Cochrane Library

Cochrane Database of Systematic Reviews

Cytology versus HPV testing for cervical cancer screening in the general population (Review)

Koliopoulos G, Nyaga VN, Santesso N, Bryant A, Martin-Hirsch PPL, Mustafa RA, Schünemann H, Paraskevaidis E, Arbyn M

Koliopoulos G, Nyaga VN, Santesso N, Bryant A, Martin-Hirsch PPL, Mustafa RA, Schünemann H, Paraskevaidis E, Arbyn M.
Cytology versus HPV testing for cervical cancer screening in the general population.
Cochrane Database of Systematic Reviews 2017, Issue 8. Art. No.: CD008587.
DOI: 10.1002/14651858.CD008587.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Cytology versus HPV testing for cervical cancer screening in the general population

George Koliopoulos[1], Victoria N Nyaga[2], Nancy Santesso[3], Andrew Bryant[4], Pierre PL Martin-Hirsch[5], Reem A Mustafa[3], Holger Schünemann[6], Evangelos Paraskevaidis[7], Marc Arbyn[8]

[1]Department of Obstetrics and Gynaecology, Elena Venizelou Maternity Hospital, Athens, Greece. [2]Unit of Cancer Epidemiology, Scientific Institute of Public Health, Brussels, Belgium. [3]Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada. [4]Institute of Health & Society, Newcastle University, Newcastle upon Tyne, UK. [5]Gynaecological Oncology Unit, Royal Preston Hospital, Lancashire Teaching Hospital NHS Trust, Preston, UK. [6]Departments of Clinical Epidemiology and Biostatistics and of Medicine, McMaster University, Hamilton, Canada. [7]Department of Obstetrics and Gynaecology, Ioannina University Hospital, Ioannina, Greece. [8]Unit of Cancer Epidemiology, Belgian Cancer Centre, Scientific Institute of Public Health, Brussels, Belgium

Contact address: George Koliopoulos, Department of Obstetrics and Gynaecology, Elena Venizelou Maternity Hospital, 1 Elena Venizelou Square, Athens, 11521, Greece. georgekoliopoulos@yahoo.com.

Editorial group: Cochrane Gynaecological, Neuro-oncology and Orphan Cancer Group.
Publication status and date: New, published in Issue 8, 2017.

Citation: Koliopoulos G, Nyaga VN, Santesso N, Bryant A, Martin-Hirsch PPL, Mustafa RA, Schünemann H, Paraskevaidis E, Arbyn M. Cytology versus HPV testing for cervical cancer screening in the general population. *Cochrane Database of Systematic Reviews* 2017, Issue 8. Art. No.: CD008587. DOI: 10.1002/14651858.CD008587.pub2.

ABSTRACT

Background

Cervical cancer screening has traditionally been based on cervical cytology. Given the aetiological relationship between human papillomavirus (HPV) infection and cervical carcinogenesis, HPV testing has been proposed as an alternative screening test.

Objectives

To determine the diagnostic accuracy of HPV testing for detecting histologically confirmed cervical intraepithelial neoplasias (CIN) of grade 2 or worse (CIN 2+), including adenocarcinoma in situ, in women participating in primary cervical cancer screening; and how it compares to the accuracy of cytological testing (liquid-based and conventional) at various thresholds.

Search methods

We performed a systematic literature search of articles in MEDLINE and Embase (1992 to November 2015) containing quantitative data and handsearched the reference lists of retrieved articles.

Selection criteria

We included comparative test accuracy studies if all women received both HPV testing and cervical cytology followed by verification of the disease status with the reference standard, if positive for at least one screening test. The studies had to include women participating in a cervical cancer screening programme who were not being followed up for previous cytological abnormalities.

Data collection and analysis

We completed a 2 x 2 table with the number of true positives (TP), false positives (FP), true negatives (TN), and false negatives for each screening test (HPV test and cytology) used in each study. We calculated the absolute and relative sensitivities and the specificities of the tests for the detection of CIN 2+ and CIN 3+ at various thresholds and computed sensitivity (TP/(TP + TN) and specificity (TN/ (TN + FP) for each test separately. Relative sensitivity and specificity of one test compared to another test were defined as sensitivity of test-1 over sensitivity of test-2 and specificity of test-1 over specificity of test-2, respectively. To assess bias in the studies, we used the Quality Assessment of Diagnostic test Accuracy Studies (QUADAS) tool. We used a bivariate random-effects model for computing pooled accuracy estimates. This model takes into account the within- and between-study variability and the intrinsic correlation between sensitivity and specificity.

Main results

We included a total of 40 studies in the review, with more than 140,000 women aged between 20 and 70 years old. Many studies were at low risk of bias. There were a sufficient number of included studies with adequate methodology to perform the following test comparisons: hybrid capture 2 (HC2) (1 pg/mL threshold) versus conventional cytology (CC) (atypical squamous cells of undetermined significance (ASCUS)+ and low-grade squamous intraepithelial lesions (LSIL)+ thresholds) or liquid-based cytology (LBC) (ASCUS+ and LSIL+ thresholds), other high-risk HPV tests versus conventional cytology (ASCUS+ and LSIL+ thresholds) or LBC (ASCUS+ and LSIL+ thresholds). For CIN 2+, pooled sensitivity estimates for HC2, CC and LBC (ASCUS+) were 89.9%, 62.5% and 72.9%, respectively, and pooled specificity estimates were 89.9%, 96.6%, and 90.3%, respectively. The results did not differ by age of women (less than or greater than 30 years old), or in studies with verification bias. Accuracy of HC2 was, however, greater in European countries compared to other countries. The results for the sensitivity of the tests were heterogeneous ranging from 52% to 94% for LBC, and 61% to 100% for HC2. Overall, the quality of the evidence for the sensitivity of the tests was moderate, and high for the specificity.

The relative sensitivity of HC2 versus CC for CIN 2+ was 1.52 (95% CI: 1.24 to 1.86) and the relative specificity 0.94 (95% CI: 0.92 to 0.96), and versus LBC for CIN 2+ was 1.18 (95% CI: 1.10 to 1.26) and the relative specificity 0.96 (95% CI: 0.95 to 0.97). The relative sensitivity of HC2 versus CC for CIN 3+ was 1.46 (95% CI: 1.12 to 1.91) and the relative specificity 0.95 (95% CI: 0.93 to 0.97). The relative sensitivity of HC2 versus LBC for CIN 3+ was 1.17 (95% CI: 1.07 to 1.28) and the relative specificity 0.96 (95% CI: 0.95 to 0.97).

Authors' conclusions

Whilst HPV tests are less likely to miss cases of CIN 2+ and CIN 3+, these tests do lead to more unnecessary referrals. However, a negative HPV test is more reassuring than a negative cytological test, as the cytological test has a greater chance of being falsely negative, which could lead to delays in receiving the appropriate treatment. Evidence from prospective longitudinal studies is needed to establish the relative clinical implications of these tests.

PLAIN LANGUAGE SUMMARY

Human papillomavirus (HPV) test compared to the Papanicolaou (Pap) test to screen for cervical cancer

Review question

We assessed studies comparing two tests to screen for cervical cancer: the HPV test (Human papillomavirus test) and the Pap test otherwise known as cervical smear or Papanicolaou test. The aim was to find out which test detects precancerous changes of the cervix more accurately.

Background

The HPV and the Pap tests are tests that a doctor performs to check for the development of cervical cancer or precancerous changes to the cells of the cervix (called lesions). These lesions can develop into cervical cancer within about 10 to 20 years. The HPV test checks whether a woman has an HPV infection which may lead to cervical cancer. If the HPV test is positive, it may mean that there are precancerous changes in the cervix. There are many types of HPV tests. One of them is called the HC2 test. The Pap test checks for whether cells in the cervix are abnormal. Abnormal cervical cells that are tested as 'low grade to high grade' may mean that there are precancerous changes in the cervix that may lead to cervical cancer. One type of Pap test is 'conventional cytology' and another is 'liquid-based cytology'. Depending on the test, if it is positive a woman may need to have the cervix examined or could receive surgery to have the precancerous lesion removed.

Study characteristics

We searched for all relevant studies up to November 2015. Forty studies compared the HPV test to the Pap test on over 140,000 women between 20 to 70 years old who attended for their routine cervical screening. The studies examined which test can detect precancerous cervical changes which are called cervical intraepithelial neoplasias (CIN 2 and CIN 3).

Quality of the evidence

There were enough studies with enough women in them to allow us to draw conclusions. However, some of the results from the studies were different from each other. For example, tests were more accurate in studies in Europe than in Asia or Central or South America. Overall, the quality of the evidence was moderate to high.

Key results

A perfect test would correctly say if a woman has precancerous changes or if a woman does not. But most tests are not perfect.

This review found that for every 1000 women screened, around 20 women will have precancerous changes. The HPV test will correctly identify 16 of these women (but will miss 4 women). The Pap test will identify 12 of the women (but will miss 8 women). The women who are missed could develop cervical cancer.

For every 1000 women screened, there will be 980 women who will not have precancerous changes. The HPV test will correctly identify 879 women (but 101 women will be incorrectly told that they have a lesion). The Pap test will correctly identify 951 women (but 29 will be incorrectly told that they have a lesion). Women who are incorrectly told that they have a lesion may have their cervix examined or may receive surgery unnecessarily.

BACKGROUND

Screening for cervical cancer meets the prerequisites that the World Health Organization (WHO) dictates as necessary for a useful mass screening programme (Wilson 1968). The disease is common enough to justify mass screening, it is associated with significant mortality, effective treatment is available for pre-invasive or early invasive disease and, finally, detection and treatment of a presymptomatic state results in benefits beyond those obtained through treatment of symptomatic disease. An effective mass screening test, the Pap test, was introduced in the 1940s by George Papanicolaou and is based on the cytological morphology assessment of exfoliated cervical cells (Papanicolaou 1941). Organised screening programmes based on the Pap test have been successful in reducing the incidence of and mortality from the disease, although cancer still does occurs in women who attend for screening (Laara 1987). It has been established that cervical cancer has a strong causal relationship with persistent infection with high-risk human papillomavirus (HPV) types (IARC 2007). Since then, research efforts have focused on the evaluation of a test for the detection of HPV DNA as an alternative method of screening for cervical cancer precursors.

Target condition being diagnosed

Worldwide, there are approximately half a million cases of cervical cancer annually and 85% of cases occur in low- and middle-income countries. Cervical cancer accounts for 10% of all female cancers, making it the fourth leading cause of cancer death in women (Arbyn 2011). It is the third most common gynaecological cancer in the UK, after ovarian and endometrial cancer, although before the introduction of the screening programme it was the most common (Quinn 1999). In high-income countries, the incidence of and mortality from cervical cancer appears to be falling, particularly in countries with systematic screening programmes (Arbyn 2009). Despite this trend, cervical cancer remains the second most common cancer in women in high-income countries under 45 years of age (Arbyn 2011).

Infection of the uterine cervix with the high-risk types of HPV is necessary for the development of cervical cancer, although the HPV infection alone is usually not sufficient to cause cancer. The presence of additional co-factors is required (Bosch 2002; IARC 2007). Most high-risk HPV infections clear spontaneously but in a small proportion of women the infection persists. It is these women who are at risk of developing high-grade cervical intraepithelial neoplasia (CIN) grades 2 or 3 and adenocarcinoma in situ, which are cancer precursors (Schiffman 2007). CIN 2 and 3 can be effectively treated by excision or ablation of the lesion. Over a period of 30 years, untreated CIN 3 has a risk of progressing to

invasive disease in approximately 25% to 30% of cases (McCredie 2008; McIndoe 1984).

Index test(s)

HPV test

Considering that HPV cannot be grown in conventional cell cultures, and serological assays have only limited sensitivity (Dilner 1999), the diagnosis of HPV infection requires the detection of its genome in cellular samples collected from the site under investigation. In the case of the uterine cervix the test is performed by collecting exfoliated cervical cells, similar to the Pap test. Specimens can be collected either by a healthcare provider during a pelvic examination, or through self-sampling in the convenience of the woman's home. Molecular technologies for the detection of HPV DNA can be broadly divided into amplified and non-amplified. The tests mainly used in clinical research use amplification methods, which are further divided into signal amplified and target amplified. The main representative techniques of each category are the hybrid capture 2 (HC2; Digene Corporation, Gainthersburg, MD, USA) assay and polymerase chain reactions (PCR), respectively.

HC2 is a Food and Drug Administration- (FDA) approved test for HPV detection. The B probe of HC2 can detect infection from any of 13 high-risk types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and 68) but separate typing is not possible. The number of viral copies that have to be present per sample in order to obtain a positive result is 5000. HC2 succeeded an earlier test, the hybrid capture tube, which detected four fewer high-risk types and had a higher threshold for positivity (50,000 viral copies per sample). That is, it had lower sensitivity than HC2 and is therefore not currently used.

PCR is a chemical reaction resulting in the synthesis of a large number of target HPV DNA copies. It allows testing on scanty cell samples, small amounts of DNA, or few viral copies and consists of two main steps. The first step is the amplification of the target DNA. This is performed with a thermocycling process (heating and cooling) and the use of oligonucleotide primers. The primers are usually consensus or general, meaning that they can be used to amplify a broad spectrum of HPV genotypes. They are aimed mainly at the L1 region of their genome. More recently PCR assays have been developed that target sequences of the E genes of the virus. Type-specific primers that amplify a particular HPV genotype can also be used, though rarely. There are various designs of general primers currently available. They differ in the size of the DNA region they amplify and in measures taken to compensate for the problem of intertypic sequence variation of the target DNA sites. The GP5+/6+ primers amplify a 150 bp fragment and have to be used at a low annealing temperature in order to compensate for the mismatches with different genotypes. The MY09/11 primers amplify a 450 bp fragment and consist of a complex mixture of oligonucleotides in order to make up for intertypic variation. The PGMY primers amplify the same region of DNA as MY09/11 but contain inosine, which matches any nucleotide. The SPF10 system is another example of inosine-containing primers and targets a 65 bp region. Finally the CPI/II primers amplify a 188 bp region of the E1 gene.

