cervical precancerous lesions, but also to inflammation and infections other than HPV infection (Vedantham et al., 2010). Meanwhile, the quality of cervical cytology depends on the quality of the sample collection and the competence of the cytologist in interpreting the result (Denny et al., 2006). As a result, there is in a large variation of the performance of the test, with both VIA and cytology, not only between countries, but also inside countries. For instance, it has been shown that the sensitivity of cytology varied from 28.9 to 76.9% at LSIL threshold in India (Sankaranarayanan et al., 2004).

DOR results lead to the same conclusion as LR+, indicating that the combined test in the both-positive case is the most accurate diagnostic test. The increase of DOR indicates an increase in the discriminating power of the tests (Bossuyt et al., 2013). The highest DOR in the combined test in the both-positive case might be explained by its highest specificity, which was nearly 1 despite its lowest sensitivity.

The combined test in both, the either-positive result and in both-positive result cases had advantages and limits to detect and rule out the disease. Our meta-analysis found a high probability of false positive results (1- specificity) in the either-positive result case, and of high false negative results in the both-positive result case (1-sensitiviy). The false positive result could lead to anxiety and further unnecessary invasive investigation or treatment, which are harmful in terms of physical, psychological and economic burden. In contrast, false negative results yield to considerable delay in diagnostic and treatment particularly when screening interval spreads over several years. This delay might lead to more complicated and advanced stages of the disease, requiring more advanced diagnostic investigations, and consequently delayed treatment and a higher risk of death as found in countries with high incidence and mortality rates of invasive cervical cancer (Bossuyt et al., 2013).

The performance of the combined test varied across studies. This variability might occur as a result of the variability of the performance of both VIA and cervical cytology tests. The result of I2 statistic found consistently large variations between studies in meta-regression analysis. Indeed, meta-regression analysis confirmed this significant variability by exploring the influence of covariates on DOR in restriction analyses, which consisted in including only studies with CIN2+ as the threshold of disease. Our finding is consistent with the study by Chen et al (Chen et al., 2012) that shows that the setting and the size of the population were significantly associated with DOR of VIA in restriction analyses. These covariates did not significantly influence the DOR in non-restriction analyses. This indicated that the influence of covariates depended on study characteristics, particularly the threshold of the disease. To better clarify and rule out the variability of the diagnostic test accuracy, more restriction is probably needed, for instance: restricting the analyses to articles with similar characteristics of test performance (setting, capacity of interpreter and etc.). However, we could not conduct this restriction analysis in our meta-analysis due to the limited number of relevant studies. Further individual studies on the performance of VIA and cervical cytology combined test are apparently required.

The specificity of the combined test in the either-positive result case decreased when analyses were restricted to studies without partial verification bias. This indicates an overestimate of specificity for the combined test in the either-positive result case. Evidently, a partial verification bias can lead to an overestimate of the sensitivities and specificities as a result of a lower proportion of false negatives. The verification biased could be corrected using a Bayesian approach, multiple imputation and the conventional correction method

proposed by Begg and Greenes (de Groot et al., 2011).

As noted, the performance of colposcopy exam is not a perfect test for diagnosing cervical precancerous lesions. Meta-analyses showed that colposcopy had sensitivities ranging from 64% to 99% and specificities from 30% to 93% in the detection of high-grade CIN (Mitchell et al., 1998). In none of the included studies did all women receive a biopsy. The subjectivity of the colposcopy-directed biopsy exam could have affected the pooled estimated sensitivity and specificity found in our meta-analysis (Sideri et al., 1995). Due to the limited number of included studies, the restriction analysis could not be done for this case.

This meta-analysis does have some limitations, which could affect the interpretation of results. First, due to the limited number of studies included, we could not assess the change of sensitivity and specificity among women with ASCUS as positive result of cervical cytology, a low-grade CIN as a disease, the geographical region and symptomatic women. However, the performance of the combined test did not change when the analyses were restricted to articles without partial verification bias and CIN2+, with the exception of the specificity of the combined test in the either-positive result case, which was high compared to non-restriction analysis. This might reflect an overestimation of the specificity of this test.

Second, VIA is recommended only for women aged of 30-45 years. But we could not conduct the analysis in this subgroup due to lack of information on test performance according to the age. This could underestimate the sensitivity due to a greater number of false negative results (FIGO, 2009).

Third, due to the limited number of studies focusing on the diagnostic accuracy of the combination VIA and cervical cytology tests for the detection of cervical precancerous and cancerous lesions, we could not explore the performance of sequential testing cervical cytology in positive VIA cases. This strategy might diminish the false positive rate of VIA, particularly in settings where VIA screening is implemented. Further individual and meta-analytic studies are therefore needed to answer this question.