The second step of the PCR process is the detection and analysis of the PCR products. The amplified DNA sequence can be detected by agarose gel electrophoresis. However type-specific analysis is possible and this can be achieved by a variety of methods such as restriction fragment length polymorphism, Southern blotting, microtiter plate hybridisation, direct sequence analysis and reverse hybridisation.

There are several new HPV assays that identify separate HPV genotypes: only the two main oncogenic types HPV16, HPV18 (for instance Cobas 4800, Cervista, Abbott RTPCR) or the full range of high-risk HPV types and even a limited or extended number of non high-risk HPV types. A review of the range of new HPV tests In the current review can be found in Poljak 2012. Only detection of high-risk HPV types is addressed in the current review.

The basic disadvantage of HPV DNA detection methods in clinical practice is their low specificity. This is because HPV infections are usually transient and most of them do not cause any serious consequences. Only a small proportion of HPV infections initiate an oncogenic process that will eventually lead to the development of precancer CIN and invasive cancer. Women with active HPV infection will express E6/E7 oncogenes. These are required for malignant transformation, by inhibiting the tumour suppressors p53 and RB. The E6/E7 mRNA transcripts are detected by mRNA-based molecular techniques and may therefore be of higher prognostic value, improving the specificity and positive prognostic value compared with the HPV DNA testing used in screening. The most widely used mRNA tests, are the PreTect HPV Proofer assay (NorChip AS, Klokkarstua, Norway), which detects only five (16, 18, 31, 33 and 45) high-risk HPV types (Chan 1999) and the APTIMA test (Hologic, Add Cyty, USA), which detect E6/E7 RNA of 14 high-risk HPV types (Arbyn 2012).

This review will not examine other molecular markers of HPV infection such as P16 and L1 immunostaining.

Comparator test: Pap test

Until recently, in the developed world, screening for cervical cancer was carried out by means of cytological examination of a cervical smear (the Pap test). After visualisation of the cervix with the use of a speculum the specimen is obtained with a sampling device, usually a spatula or a brush, which is rotated on the cervix. The collected material is applied to a glass slide (for conventional cytology) or the sampling device is rinsed in or left in a preservative solution (for liquid-based cytology (LBC)).

Cytologists reading the Pap tests usually follow the Bethesda classification system for reporting cervical cytologic diagnoses (Solomon 2002). In this system the smears are reported as negative for intraepithelial lesion or malignancy; atypical squamous cells of undetermined significance (ASC-US); atypical squamous cells, cannot exclude high grade lesion (ASC-H); low-grade squamous intraepithelial lesion (LSIL); high-grade squamous intraepithelial lesion (HSIL); squamous cell carcinoma; atypical glandular cells (ACG); adenocarcinoma in situ (AIS); or adenocarcinoma. Women with an abnormal Pap test should be referred for further investigation, which includes either repetition of the cytology, HPV triage or colposcopy (Jordan 2008; Wright 2006). Cervical smears in the UK are reported using the British Society of Cervical Cytopathology (BSCC) terminology, which includes the categories of negative, inadequate, mild dyskaryosis, moderate dyskaryosis, severe dyskaryosis, possible invasive cancer, glandular neoplasia, and borderline changes. Women in the UK are referred for colposcopy if three consecutive smears are reported as inadequate; two consecutive smears as borderline; or any smear is reported as mild, moderate or severe dyskaryosis, possible invasive cancer or glandular neoplasia (NHSCSP 2004).

The European executive policy is that women between the ages of 25 and 65 years are invited to have a cervical smear test every three to five years (Arbyn 2010). The establishment of a population-based screening programme with the ideal screening interval involves considerable infrastructure, workforce and equipment costs, which can be a barrier for implementation in low- and middle-income countries.

Rationale

It is proven that 80% of cervical cancer can be prevented by well-organised, high-quality screening programmes using Pap smears with three- to five-year screening intervals (IARC 2005). With well-organised programmes, mortality from the disease can be reduced by up to 80% (IARC 2005). Some of the Nordic countries are good examples in this respect (Sigurdsson 1999). On the other hand, in several countries a decrease in cervical cancer incidence of only 40% to 65% has been documented. There are still countries with very high death and morbidity rates from this disease and with no historical decrease in the rates (Arbyn 2009; IARC 2005). Various shortcomings of cervical cytology screening have been suggested as the source of this observation. One of them is the relatively low sensitivity of a single Pap test, even though the longitudinal sensitivity of repeated cytology is higher. In cancer screening a high rate of false negative results is a serious weakness. Therefore a more sensitive screening test is desirable. A systematic review of cervical screening failures in countries with organised screening programmes showed that, among the women who developed cervical cancer, 20% to 55% had had false-negative smears 0 to 6 years prior to the diagnosis (Spence 2007). However, this result should be interpreted cautiously as the percentage of cancers that are cytologically negative is in direct proportion to screening coverage. In a population with complete coverage all cervical cancers that still occur would necessarily be due to screening or follow-up failures.

Apart from the issue of low sensitivity, there are other concerns about the Pap smear test. There is considerable variation in the organisation and implementation of cervical cancer screening programmes within European countries (Anttila 2004). Infrastructure and resources in health care are not sufficient in many areas to build up an effective programme based on cytology. Even in several high-income countries, a large proportion of the target women remain totally unscreened (Breitenecker 2004), forming a high-risk group for cervical cancer. Moreover, very frequent screening intervals of young women may be associated with growing anxiety, over-treatment and unfounded costs. Finally, there are concerns about the quality of the Pap test.

Given that HPV is the cause of cervical cancer and that HPV DNA is detected in virtually all cervical cancers (Walboomers 1999), new screening techniques based on HPV DNA testing have raised hopes and expectations for better prevention of the disease. Testing for HPV DNA is one of the most intensively studied alternatives to cervical cytology screening. The role of HPV testing has already been established and its use has gained wide acceptance in certain areas such as the triage of Pap smears with atypical squamous cell changes (ASCUS smears) and follow-up after treatment (Arbyn 2004; Arbyn 2006). Its role in general population screening is still being discussed.

OBJECTIVES

The main objective of this review was to determine the diagnostic accuracy of HPV testing for detecting histologically confirmed CIN 2 or worse (CIN 2+), including adenocarcinoma in situ, in women participating in primary cervical cancer screening; and how it compares to the accuracy of cytological testing (liquid-based and conventional) at various thresholds.

Secondary objectives

Secondary objectives of the review were:

- to determine the accuracy for each test at prespecified thresholds and the accuracy of different HPV testing techniques;
- to investigate sources of heterogeneity of test accuracy in the included studies. As possible sources of heterogeneity we assessed the influence of the following covariates: the geographical location where the study was conducted, the age limits of the study population, the number of HPV types that the HPV test detects, and the likelihood of verification bias.

METHODS

Criteria for considering studies for this review

Types of studies

We looked for comparative test accuracy studies where all participants had received both HPV testing and cervical cytology (paired studies) followed by partial or complete verification of the disease status with the reference standard (see below). Studies where participants were randomised to receive either only the index test or only the comparator test were not eligible to be included. Our review focuses on paired studies because the comparison of index tests in such studies is (potentially) more valid because key factors can be held similar, including population and reference standard procedure.

Participants

Women participating in a cervical cancer screening programme who were not being followed up for previous cytological abnormalities. The study population could not be part of a case-control design (with a predetermined proportion of known disease positives to known disease negatives). Rather, women had to form a consecutive series; they had to be recruited as a single group with their disease status being unknown at the time of recruitment. The women had to be close to or within the age range suitable for cervical screening according to international guidelines (20 to 70 years).

Index tests

Only HPV tests that are still currently used in clinical research practice were considered. These are:

- HC2 or newer improved signal amplification methods;
- PCR using the following primers GP5+/GP6+, MY09/11, SPF10, or CPI/II;
- Aptima (HPV E6/E7 mRNA testing);
- other techniques that were identified during the search process.

For the HC2 method we considered two thresholds for the definition of a positive result: 1 pg/mL and 2 pg/mL; and for the other techniques the threshold used by the researchers.

Comparator tests

For conventional cytology or liquid-based cytology we considered two thresholds that define an abnormal Pap smear: ASCUS or worse, and LSIL or worse (Solomon 2002). In studies where the cytology was reported in other systems (that is the BSCC terminology or the Second Munich Cytological Classification) we converted the results to the nearest equivalent in the Bethesda system (Solomon 2002). We considered the borderline category of the BSCC and the Pap IIw category of the Munich classification as equivalent to the ASC-US category. We considered the mild dyskaryosis category of the BSCC and the Pap IIID category of the Munich classification as equivalent to the LSIL category.

Target conditions

The target condition was high grade CIN 2 or worse. Some studies used the threshold of CIN 3. We included these in the review but analysed them separately.

Reference standards

As a reference standard, we used the combination of colposcopy and histology. If colposcopy was normal, we did not require a histologic result for proof of absence of disease. If colposcopy was abnormal and a biopsy was taken, then we used the histologic result as the reference standard. We assumed that the histologic examination of material obtained by colposcopy-directed biopsy, loop excision or endocervical curettage provided complete assessment of the considered disease status.

Colposcopy as a reference standard is a subjective examination and has low sensitivity for the detection of small CIN 3 lesions (Jeronimo 2006). On the other hand its performance in quality-assured settings is not at all insensitive for clinically important CIN 3. The ideal reference standard for the evaluation of a cervical screening test would be the excision of the whole transformation zone and its subsequent histopathological examination. Given that such a procedure in healthy women is ethically unjustifiable, due to its morbidity, studies have to rely on colposcopy with directed biopsies even with its limitations.

In this review we included studies where the reference standard was used in one of three ways:

- applied to all women;
- applied to all women with a positive screening test and to a random sample of screen-negative women in order to correct for verification bias;
- restricted to those with a positive screening test.

This last category of studies is prone to verification bias if the double test negatives are considered to be true negatives. However, verification bias will be limited when one of the screen tests is very sensitive. These studies can produce unbiased estimates of relative sensitivity and relative false positive rates (Arbyn 2009a; Schatzkin 1987).

Search methods for identification of studies

Electronic searches

We performed a systematic literature search of articles (1992 to November 2015) that contained quantitative data. We started our search from 1992 because HPV testing for clinical use was not introduced until a few years later.

We retrieved articles from the electronic bibliographic databases:

- MEDLINE, through PubMed (January 1992 to November 2015);
- Embase (January 1992 to November 2015).

The search strategies for MEDLINE and Embase are given in Appendix 1. The service provider that we used to access Embase was Ovid. We used studies that we had identified as relevant as seeds in Scopus to identify articles citing the relevant studies, and used the 'related articles' feature in PubMed, to retrieve articles which were similar in terms of keywords and database subject headings to the original included studies.

The search was restricted to articles written in the English language.

Searching other resources

We checked the reference lists of articles identified as relevant for additional relevant articles, and the reference lists of these were in turn checked for relevance. We contacted authors of relevant articles in order to obtain missing data.

Data collection and analysis

Selection of studies

One review author (GK) assessed the titles and abstracts from the literature search to determine whether they met the eligibility criteria. If there was any doubt we retrieved the full text of the article. Another review author (PMH) then reviewed the search results and the articles detected by the first review author in order to increase the specificity of the search. For any disagreements the third review author (MA) was consulted. The selection process was not blind (that is the names of the authors and institutions were not concealed). A list of the excluded studies is provided including the reasons for exclusion (Characteristics of excluded studies).

Data extraction and management

One review author (GK) collected data on the following using an electronic data collection form:

- study design;
- number of participants;
- age range of participants;
- threshold for the definition of a positive screening result;
- index and comparator tests;
- method used as reference standard;
- threshold used for the definition of disease (e.g. CIN 2+, or CIN 3+);
- the number of true positives, false positives, true negatives, and false negatives in a 2 x 2 table completed for each screening test used in each study.

A second review author (PMH) double-checked the electronic data collection form.

Assessment of methodological quality

To assess the methodological quality of the included studies, two review authors (GK, PMH) used the 'Quality Assessment of Diagnostic test Accuracy Studies (QUADAS) tool (Whiting 2003). The results for each study are presented in table form. The application of QUADAS items to the current review is explained in Appendix 2.

Statistical analysis and data synthesis

We extracted the numbers of true positives, false negatives, false positives and true negatives defined at the considered thresholds from each study. We calculated the absolute and relative sensitivities and the specificities of the tests for the detection of CIN 2+ and CIN 3+ at various thresholds and we computed sensitivity (TP/(TP + TN) and specificity (TN/ (TN + FP) separately for each test. Relative sensitivity and specificity of one test compared to another test were defined as sensitivity of test-1 over sensitivity of test-2 and specificity of test-1 over specificity of test-2, respectively.

We used a bivariate random-effects model analysis (BRMA) as has been described by Chu 2006 and Reitsma 2005. The BRMA preserves the two-dimensional nature of the original data. It allows the meta-analyst to take into account the within- and between-study variability and the intrinsic correlation between sensitivity and specificity. When there were only three studies, a reduced BRMA model with zero covariance component was fitted (univariate random-effects model; URMA). When there were only two studies, we further reduced the BRMA by excluding the random-effects to a univariate fixed-effects model (UFMA). The BRMA, URMA and UFMA were all programmed and fitted using SAS PROC NLMIXED in SAS 9.4 (SAS 9.4; Takwoingi 2010). We performed direct comparison, comparing two tests, by including one test as a covariate in the BRMA model. We first derived the relative measures from the parameters of the models in the log scale and later exponentiated. Using STATA 14 (STATA 14), the binomial distribution using the cii command was used to compute the exact confidence intervals when there was only one study. The standard errors for the log relative sensitivity and specificity were obtained using the delta method, which was internally implemented in SAS. For one study, the asymptotic standard error of the log relative sensitivity and specificity was computed in STATA

14. The resulting 95% Wald confidence intervals for the log relative sensitivity and specificity were subsequently exponentiated to yield relative accuracy measures within the 0 to infinity range.
Given that heterogeneity is likely to be present in many meta-analyses, we considered that a mixed model that uses all of the available data seemed preferable to conducting multiple analyses on subsets of the data using a range of statistical methods. In particular, in the studies where a random sample of test negatives was verified (Reference standards category 2) we did not put the 2 x 2 data directly into Review Manager 5 (RevMan 5) but first calculated the adjusted number of screening test true and false positives and negatives given the proportion of the verified population (RevMan 2014).

Investigations of heterogeneity

For investigation of the sources of heterogeneity, we performed multiple BRMA regressions, each one including one covariate. The covariates that we considered for such analyses were the geographical location where the study was conducted (continent or sub-continent), the age limits of the study population, the number of HPV types that the HPV test detects, and the likelihood of verification bias.
The possible effect of some other important possible quality-related variables that could cause heterogeneity, such as the type of cytology (liquid-based or conventional), the type of HPV testing (HC2 or PCR), and the positivity thresholds was avoided by considering conventional cytology, LBC, HC2, PCR at different thresholds as separate tests.

Sensitivity analyses

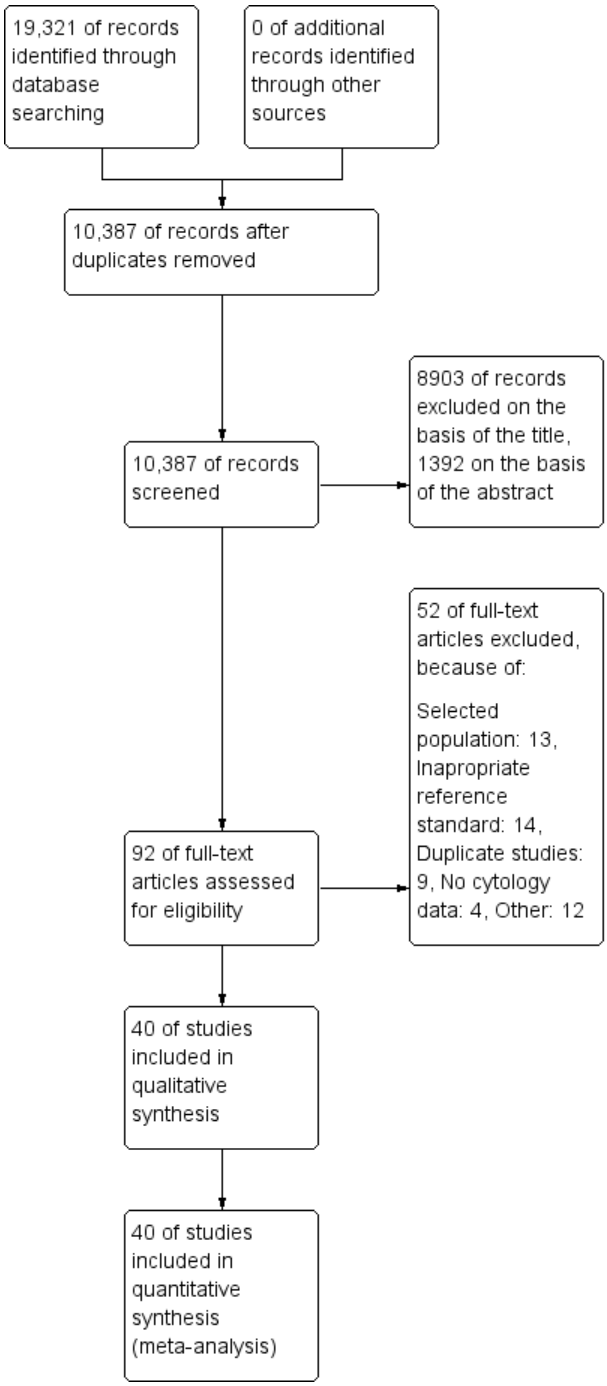
We performed a separate analysis on the accuracy of HPV testing in women over 30 years of age. Studies where the population was strictly over 30 years of age were included in this analysis. This age group was selected as the likelihood of persistent HPV infection and subsequent development of (pre-cancer) is substantially higher in women older than 30.

RESULTS

Results of the search

Total hits from MEDLINE: 9387
Total hits from Embase: 9934
The PRISMA flow chart with details of the results of the search, the exclusions and the inclusions is given in Figure 1 (Moher 2009).

Figure 1. Study flow diagram detailing the number of the initially retrieved articles and consequent exclusions



Ultimately 40 studies were included in this systematic review, which used the following tests.

- Conventional cytology was used in 22 studies
- Liquid-based cytology in 20 studies
- HC2 for HPV DNA testing was used in 27 studies
- HC2+4 for HPV DNA testing (an expanded version of HC2 that tests for 4 additional HPV types) in one study
- PCR for HPV DNA testing in 10 studies
- The Cobas HPV DNA test in two studies
- The Care HPV DNA test in two studies
- The SNIPER HPV DNA test in one study
- The NASBA HPV E6/E7 mRNA test in one study
- The Aptima HPV E6/E7 mRNA test in four studies

Regarding the geographical location of the studies, 18 studies took place in European countries, three studies were in Africa, four studies in Central and South America, 10 studies in Asia (China and India), two studies in the Pacific, and three studies in North America. The earliest study was published in 1995, with the majority of the studies published between 2002 and 2011.

Methodological quality of included studies

A description of each QUADAS item is given in Appendix 2. The first QUADAS item was answered 'yes' (i.e. the tests are done on a random sample of women within the cervical screening age range (20 to 70 years) not being followed up for cervical abnormalities) in 37 of the 40 studies. An appropriate reference standard (QUADAS 3) was used in all studies (colposcopy with directed biopsies as minimum). The fourth QUADAS item (i.e. the total interval between cytology, HPV testing and verification with the reference standard was less than 12 weeks) was answered 'yes' in 38 of the40 studies, the fifth (i.e. all women or at least a random sample of all women tested with cytology or HPV testing had disease status verification by the reference standard) in 26 of the 40 studies, the sixth (i.e. all women who had disease status verification, had this done by the same method) in 39 of the 40 studies, the seventh (i.e. the reference standard used for disease status verification is not composed in any part by cervical cytology or HPV testing) in 39 of the 40 studies, the tenth (i.e. the cytologists and the technicians interpreting the Pap smear and the HPV test were not aware of the colposcopy/biopsy results) in 19 of the 40 studies, the eleventh (i.e. the colposcopists and the pathologists were not aware of the cytology and HPV test results when interpreting the results of the reference standard) in 34 of the 40 studies, the twelfth (i.e. the cytologist was aware of the woman's basic history) in 2 of the 40 studies (this item was scored yes only if it was explicitly stated in the study that the cytologists were given the relevant information about each woman), the thirteenth (i.e. the numbers of inadequate cytology and HPV test results are given) in 29 of the 40 studies, and the fourteenth (i.e. it is clear what happened to all participants who entered the study including the withdrawals) in 34 of the 40 studies.

The criteria that were the hardest to be scored as 'yes' in this meta-analysis were the fifth, the tenth, the twelfth and the thirteenth. The fifth criterion (was partial verification avoided?) was answered 'no' in 14 studies that applied the reference standard (colposcopy) only to women with a positive screening test (Belinson 2003; Belinson 2010; Clavel 2001; Cuzick 1995; Cuzick 1999; Labani 2014; McAdam 2010b; Naucler 2009, Nieves 2013; Ronco 2006, Salmeron 2003; Shipitsyna 2011; Syrjanen 2002; Wu 2010). The tenth criterion (were the reference standard results blinded?) was answered 'no' in 11 studies where colposcopists were aware of the cytology or the HPV test results, and 'unclear' in 10 studies where there was no specific mention in the paper. The twelfth criterion (relevant clinical information given to the people reporting the screening test?) was answered as 'unclear' in 38 studies, as most papers did not clarify whether the cytologists were given the routine information required for the reporting of a Pap smear (last menstrual period, relevant smear history, age etc). The thirteenth criterion (were un interpretable results reported?) was answered as 'no' or 'unclear' in 11 studies where there was no mention on the numbers of inadequate smears and invalid HPV assays Figure 2. There were 23 studies with two or fewer items answered 'no' or 'unclear' (Agorastos 2005; Agorastos 2015; Belinson 2003; Cardenas-Turanzas 2008; Castle 2011a; Cuzick 1995; Cuzick 1999; Cuzick 2003; de Cremoux 2003; Ferreccio 2013; Gravitt 2010; Iftner 2015; Mahmud 2012; McAdam 2010a; Monsonego 2011; Kulasingam 2002; Li 2009; Moy 2010; Pan 2003; Petry 2003; Qiao 2008; Sankaranarayanan 2004a; Schneider 2000). On the other hand there seven studies with more than three items answered 'no' or unclear (Clavel 2001; Depuydt 2011; Hovland 2010; Labani 2014; McAdam 2010b; Sarian 2005; Syrjanen 2002) (Figure 3). Overall the impression of the reviewers was that there was an adequate number of good quality studies for the completion of this meta-analysis. Two issues that reduced the quality of many studies were the issue of verification bias (QUADAS 5) and the issue of the blinding of the reference standard. The first was addressed in a sensitivity analysis

Figure 2. Methodological quality graph: review authors' judgements about each methodological quality item presented as percentages across all included studies

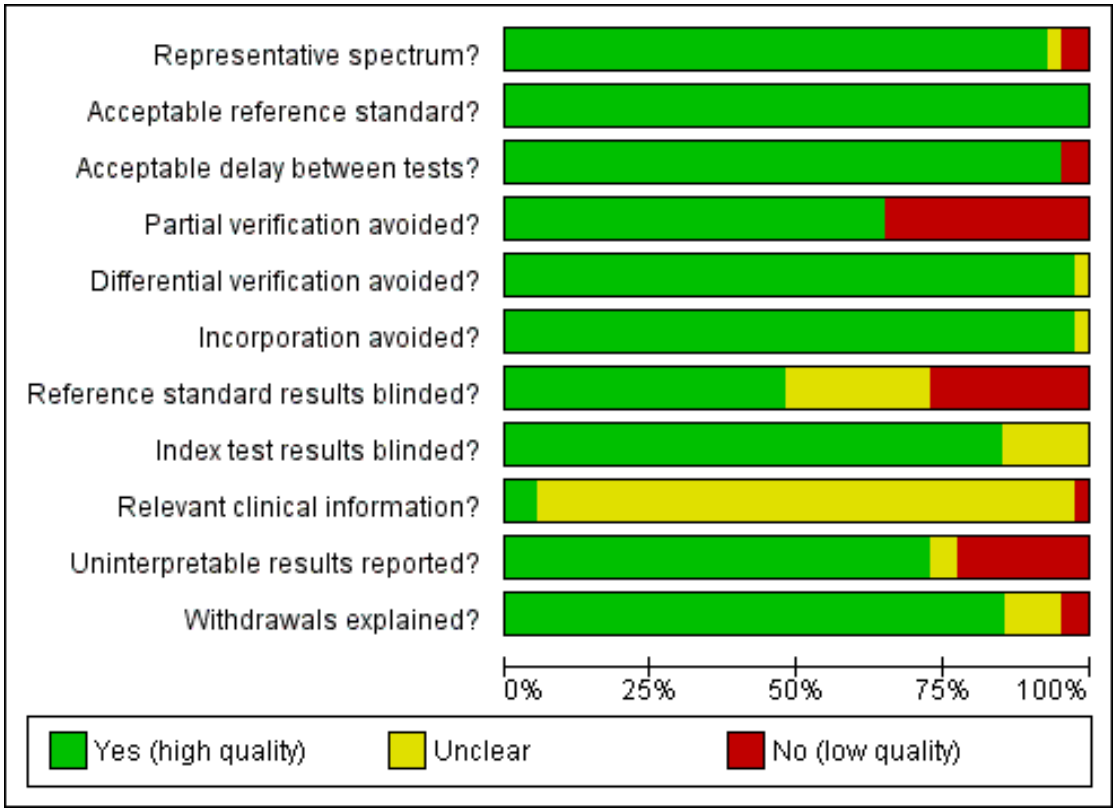
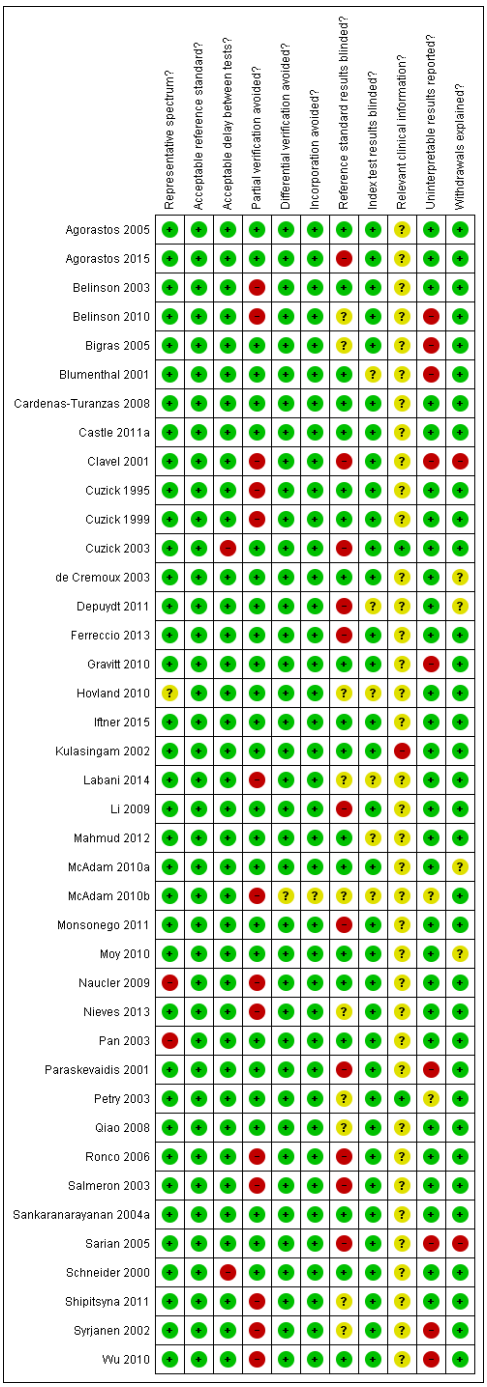


Figure 3. Methodological quality summary: review authors' judgements about each methodological quality item for each included study



Findings

Cervical cytology

Conventional cytology (CC) at the threshold of ASCUS+ for the detection of CIN 2+ and CIN 3+

There were 16 cross-sectional studies assessing CC for the detection of CIN 2+ (Data table 1) with 61,099 participants. Nine studies were conducted in Europe, two in Africa, two in Asia, one in North America, and two in Central and South America. Seven studies were undertaken in a population aged strictly over 30 years. The median sample size was 2256 (range 305 to 10,358) and the median prevalence of CIN 2+ was 1.66% (range 0.3% to 4.9%). The earliest study was published in 1995, with the majority published between 2003 and 2010.