The combination of VIA and cervical cytology in the either-positive result case gained sensitivity compared to the use a single approach, but lost specificity, contrary to combination in the both-positive result case. Our results suggest that the combined test should be considered in developing countries as a primary screening test if facilities exist to confirm, through colposcopy and biopsy a positive result in order to diminish the number of false positive cases and its consequence, unnecessary treatment.

Acknowledgements

References

Arbyn M, Sankaranarayanan R, Muwonge R, et al (2008). Pooled analysis of the accuracy of five cervical cancer screening tests assessed in eleven studies in Africa and India. *Int J Cancer*, **123**, 153-60.

Bossuyt P, Davenport C, Deeks J, et al (2013). Chapter 11 Interpreting results and drawing conclusions. In 'Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy version 9', Eds The Cochrane Collaboration,

Chen C, Yang Z, Li Z, et al (2012). Accuracy of several cervical screening strategies for early detection of cervical cancer: a meta-analysis. *Int J Gynecol Cancer*, **22**, 908-21.

Consul S, Agrawal A, Sharma H, et al (2012). Comparative study of effectiveness of Pap smear versus visual inspection with acetic acid and visual inspection with Lugol's iodine for mass screening of premalignant and malignant lesion of cervix. *Indian J Med Paediatr Oncol*, **33**, 161-5.

de Groot JA, Bossuyt PM, Reitsma JB, et al (2011). Verification problems in diagnostic accuracy studies: consequences and solutions. *BMJ*, **343**, 4770.

Deeks JJ, Bossuyt PM, Gatsonis C 2010. Cochrane handbook for systematic reviews of diagnostic test accuracy version 1.0.0., The Cochrane Collaboration.

Denny L, Quinn M, Sankaranarayanan R (2006). Chapter 8: Screening for cervical cancer in developing countries. *Vaccine*, **24**, 71-7.

Echelman D, Feldman S (2012). Management of cervical precancers: a global perspective. *Hematol Oncol Clin North Am*, **26**, 31-44.

Ferlay J, Soerjomataram I, Dikshit R, et al (2014). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*, **136**, 359-86.

FIGO (2009). Global Guidance for Cervical Cancer Prevention and Control.

Harbord RM, Whiting P (2009). metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression. *Stata J*, **9**, 211-29.

Jacob M (2009). HPV Vaccines: Characteristics, target population and safety. in 'global guidance for cervical cancer prevention and control', Eds International Federation of Gynecology and Obstetrics, 15-9

Kay M (2013). Screening with acetic acid could prevent 22 000 deaths from cervical cancer in India every year. *BMJ*, **346**, 3935.

Liberati A, Altman DG, Tetzlaff J, et al (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol*, **62**, 1-34.

Macaskill P, Gatsonis C, Deeks JJ, et al (2010). Chapter 10: analysing and presenting results. In 'Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy', Eds The Cochrane Collaboration,

Mitchell MF, Schottenfeld D, Tortolero-Luna G, et al (1998). Colposcopy for the diagnosis of squamous intraepithelial lesions: a meta-analysis. *Obstet Gynecol*, **91**, 626-31.

Nanda K, McCrory DC, Myers ER, et al (2000). Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review. *Ann Intern Med*, **132**, 810-9.

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

Sankaranarayanan R, Thara S, Sharma A, et al (2004). Accuracy of conventional cytology: results from a multicentre

screening study in India. *J Med Screen*, **11**, 77-84.

Saslow D, Solomon D, Lawson HW, et al (2012). 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*, **62**, 147-72.

Sauvaget C, Fayette JM, Muwonge R, et al (2011). Accuracy of visual inspection with acetic acid for cervical cancer screening. *Int J Gynaecol Obstet*, **113**, 14-24.

Sherris J, Wittet S, Kleine A, et al (2009). Evidence-based, alternative cervical cancer screening approaches in low-resource settings. *Int Perspect Sex Reprod Health*, **35**, 147-54.

Sideri M, Schettino F, Spinaci L, et al (1995). Operator variability in disease detection and grading by colposcopy in patients with mild dysplastic smears. *Cancer*, **76**, 1601-5.

Vedantham H, Silver MI, Kalpana B, et al (2010). Determinants of VIA (Visual Inspection of the Cervix After Acetic Acid Application) positivity in cervical cancer screening of women in a peri-urban area in Andhra Pradesh, India. *Cancer Epidemiol Biomarkers Prev*, **19**, 1373-80.

Vesco KK, Whitlock EP, Eder M, et al 2011. Screening for Cervical Cancer: A Systematic Evidence Review for the U.S. Preventive Services Task Force, Rockville MD.

Whiting PF, Rutjes AW, Westwood ME, et al (2011). QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*, **155**, 529-36.

World Bank (2014). Country and lending groups.