There were nine cross-sectional studies assessing CC using the threshold of CIN 3+ (Data table 2) with 51,857 participants. Four studies were conducted in Europe, two in Asia, one in Africa and two in Central and South America. Six studies were undertaken in a population aged strictly over 30 years. The median sample size was 6194 (range 1386 to 10,358) and the median prevalence of CIN 3+ was 0.8% (range 0.2% to 1.5%) The earliest study was published in 1999, with the majority published between 2003 and 2010.

Sensitivity of CC ranged from 43% to 96% (pooled 65.9% (95% CI 54.9 to 75.3)) for the outcome CIN 2+ and from 39% to 85% (pooled 70.3% (95% CI 57.9 to 80.3)) for the outcome CIN 3+. The specificity ranged from 86% to 98% (pooled 96.3% (95% CI 94.7 to 97.4)) for CIN 2+ and 85% to 98% (pooled 96.7% (95% CI 94.6 to 98.0)) for CIN 3+ Table 1.

CC at the threshold of LSIL+ for the detection of CIN 2+ and CIN 3+

There were nine cross-sectional studies assessing CC for the detection of CIN 2+ (Data table 3) with 41,494 overall participants. Four studies were conducted in Europe, three in Africa, one in Asia, and one in Central and South America. Three studies were undertaken in a population aged strictly over 30 years. The median sample size was 2199 (range 305 to 10,591) and the median prevalence of CIN 2+ was 2% (range 0.8% to 9.5%). The earliest study was published in 2001, with the majority published between 2004 and 2011.

There were five cross-sectional studies assessing CC using the threshold of CIN 3+ (Data table 4) with 35,648 overall participants. Two studies were conducted in Europe, one in Asia, one in Africa and one in Central and South America. Two studies were undertaken in a population aged strictly over 30 years. The median sample size was 10,138 (range 1386 to 10,591) and the median prevalence of CIN 3+ was 1% (range 0.7% to 1.5%). The earliest study was published in 2002, with the majority published between 2002 and 2005.

Sensitivities of the tests ranged from 18% to 89% (pooled 62.8%, 95% CI 46.8% to 76.5%) and 64% to 80% (pooled 74.4%, 95% CI 67.8% to 80.1%). Specificities ranged from 92% to 100% (pooled 97.7%, 95% CI 96.1% to 98.7%) and 95% to 98% (pooled 96.9%, 95 % CI 94.9% to 98.1%) for the detection of CIN 2+ and CIN 3+, respectively.

Liquid-based cytology (LBC) at the threshold of ASCUS+ for the detection of CIN 2+ and CIN 3+

There were 15 cross-sectional studies assessing LBC for the detection of CIN 2+ (Data table 5) with 82,003 overall participants. Seven studies were conducted in Europe, one in Africa, six in Asia, and one in North America. Five studies were undertaken in a population aged strictly over 30 years. The median sample size was 3843 (range 301 to 16,516) and the median prevalence of CIN 2+ was 2.3% (range 0.4% to 5%). The earliest study was published in 2001, with the majority published between 2006 and 2011.

There were 13 cross-sectional studies assessing LBC using the threshold of CIN 3+ (Data table 6) with 71,919 overall participants. Five studies were conducted in Europe, five in Asia, one in Central and South America and one in North America. Five studies were undertaken in a population aged strictly over 30 years. The median sample size was 3843 (range 979 to 16,516) and the median prevalence of CIN 3+ was 0.9% (range 0.2% to 3.5%). The earliest study was published in 2002, with the majority published between 2009 and 2011.

Sensitivities of the tests ranged from 52% to 94% (pooled 75.5%, 95% CI 66.6% to 82.7%) and 52% to 98% (pooled 76.0%, 95% CI 64.7% to 84.5%) for the detection of CIN 2+ and CIN 3+ respectively. Specificities ranged from 73% to 97% (pooled 91.9%, 95% CI 90.1% to 90.5%) for detection of CIN 2+ and from 73% to 97% (pooled 91.2%, 95% CI 90.1 to 90.5%) for CIN 3+.

LBC at the threshold of LSIL+ for the detection of CIN 2+ and CIN 3+

There were 10 cross-sectional studies assessing LBC for the detection of CIN 2+ (Data table 7) with 33,519 overall participants. Three studies were conducted in Europe, one in Africa, four in Asia, and two in Oceania and Pacific. Six studies were undertaken in a population aged strictly over 30 years. The median sample size was 2475 (range 301 to 9451) and the median prevalence of CIN 2+ was 3.6% (range 1% to 5.3%). The earliest study was published in 2003, with the majority published between 2009 and 2011.

There were five cross-sectional studies assessing LBC using the threshold of CIN 3+ (Data table 8) with 21,166 overall participants. Three studies were conducted in Europe, and two in Asia. Three studies were undertaken in a population aged strictly over 30 years. The median sample size was 2905 (range 1993 to 9451) and the median prevalence of CIN 3+ was 0.9% (range 0.4% to 2.2%). The earliest study was published in 2003, with the majority published between 2008 and 2011.

Sensitivities of the tests ranged from 42% to 87% (pooled 70.3%, 95% CI 59.7% to 79.1%) and 48% to 93% (pooled 71.9%, 95% CI 61.2% to 76%). Specificities ranged from 90% to 98% (pooled 96.2%, 95% CI 94.6% to 97.4%) and 92% to 98% (pooled 96.1%, 95% CI 93.5% to 97.6%) for the detection of CIN 2+ and CIN 3+, respectively.

HPV testing

Hybrid capture II (HC2) at the threshold of 1 pg/mL for the detection of CIN 2+ and CIN 3+

There were 25 cross-sectional studies assessing HC2 for the detection of CIN 2+ (Data table 9) with 138,230 overall participants. Nine studies were conducted in Europe, two in Africa, eight in Asia, one in North America, two in Oceania and Pacific, and three in Central and South America. Thirteen studies were undertaken in a population aged strictly over 30 years. The median sample size was 4195 (range 491 to 16,410) and the median prevalence of CIN 2+ was 1.8% (range 0.5 to 10.1%). The earliest study was published in 2001, with the majority published between 2001 and 2008.

There were 19 cross-sectional studies assessing HC2 for the detection of CIN 3+ (Data table 10) with 120,380 overall participants. Seven studies were conducted in Europe, seven in Asia, one in Africa and four in Central and South America. Nine studies were undertaken in a population aged strictly over 30 years. The median sample size was 4429 (range 1352 to 16,410) and the median prevalence of CIN 3+ was 0.8% (range 0.2% to 2.2%). The earliest study was published in 2002, with the majority published between 2002 and 2008.

Sensitivities of the tests ranged from 61% to 100% (pooled 92.6%, 95% CI 89.6% to 95.3%) and 81% to 100% (pooled 96.5%, 95% CI 94% to 97.9%). Specificities ranged from 64% to 95% (pooled 89.3%, 95% CI 87% to 91.2%) and 69% to 95% (pooled 89.2%, 95% CI 86.7% to 91.3%) for the detection of CIN 2+ and CIN 3+, respectively.

HC2 at the threshold of 2 pg/mL for the detection of CIN 2+ and CIN 3+

Only two cross-sectional studies reported diagnostic data on HC2 at the threshold of 2 pg/mL for the detection of CIN 2+ (Data table 11) and CIN 3+ (Data table 12) with 26,768 overall participants. Sensitivity was 96% in both studies for the detection of CIN 2+ and specificity was similar at 94% and 95%. The sensitivity and specificity when HC2 was assessed at the threshold of 2 pg/mL for the detection of CIN 3+ was nearly identical to the test for detection of CIN 2+ (sensitivity was 95% and 96% in the two studies and specificity was 94% and 95%).

Polymerase chain reaction (PCR) for 13 high-risk types or more (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68) for the detection of CIN 2+ and CIN 3+

There were six cross-sectional studies assessing PCR for 13 high-risk types for the detection of CIN 2+ (Data table 13) with 16,343 overall participants. Four studies were conducted in Europe, one in Africa, and one in Asia. Two studies were undertaken in a population aged strictly over 30 years. The median sample size was 2100 (range 313 to 6089) and the median prevalence of CIN 2+ was 2% (range 0.3% to 5%). The earliest study was published in 2000, with the majority published between 2009 and 2011.

There were four cross-sectional studies assessing PCR for 13 high-risk types for the detection of CIN 3+ (Data table 14) with 14,048 overall participants. Two studies were conducted in Europe, one in Asia, and one in North America. Two studies were undertaken in a population aged strictly over 30 years. The median sample size was 3490 (range 979 to 6089) and the median prevalence of CIN 3+ was 1.3% (range 0.8% to 3.3%). The earliest study was published in 2002, with the majority published between 2009 and 2011.

Sensitivities of the tests ranged from 75% to 100% and 88% to 100%, specificities from 85% to 97% and 79% to 94% for the detection of CIN 2+ and CIN 3+, respectively.

PCR for 10 to 11 high-risk types for the detection of CIN 2+ and CIN 3+

Only two cross-sectional studies reported diagnostic data on PCR for 10 to 11 high-risk types for the detection of CIN 2+ (Data table 15) with 3964 overall participants, and just one study for CIN 3+ (Data table 16). Sensitivity was 74% and 89% in both studies for the detection of CIN 2+ and specificity was 95% and 79%. The sensitivity and specificity for the detection of CIN 3+ was 79% and 95%, respectively.

Aptima (HPV E6/7 mRNA testing) for the detection of CIN 2+ and CIN 3+

Three cross-sectional studies reported diagnostic data on Aptima for the detection of CIN 2+ (Data table 17) with 15,895 overall participants and four reported data on CIN 3+ (Data table 18) with 17,944 overall participants. Sensitivity range was 91% to 100% (pooled 92.7%, 95% CI 31.7% to 99.7%) for the detection of CIN 2+ and 93% to 100% (pooled 96%, 95% CI 72.9% to 99.5%) for the detection of CIN 3+. Specificity range for CIN

2+ was 91% to 97% (pooled 93.3%, 95% CI 47.3% to 99.5%) and for CIN 3+ 90% to 96% (pooled 92.8%, 95% CI 86.2% to 96.3%).

Cobas HPV test

Two cross-sectional studies reported diagnostic data on Cobas for the detection of CIN 2+ Data table 24 and CIN 3+ Data table 25 with 11,666 overall participants. Sensitivity range for CIN 2+ was 88% to 100% and 92% to 100% for CIN 3+. Specificity range was 58% to 90% for CIN 2+ and 57% to 90% for CIN 3+.

Other tests

Only single studies reported diagnostic data on the following tests; PCR for four high-risk types for CIN 2+ (Data table 19), care HPV test (0.5 pg/mL) for CIN 2+ (Data table 20) and CIN 3+ (Data table 21)), care HPV test (1 pg/mL) for CIN 2+ (Data table 22) and CIN 3+ (Data table 23), NASBA (five types (Data table 26) and nine types (Data table 27)) for CIN 2+ and HC2+4 (1 pg/mL) for CIN 2+ (Data table 28) and CIN 3+ (Data table 29). Of these single study tests, sensitivity was lowest (81%) in the NASBA (five types) for CIN 2+ and highest (94%) in the NASBA (nine types) for CIN 2+. Similarly, specificity was lowest (83%) in the care HPV test (0.5 pg/mL) for CIN 3+ and highest (97%) in the NASBA (five types) for CIN 2+.

Comparisons between cervical cytology and HPV testing for detection of CIN 2+ and CIN 3+

Comparisons could not be made for all tests, as the number of studies evaluating some of the test types were inadequate to provide stable ROC estimates (mainly analyses which included fewer than four studies). HPV testing for all or most high-risk HPV types such as HC2 or certain PCR assays had higher pooled sensitivity for CIN 2+ or CIN 3+ than CC or LBC at any threshold (ASCUS or LSIL). The pooled sensitivity of LBC was higher than CC. Conversely HPV testing had lower pooled specificity than cytology at any threshold (ASCUS or LSIL), with the difference being more evident with CC rather than LBC. We did not compare tests when there were fewer than two studies presenting paired data Table 2. When restricting the analysis only for studies with a population strictly over age 30, HC2 had slightly improved sensitivity and specificity for CIN 2+.

CC at cut-off ASCUS versus HC2

There were nine studies comparing conventional cytology (ASCUS+) to HC2 (1 pg/mL) for the detection of CIN 2+ (Figure 4) and six studies for CIN 3+ (Figure 5). Only the Cuzick 2003 study examined the accuracy of conventional cytology (ASCUS+) versus HC2 (2 pg/mL) for detection of both CIN 2+ and CIN 3+.

Figure 4. Summary ROC plot of 2 tests for detection of CIN 2+ (verified with histology): Conventional Cytology (ASCUS+) and HPV testing with hybrid capture 2 (1pg/mL). The black and red solid circles correspond to the summary estimates of sensitivity and specificity, and are shown with a 95% confidence region.

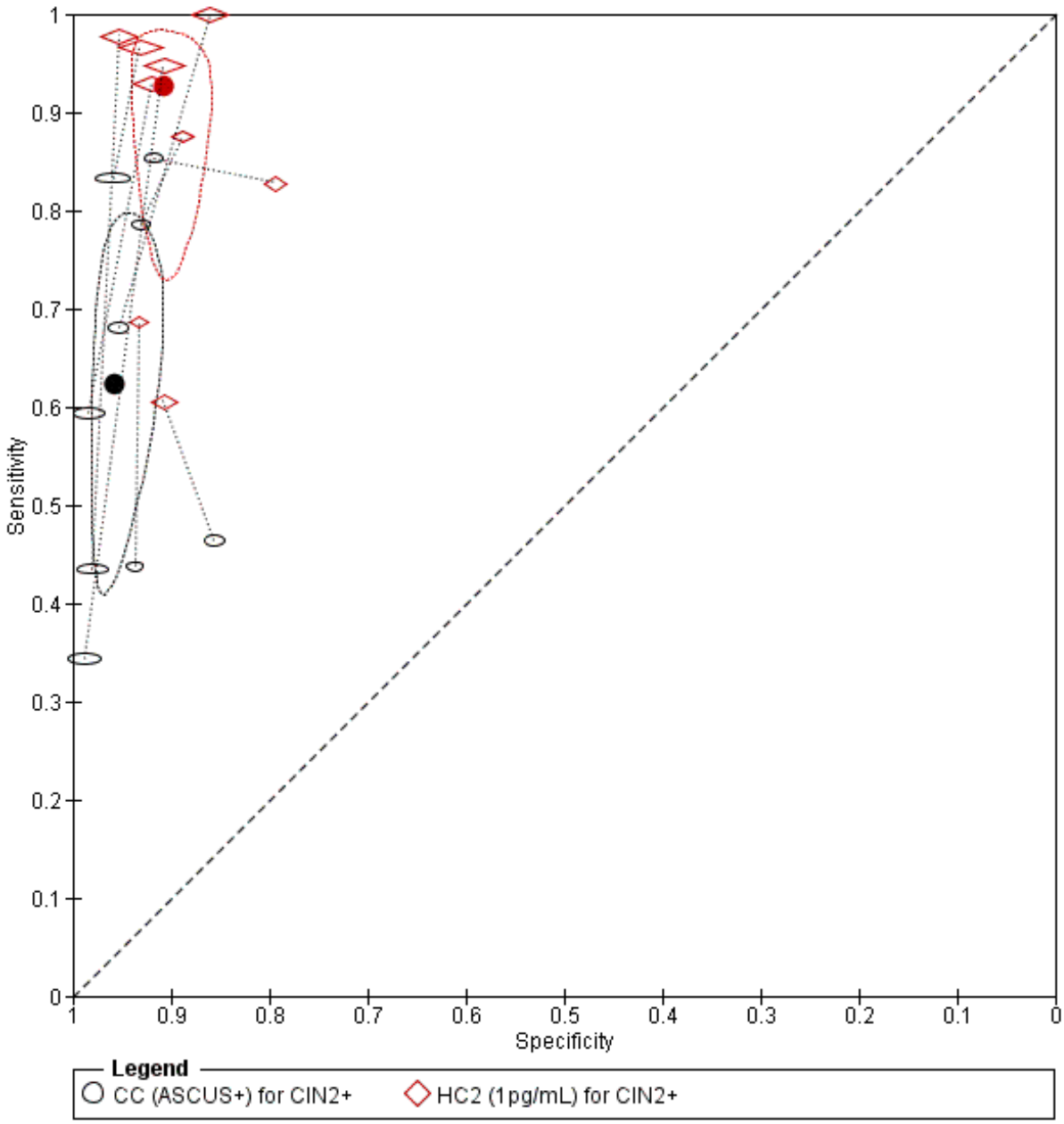
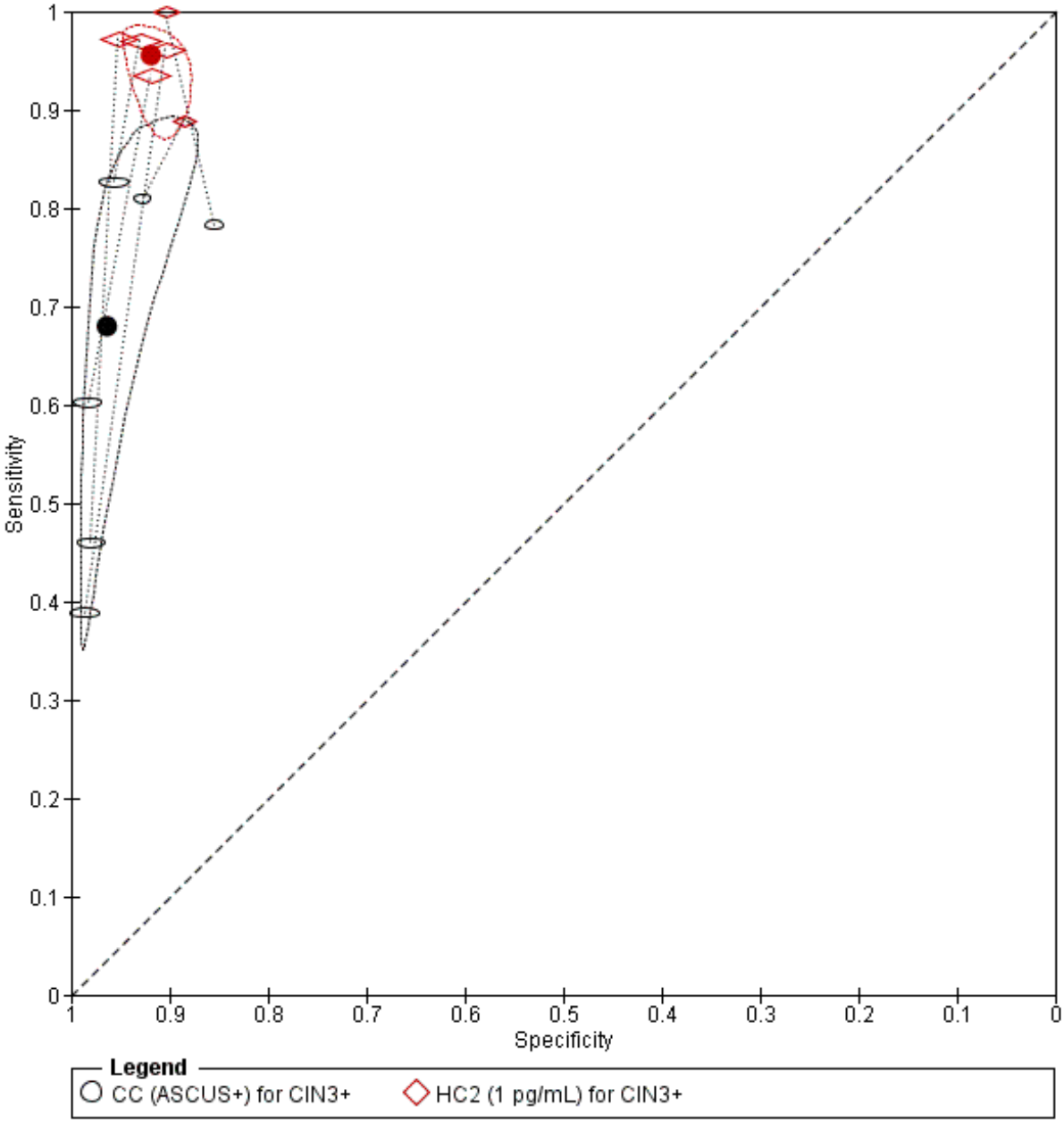


Figure 5. Summary ROC plot of 2 tests for detection of CIN 3+ (verified with histology): Conventional Cytology (CC) (ASCUS+) and HPV testing with hybrid capture (HC) 2 (1pg/mL). The black and red solid circles correspond to the summary estimates of sensitivity and specificity, and are shown with a 95% confidence region.



The HC2 test at the 1 pg/mL threshold appeared to be a better test than CC at the threshold of ASCUS (for CIN 2+) in terms of summary (S)ROC curve, and the meta-analytic sensitivity was considerably lower than in the HC2 tests. However, the specificity was slightly higher in the CC ASCUS test compared to the HC2 tests.

The relative sensitivity of HC2 versus CC for CIN 2+ was 1.52 (95% CI: 1.24 to 1.86) and the relative specificity 0.94 (95% CI 0.92 to 0.96).The relative sensitivity of HC2 versus CC for CIN 3+ was 1.46 (95% CI 1.12 to 1.91) and the relative specificity 0.95 (95% CI 0.93 to 0.97).

CC ASCUS versus PCR (for more than 12 high-risk types)

There were three studies comparing CC (ASCUS+) to PCR for the detection of CIN 2+ (Figure 5) and just one (Naucler 2009) for CIN 3+.

The PCR SROC curve for detection of CIN 2+ appeared to indicate a better test than CC (ASCUS+) but this was only based on three studies that offered paired data. The meta-analytic sensitivity and specificity were reasonably high and the PCR test seemed to have better overall discrimination than the CC test. The specificity in the CC test was very high but sensitivity was too low to make this test acceptable based on the limited data available.

The relative sensitivity of PCR (more than 12 types) versus CC for CIN 2+ was 1.37 (95% CI 0.58 to 3.21) and the relative specificity was 0.95 (95% CI 0.76 to 1.19). The relative sensitivity of PCR (more than 12 types) versus CC for CIN 3+ was 1.30 (95% CI 1.09 to 1.54) and the relative specificity was 0.95 (95% CI 0.94 to 0.96).

CC LSIL versus HC2

There were six studies comparing conventional cytology (LSIL+) to HC2 (1 pg/mL) for the detection of CIN 2+ (Figure 6) and five studies for CIN 3+ (Figure 7). Only the Cuzick 2003 study examined the accuracy of CC (LSIL+) versus HC2 (2 pg/mL) for detection of CIN 2+. There were no studies reporting data with CIN 3+ outcome for this test comparison.

Figure 6. Summary ROC plot of 2 tests for detection of CIN 2+ (verified with histology): Conventional Cytology (CC) (LSIL+) and HPV testing with hybrid capture (HC) 2 (1pg/mL). The black and red solid circles correspond to the summary estimates of sensitivity and specificity, and are shown with a 95% confidence region.

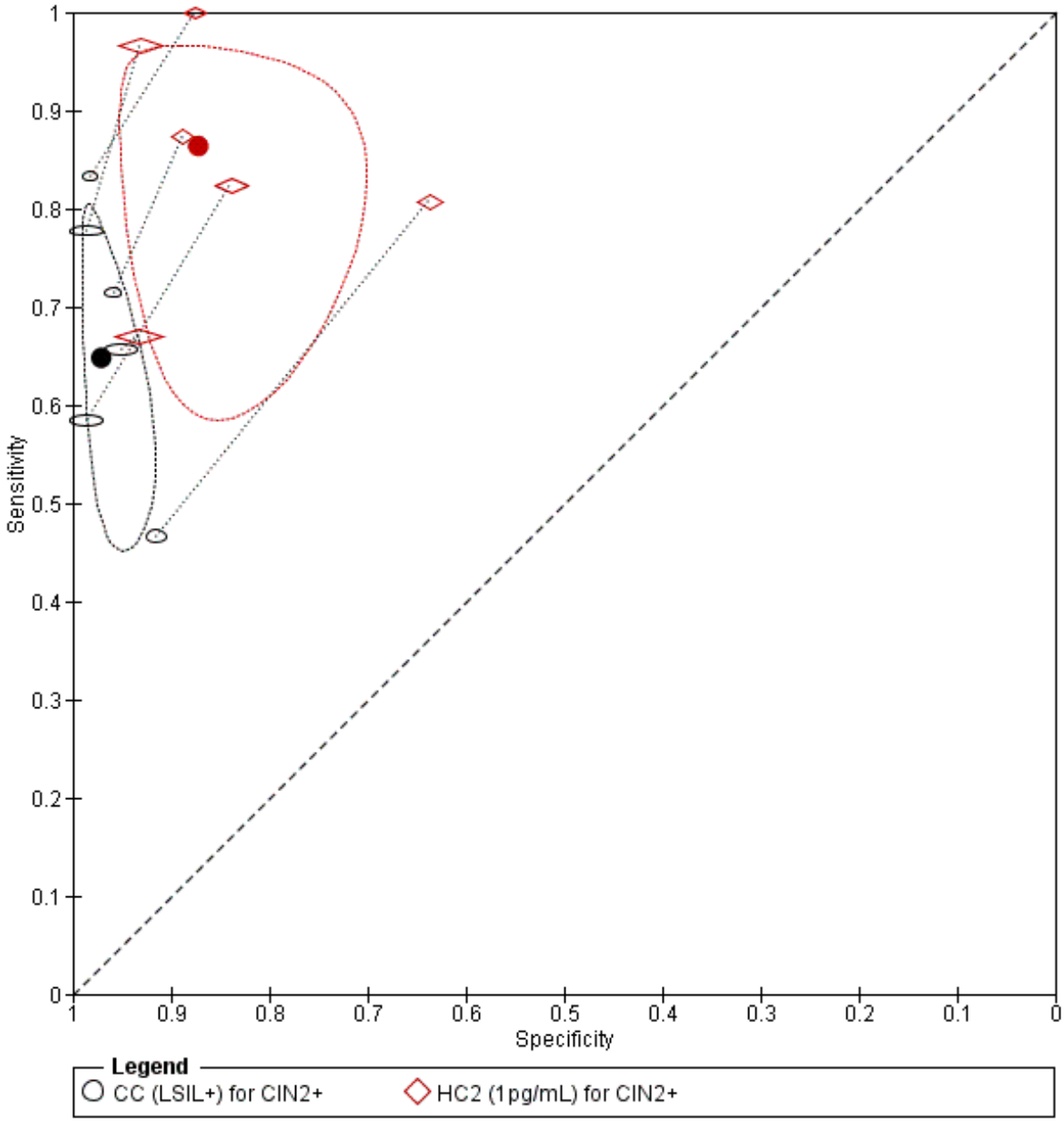
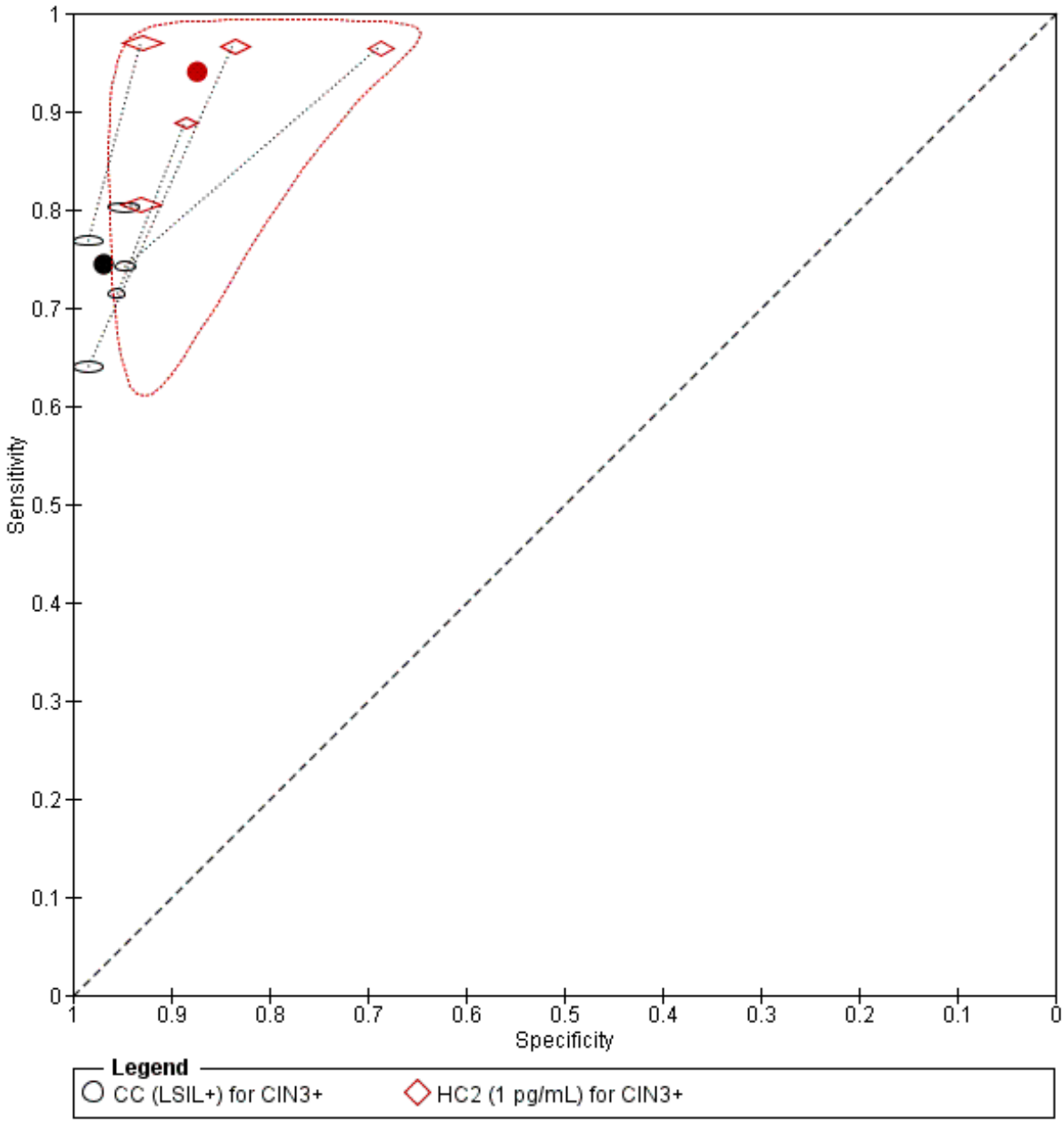


Figure 7. Summary ROC plot of 2 tests for detection of CIN 3+ (verified with histology): Conventional Cytology (CC) (LSIL+) and HPV testing with hybrid capture (HC) (1pg/mL). The black and red solid circles correspond to the summary estimates of sensitivity and specificity, and are shown with a 95% confidence region.



The SROC curves show that both tests are not accurate at correctly classifying women with and without the disease. Although the meta-analytic sensitivity was higher in the HC2 tests compared to the CC test, the specificity was considerably lower. Specificity was very high in the CC test but sensitivity was not at an acceptable level.

The relative sensitivity of HC2 versus CC for CIN 2+ was 1.28 (95% CI 1.15 to 1.41) and the relative specificity was 0.91 (95% CI 0.87 to 0.95).The relative sensitivity of HC2 versus CC for CIN 3+ was 1.22 (95% CI 1.12 to 1.32) and the relative specificity was 0.91 (95% CI 0.87 to 0.95).

CC LSIL versus PCR (for more than 12 high-risk types)

There were two studies comparing CC (LSIL+) to PCR for the detection of CIN 2+ and none for CIN 3+. From the SROC the PCR test seemed far superior at detecting CIN 2+ compared to the CC test, but this was only based on two studies so it is difficult to draw any conclusions.

LBC ASCUS versus HC2

There were 10 studies comparing LBC (ASCUS+) to HC2 (1 pg/mL) for the detection of CIN 2+ (Figure 8) and seven studies for CIN 3+ (Figure 9). Only the Ronco 2006 study examined the accuracy of LBC (ASCUS+) versus HC2 (2 pg/mL) for detection of both CIN 2+ and CIN 3+.

Figure 8. Summary ROC plot of 2 tests for detection of CIN 2+ (verified with histology): Liquid Based Cytology (LBC) (ASCUS+) and HPV testing with hybrid capture (HC) 2 (1pg/mL). The black and red solid circles correspond to the summary estimates of sensitivity and specificity, and are shown with a 95% confidence region.

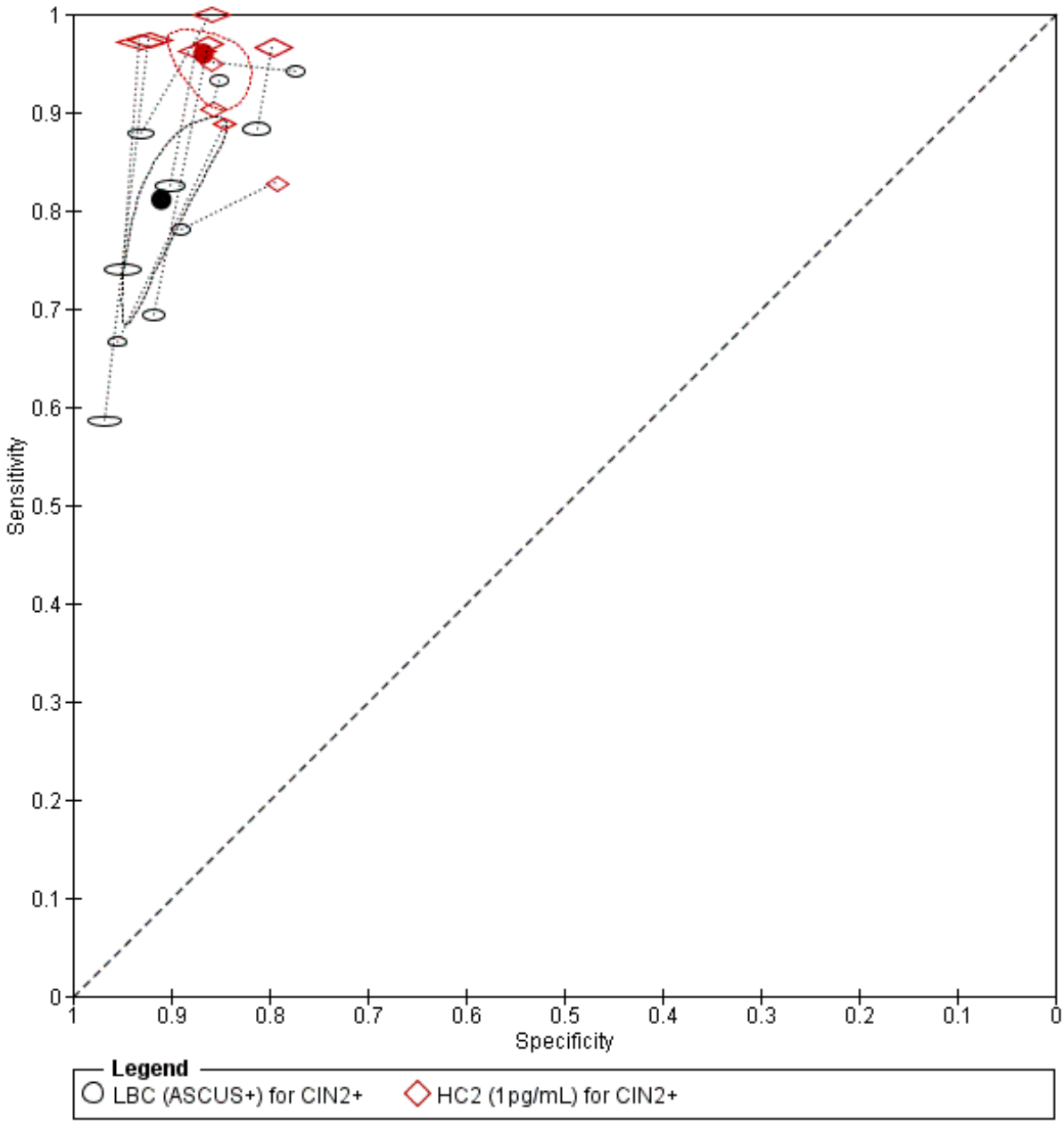
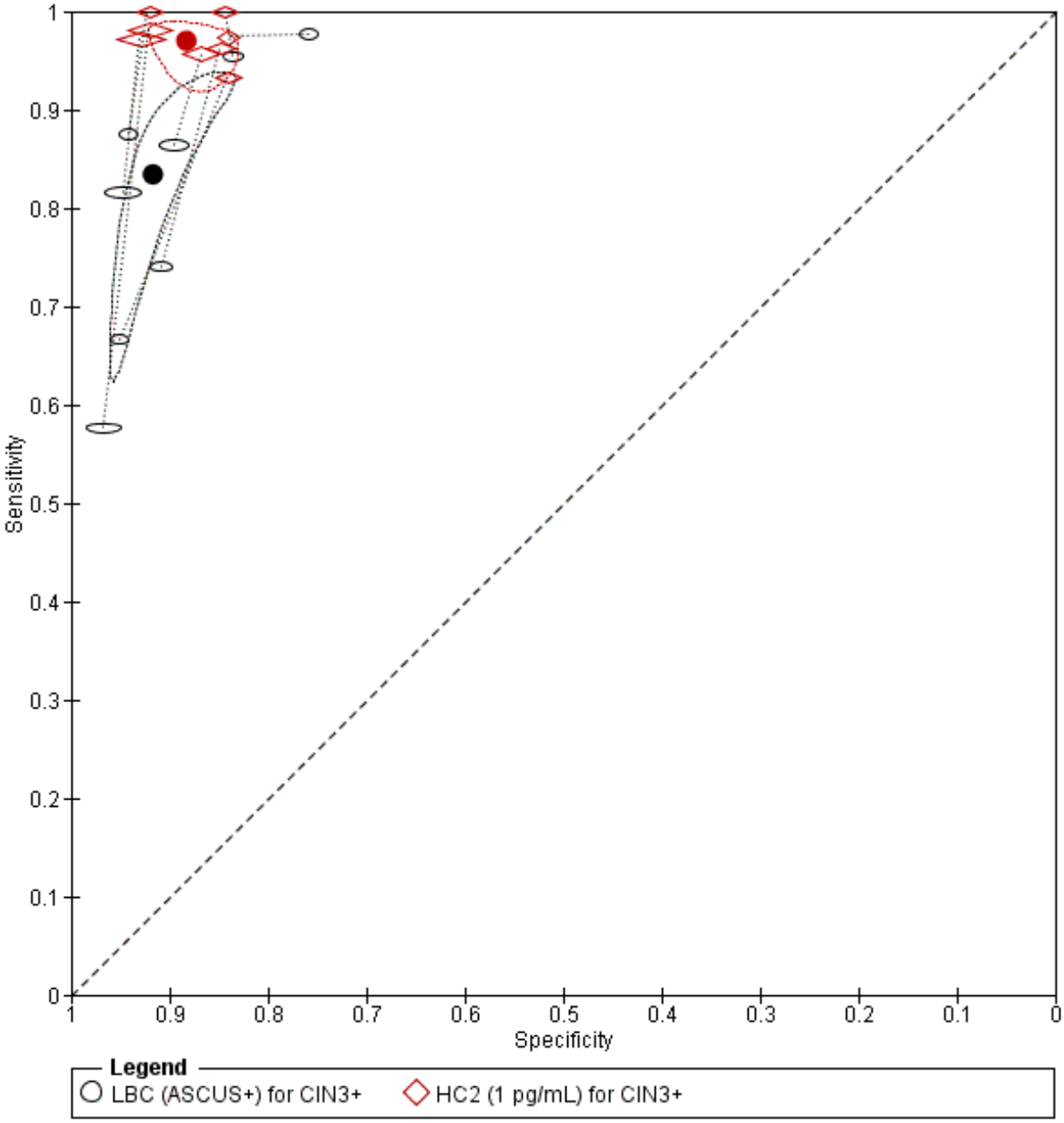


Figure 9. Summary ROC plot of 2 tests for detection of CIN 3+ (verified with histology): Liquid Based Cytology (LBC) (ASCUS+) and HPV testing with hybrid capture (HC) 2 (1pg/mL). The black and red solid circles correspond to the summary estimates of sensitivity and specificity, and are shown with a 95% confidence region.

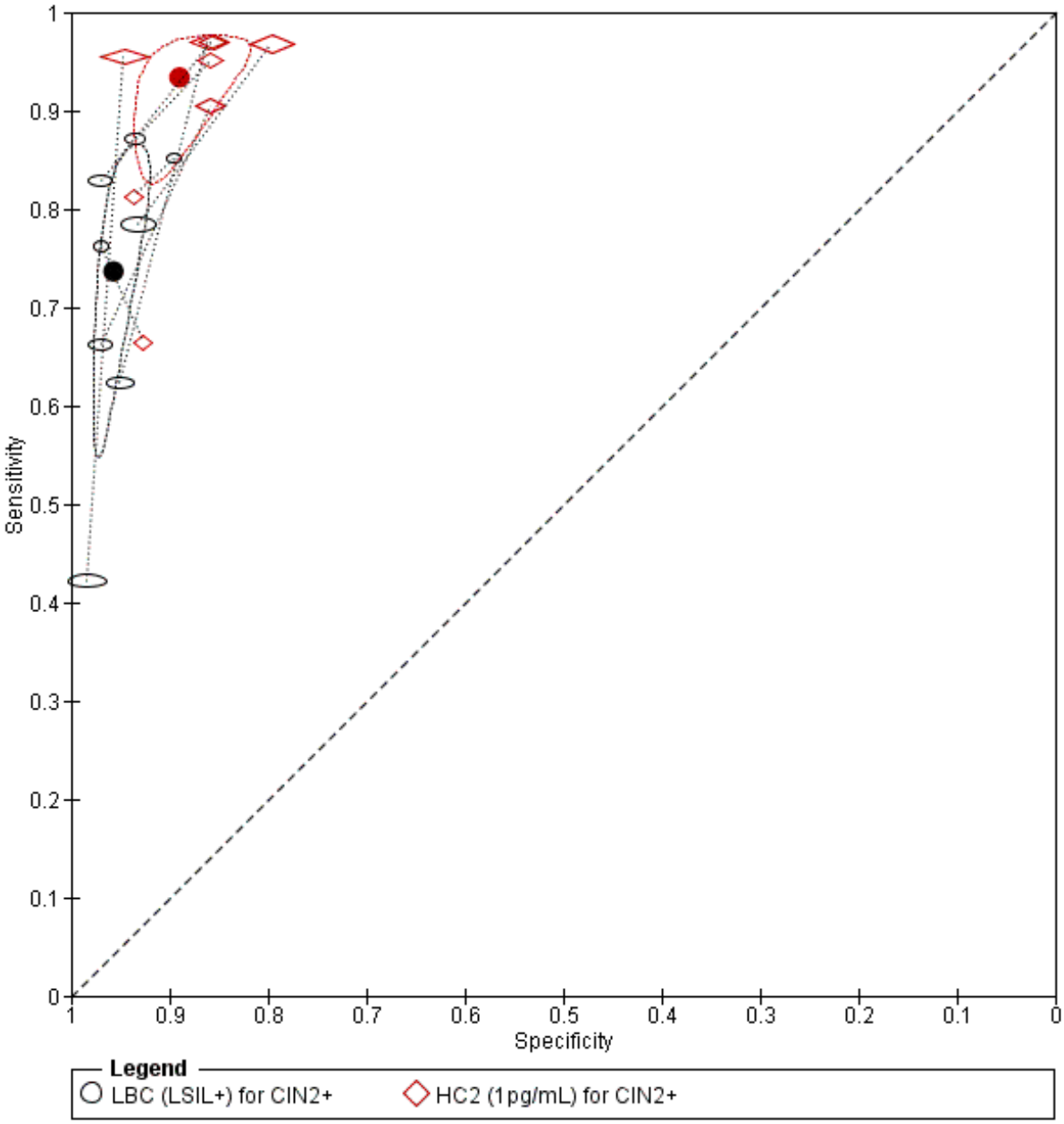


The HC2 SROC curves at all thresholds appeared to represent good tests, whereas the LBC test at the threshold of ASCUS (for both CIN 2+ and CIN 3+) appeared inferior. The meta-analytic sensitivity was considerably lower than in the HC2 tests whereas the specificity is not much lower in the HC2 test compared to LBC. The HC2 tests appeared to have better overall discrimination.
The relative sensitivity of HC2 versus LBC for CIN 2+ was 1.18 (95% CI 1.10 to 1.26) and the relative specificity was 0.96 (95% CI 0.95 to 0.97). The relative sensitivity of HC2 versus LBC for CIN 3+ was 1.17 (95% CI 1.07 to 1.28) and the relative specificity was 0.96 (95% CI 0.95 to 0.97).

LBC ASCUS versus PCR (for more than 12 high-risk types)

There were three studies comparing LBC (ASCUS+) to PCR for the detection of CIN 2+ (Figure 10) and three for CIN 3+.

Figure 10. Summary ROC plot of 2 tests for detection of CIN 2+ (verified with histology): Liquid Based Cytology (LBC) (LSIL+) and HPV testing by hybrid capture (HC) 2 (1pg/mL). The black and red solid circles correspond to the summary estimates of sensitivity and specificity, and are shown with a 95% confidence region.



From the SROC the PCR test seemed superior at detecting CIN 2+ compared to the LBC test, but this was only based on two studies so it is difficult to draw any conclusions. The meta-analytic sensitivity was very high for detection of CIN 2+ and CIN 3+ but specificity was much lower, based on these limited data. Sensitivity in the LBC test was very low.

The relative sensitivity of PCR (more than 12 types) versus LBC for CIN 2+ was 1.53 (95% CI 0.53 to 4.44) and the relative specificity was 0.90 (95% CI 0.82 to 0.99).The relative sensitivity of PCR (more than 12 types) versus LBC for CIN 3+ was 1.47 (95% CI 0.64 to 3.35) and the relative specificity was 0.94 (95% CI 0.80 to 1.09).

LBC LSIL versus HC2

There were eight studies comparing LBC (LSIL+) to HC2 (1 pg/mL) for the detection of CIN 2+ (Figure 9) and four studies for CIN 3+. No study examined the accuracy of LBC (LSIL+) versus HC2 (2 pg/mL) for detection of CIN 2+ or CIN 3+.

The HC2 SROC curves at all thresholds appeared very sensitive, but specificity was lower. The LBC tests at the threshold of LSIL (for both CIN 2+ and CIN 3+) appeared superior in specificity but sensitivity was too low. The HC2 test appeared to have better overall discrimination.

The relative sensitivity of HC2 versus LBC for CIN 2+ was 1.35 (95% CI 1.19 to 1.53) and the relative specificity was 0.92 (95% CI 0.89 to 0.95).The relative sensitivity of HC2 versus LBC for CIN 3+ was 1.30 (95% CI 0.49 to 1.96) and the relative specificity was 0.92 (95% CI 0.84 to 1.00).

LBC LSIL versus PCR (for more than 12 high-risk types)

There were two studies comparing LBC (LSIL+) to PCR for the detection of CIN 2+ and one (Depuydt 2011) for CIN 3+.

From the SROC the PCR test seemed superior at detecting CIN 2+ compared to the LBC test, but this was only based on two studies so it is difficult to draw any conclusions.

LBC ASCUS versus APTIMA

There were three studies comparing LBC (ASCUS+) to APTIMA for the detection of CIN 3+.

The APTIMA test appeared to have superior sensitivity to LBC with similar specificity. The relative sensitivity of APTIMA versus LBC for CIN 3+ was 1.30 (95% CI 0.49 to 3.41) and the relative specificity was 0.98 (95% CI 0.93 to 1.04).

Investigations of heterogeneity

The influence of factors on the accuracy of HC2 (defined at cut-off 1 pg/mL) for CIN 2+, and CIN 3+, assessed by bivariate random-effects meta-analyses with one covariate each time is shown in Table 3. The sensitivity was significantly higher in studies enrolling women older than 30 years than in studies enrolling women of any age: relative sensitivity of 1.13 (95% CI 1.03 to 1.25) and 1.10 (95% CI 1.02 to 1.19) for outcomes CIN 2+ and CIN 3+, respectively. The specificity was also higher in women older than 30 years, but the difference was only significant for the outcome CIN 3+ (relative specificity of 1.04, 95% CI 1.00 to 1.08.

The sensitivity was higher in studies with high versus low risk for verification bias, but the difference was only significant for the outcome of CIN 3+ (relative sensitivity of 1.09, 95% CI 1.01 to 1.18). The specificity estimates were not affected by risk of verification bias.

To assess geographical effect, locations were recorded as high-income (North-America, Australia/New Zealnd, Europe) or middle- and low-income (other countries). We could not identify any significant effects on accuracy estimates.

The effect of the number of HPV types targeted by the HPV assay could not be assessed for HC2, since this test always detects 13 high-risk types. Also for other assays, insufficient data were available to assess the effect of the number of types by BRMA analyses.

Sensitivity analysis

The pooled sensitivity and specificity of HC2 at threshold of 1 pg/mL for CIN 2+ was 89.0% (95% CI 81.1% to 93.9%) and 88.6% (95% CI 84.2% to 91.9%), respectively, when the meta-analysis included only the 11 studies where the reference standard was used on all women (Data table 28). The sensitivity and specificity of CC or LBC at cut-off ASCUS+ for CIN 2+, pooled from nine studies (Data table 29), where all women were submitted to the reference standard, was 72.2% (95% CI 57.5% to 83.3%) and 93.6% (95% CI 88.9% to 96.4%).

When the meta-analysis was restricted to 13 studies where only women of age 30 or older were enrolled (Data table 30), the pooled sensitivity of HC2 at 1 pg/mL for CIN 2+ was 93.9% (95% CI 89.3% to 96.6%), whereas the specificity for CIN 1 or below at the same cut-off was 91.3% (95% CI 88.9% to 93.2%).

Influence of the number of types targeted by HPV assays is included in the results by HPV test (see above). HC2 always included 13 high-risk types, APTIMA and Cobas 4800 always included 14 types. Several distinct PCR-based assays were used targeting different high-risk HPV types: four high-risk HPV types in one study (Cuzick 1995), 10-11 high-risk HPV types in two studies (Cuzick 1999, Paraskevaidis 2001) and 13 or more HPV types in six studies (Data table 13). All these PCR systems were distinct assays, so the effect of the choice and number of high-risk HPV types could not be assessed separately from the test platform. Only

for the NASBA test did we retrieve one study where the same assay targeted five types (HPV16, 18, 31, 33 and 45) or nine types (the same five types plus HPV35, 51, 52 and 58). The sensitivities were 13/16 (81%) and 15/16 (94%) and the specificities were 287/297 (97%) and 279/297 (84%), respectively.

Summary of findings

Human papillomavirus (HPV) compared to Papanicolaou (Pap) test for detection of cervical intraepithelial neoplasia (CIN 2+) in asymptomatic women					
Patient or population: adult asymptomatic women **Settings:** outpatient screening programmes **New Test:** HPV, HC2 test **Cut-off value:** 1 pg/mL **Comparison Test:** Pap, liquid-based cytology (LBC) test **Cut-off value:** atypical squamous cells of undetermined significance (ASCUS) **Reference Test:** a colposcopy exam with or without biopsy as clinically indicated					
HPV	138,230 women (25 studies)	**Pooled sensitivity (95% CI)**	89.9% (88.6 to 91.1%)	**Pooled specificity (95% CI)**	89.9% (89.7 to 90.0%)
Pap	82,003 women (15 studies)	**Pooled sensitivity (95% CI)**	72.9% (70.7 to 75%)	**Pooled specificity (95% CI)**	90.3% (90.1 to 90.5%)
Test results	**Number of results per 1000 women tested (95% CI)**		**Quality of the evidence (GRADE)**	**Comments**	
	Prevalence of CIN 2+, 2%[1]				
	HPV	**Pap**			
True positives (TP)	**18** (17 to 19)	**14** (13 to 15)	⊕⊕⊕○ **moderate** due to inconsistency[2,3]	Women will be correctly classified and will receive further confirmatory testing or treatment	
TP absolute difference	**4 more**				
False negatives (FN)	**2** (1 to 3)	**6** (5 to 7)		Women will be falsely reassured that they do not have CIN 2+, and the potentially beneficial treatment may be missed or will be delayed	
FN absolute difference	**4 fewer**				
True negatives (TN)	**981**	**985**	⊕⊕⊕○	Women will be correctly reassured that they do not have	
TN absolute difference	**4 fewer**				

False positives (FP)	**99** (98 to 101)	**95** (93 to 97)		Women will likely receive unnecessary further testing and possibly also unnecessary treatment; additionally further
FP absolute difference	**4 more**			

CI: Confidence interval; **HPV** human papillomavirus; Pap: Papanicolaou test, **CIN**: cervical intraepithelial neoplasia

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Prevalence of 2% (20 women out of 1000) was assumed to be the average prevalence of cervical intraepithelial neoplasia 2+ in non HIV asymptomatic women.
[2]Serious inconsistency in sensitivity among studies with sensitivity ranging from 52%-94% for Pap, and 61% to 100% for HPV.
[3]We did not downgrade for risk of bias, but the few limitations with studies were considered with inconsistency.

DISCUSSION

Summary of main results

We evaluated the accuracy of two cervical cancer screening methods: cervical cytology and HPV testing in a large body of clinical studies. We focused mainly on the sensitivity and the specificity of the tests. Predictive values depend on the local disease prevalence, and therefore generalisation of the results is usually of limited value.

The results show that the HC2 and the PCR (for more than 12 HPV types) have higher sensitivity than cytology even at the lowest cytological positivity threshold of ASCUS, showing that these HPV tests are less likely to miss CIN 2+ (and CIN 3+) than cytological tests. The specificity on the other hand is significantly higher for cytology at the threshold of LSIL than either for HC2 or PCR. Whilst the predictive value of a negative HPV test approached 100%, HPV tests are associated with more unnecessary referrals (for false positives) than cytological tests. The accuracy of the test depends on how well the test separates the group being tested into those with and without the disease in question. The HC2 tests appeared to have better overall discrimination. The same conclusion can be extrapolated for other methods of DNA testing for all or most high-risk HPV types by PCR although there were fewer studies available for robust conclusions. The four studies that used E6 and E7 mRNA detection with APTIMA showed higher sensitivity and equivalent specificity to cytology.

There was large inter-study variation in the sensitivity and specificity estimations of all screening tests. For cytology, this was partly expected because the reproducibility of the cytological interpretation of smears is often problematic (Stoler 2001). The sensitivity of cytology is exceptionally low in the German studies (Petry 2003; Schneider 2000), which raises concerns regarding the sampling technique and cytologic interpretation in these studies. In one of these studies (Petry 2003), the instrument used for sampling was a cotton-tipped swab, which is not recommended, but this alone cannot explain the very low sensitivity of cytology. For PCR, the inter-study variation could be explained by the different primers used and other variations in the technique between laboratories. For the HC2 method, however, the variation was surprising and was mainly caused by the low sensitivity of HPV testing, which was often observed in low- and middle-income countries (Blumenthal 2001; Gravitt 2010; McAdam 2010a; McAdam 2010b; Sankaranarayanan 2004a). This could be attributed to the variable quality of verification procedures (Arbyn 2008; Arbyn 2009). In one particular study classified as low risk of bias, CC and HC2 are shown to have almost equal sensitivity (Sankaranarayanan 2004a). Possible explanations for the low HC2 sensitivity in our study could be contamination of the sample by acetic acid or Lugol's iodine or deterioration of the sample because of exposure at high temperature. Contamination of the sample by acetic acid or Lugol's iodine could normally not have occurred, since, according to the protocol, the sample for HC2 was collected before application of vinegar or iodine solution. Finally, misclassification of the outcome could also explain the low observed sensitivity of HC2. The policy of random biopsies, which was employed in some studies, might have increased the detection rate of lesions, although the value of random biopsies is disputed (Wentzensen 2015). The low pooled specificity of HC2, which was observed in studies conducted in Africa, is mainly due to the outlying specificity of one study (Blumenthal 2001).

It has been proposed that the specificity of HPV testing is age-related and higher in older women, something that should be borne in mind when evaluating the cost-effectiveness of HPV test screening. This was not confirmed by this meta-analysis, which showed only a mild but not significant increase in sensitivity and specificity for the HC2 test in women over 30 years compared to the general population. However this particular analysis was not done on paired studies.

Strengths and weaknesses of the review

Strengths of the review

A relatively large number of studies fulfilling the inclusion criteria was identified. Many studies reported results for various cytologic, virologic and histologic thresholds. This enabled us to perform meta-analyses for many of these thresholds. The studies had a wide geographical distribution with all continents being represented in more than one study, with the exception of Oceania. Most of the studies were of good methodological quality according to the QUADAS criteria. Also the design of concomitant testing of the subjects with both tests limited the risk of selection bias.

Limitations of the review

In most studies, the presence or absence of disease was not verified with colposcopy and histology in all women, leading to potential verification bias. It is likely that false-negative results are missed for either test without adequate verification of test negatives. Theoretically this causes an overestimation of the sensitivity of the tests, but it should not affect the relative sensitivity (sensitivity ratio) or false-positivity rate. The pooled sensitivity of HC2 was higher in the high-risk-of-verification-bias group but surprisingly, the pooled sensitivity of cytology was higher in the low-risk group. This was mainly the effect of one study classified as low risk of bias where the two tests were shown to have almost equal sensitivity (Sankaranarayanan 2004a).

It is likely that the contrast between HPV testing and cytology was inflated by the inclusion of the two German studies, where cytology had very low sensitivity. In most studies colposcopists were aware of the screening test results, which could bias their

colposcopic diagnosis and their decision to take a biopsy or not. Finally, the use of colposcopy or even punch biopsies as a gold standard can also be sub-optimal as their performance is operator-dependent (Stoler 2001) and could have influenced the findings of each study and subsequently these meta-analyses.

The review included only studies in the English language. It was presumed that this would cover the vast majority of the existing studies on the subject. The thoroughness of the search would have otherwise been questionable. However, during the search process we did not identify any relevant studies in a non-English language that were excluded solely because of it.

Applicability of findings to the review question

Our study was restricted to cross-sectional outcomes such as sensitivity and specificity where the performance of one application of the screening tests is compared to a gold standard (colposcopy and histology). It is known that the precancerous lesions of the cervix take several years to progress to cancer. During this time, women are subjected to a number of cytological examinations. Therefore, even though it is likely, it cannot be argued that the superior cross-sectional sensitivity of HPV testing will certainly mean superiority within an actual cervical screening setting. In addition, since CIN2-3 is potentially regressive, it has not yet been shown that HPV screening does more than just finding more small-size lesions, which would clear without intervention. For this reason, high-risk HPV-based cervical cancer screening was not yet recommended in the second edition of the European Guidelines for Cervical Cancer Screening, considering evidence available in 2006 (Arbyn 2010).

In conclusion, this meta-analysis has shown clearly that HC2 has a superior sensitivity to CC and LBC. However, the improved cross-sectional accuracy does not guarantee a better performance in terms of reduction on the incidence of cervical cancer if the HPV test is implemented in primary screening. The longitudinal outcomes of the ongoing randomised studies should clarify this issue. In the meantime, data from the second screening round of RCTs, comparing cytology with HPV screening, have demonstrated a significant reduction of CIN 3+, and even of invasive cancer, among women in the first round who had a negative HPV test, compared to women in the control arm who had a negative Pap smear (Arbyn 2012; Ronco 2014).

AUTHORS' CONCLUSIONS

Implications for practice

For a screening test, a high sensitivity such as the one produced by human papillomavirus (HPV) testing is very important as it reduces false-negative results. Under the assumption that HPV also detects more progressive lesions not detectable by cytology, one may expect that HPV-based screening would result in a lower incidence of and mortality from cervical cancer. On the other hand, its lower specificity could have cost implications because of the referral of a large number of women with false-positive results to colposcopy (Mandelblatt 2002). Apart from producing more referrals for colposcopy, the hybrid capture 2 (HC2) test was, until recently, more expensive as a test than the cervical cytologic examination (Meera 2002). However, since recent years, the cost price of HPV assays has decreased dramatically. Massive centralised purchase of HPV tests could even make virological screening cheaper than cytological screening.

The increased false positive rate of HPV testing exposes women to unnecessary psychological morbidity (McCaffery 2004) and an increase in the referral rate for colposcopy. The number of colposcopy referrals could be limited by offering cytological triage of HPV-positive results, or reflex testing for HPV16 or 18 (Castle 2011b; Cuzick 2003; Dijkstra 2013; Rijkaart 2012a).The negative predictive value of HPV testing approaches 100% in most studies. It has been shown that the five-year disease-free rate following a negative HPV test is equivalent to the two-year disease-free rate following negative cytology (Kjaer 2004). This suggests that the use of HPV testing could allow the lengthening of screening intervals with subsequent reduction in costs. Although efficiency of cervical cancer screening may be optimised by switching to HPV-screening at longer intervals, greater gains might be obtained by increasing coverage, access to treatment and follow-up.

Based on the accuracy data from nine of the 35 cross-sectional studies included in our meta-analyses, and considering also longitudinal results from the Sherman 2003 study, the US Food and Drug Administration approved the use of a high-risk probe cocktail of HC2 as an adjunct to cervical cytology screening in women aged 30 years or more (Saslow 2012). In Europe, however, use of HPV tests was not recommended in the 2008 EU guidelines for quality assurance in cervical cancer screening (Arbyn 2010). In these guidelines, a possible switching to HPV-based screening was going to be proposed only when randomised controlled trials (RCTs) would demonstrate lower incidence of cervical intraepithelial neoplasia (CIN) 3+ in the second screening round in women screened with HPV; This new evidence is currently being translated in the updated EU guidelines recommending HPV-based screening as the primary screening test.

Implications for research

As the aim of the cervical screening programme is to reduce the incidence of and mortality from cervical cancer, the most desired outcome measure would be the effect of each screening test on incidence and mortality in the long term and detection rates of pre-invasive disease in the short term, together with other outcomes such as economic and psychological morbidity of awareness of HPV status. Longitudinal studies are required to examine whether

the relatively low sensitivity of cytology would be improved by the repeated cytological examinations, which would detect initially-missed lesions.

The role of HPV-related markers other than HPV DNA testing, such as HPV genotyping, E6 and E7 mRNA expression, E6 and E7 protein and p16 will have to be investigated in a similar manner. Considering that simultaneous HPV and Pap testing (co-testing) is used for primary screening in the USA and Canada, it would be useful to compare the accuracy of co-testing to HPV testing alone in another meta-analysis.

Another important issue is that most of the studies were performed before the introduction of the HPV vaccine. It will be interesting to study how the accuracy of the two tests compares in a widely-vaccinated population.

ACKNOWLEDGEMENTS

This project was supported by the National Institute for Health Research (NIHR), via Cochrane infrastructure funding to the Cochrane Gynaecological, Neuro-oncology and Orphan Cancers Group. The views and opinions expressed therein are those of the review authors and do not necessarily reflect those of the Systematic Reviews Programme, the NIHR, the NHS, or the Department of Health.

We thank Jo Morrison, Gail Quinn, Clare Jess and Jane Hayes of the Cochrane Gynaecological, Neuro-oncology and Orphan Cancers editorial team for their contribution to the editorial process. We thank the peer reviewers for their many helpful suggestions.

REFERENCES

References to studies included in this review

Agorastos 2005 *{published data only}*
Agorastos T, Dinas K, Lloveras B, de Sanjose S, Kornegay JR, Bonti H, et al. Human papillomavirus testing for primary screening in women at low risk of developing cervical cancer. The Greek experience. *Gynecologic Oncology* 2005;**96**:714–20.

Agorastos 2015 *{published data only}*
Agorastos T, Chatzistamatiou K, Katsamagkas T, Koliopoulos G, Daponte A, Constantinidis T, et al. Primary screening for cervical cancer based on high-risk human papillomavirus (HPV) detection and HPV 16 and HPV 18 genotyping, in comparison to cytology. *PLoS ONE* 2015; **10**(3):e0119755.

Belinson 2003 *{published data only}*
Belinson JL, Qiao YL, Pretorius RG, Zhang WH, Rong SD, Huang MN, et al. Shanxi Province cervical cancer screening study II: self-sampling for high-risk human papillomavirus compared to direct sampling for human papillomavirus and liquid based cervical cytology. *International Journal of Gynecological Cancer* 2003;**13**:819–26.

Belinson 2010 *{published data only}*
Belinson SE, Wulan N, Li R, Zhang W, Rong X, Zhu Y, et al. SNIPER: a novel assay for human papillomavirus testing among women in Guizhou, China. *International Journal of Gynecological Cancer* 2010;**20**:1006–10.

Bigras 2005 *{published data only}*
Bigras G, de Marval F. The probability for a Pap test to be abnormal is directly proportional to HPV viral load: results from a Swiss study comparing HPV testing and liquid-based cytology to detect cervical cancer precursors in 13, 842 women. *British Journal of Cancer* 2005;**93**:575-81.

Blumenthal 2001 *{published data only}*
Blumenthal PD, Gaffikin L, Chirenje ZM, McGrath J, Womack S, Shah K. Adjunctive testing for cervical cancer in low resource settings with visual inspection, HPV, and the Pap smear. *International Journal of Gynaecology and Obstetrics* 2001;**72**:47–53.

Cardenas-Turanzas 2008 *{published data only}*
Cardenas-Turanzas M, Nogueras-Gonzalez GM, Scheurer ME, Adler-Storthz K, Benedet JL, Beck R, et al. The performance of Human papillomavirus high-risk DNA testing in the screening and diagnostic settings. *Cancer Epidemiology, Biomarkers & Prevention* 2008;**17**:2865–71.

Castle 2011a *{published data only}*
Castle PE, Stoler MH, Wright TC Jr, Sharma A, Wright TL, Behrens CM. Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical cancer screening of women aged 25 years and older: a sub analysis of the ATHENA study. *Lancet Oncology* 2011;**12**:880-90.

Clavel 2001 *{published data only}*
Clavel C, Masure M, Bory JP, Putaud I, Mangeonjean C, Lorenzato M, et al. Human papillomavirus testing in primary screening for the detection of high-grade cervical lesions: a study of 7932 women. *British Journal of Cancer* 2001;**84**:1616–23.

Cuzick 1995 *{published data only}*
Cuzick J, Szarewski A, Terry G, Ho L, Hanby A, Maddox P, et al. Human papillomavirus testing in primary cervical screening. *Lancet* 1995;**345**:1533–6.

Cuzick 1999 *{published data only}*
Cuzick J, Beverley E, Ho L, Terry G, Sapper H, Mielzynska I, et al. HPV testing in primary screening of older women. *British Journal of Cancer* 1999;**81**:554–8.

Cuzick 2003 *{published data only}*
Cuzick J, Swarewski A, Cubie H, Hulman G, Kitchener H, Luesley D, et al. Management of women who test positive for high-risk types of human papillomavirus: the HART study. *The Lancet* 2003;**362**:1871–6.

de Cremoux 2003 *{published data only}*
de Cremoux P, Coste J, Sastre-Garau X, Thioux M, Bouillac C, Labbe S, et al. Efficiency of the hybrid capture 2 HPV DNA test in cervical cancer screening. *American Journal of Clinical Pathology* 2003;**120**:492–9.

Depuydt 2011 *{published data only}*
Depuydt CE, Makar AP, Ruymbeke MJ, Benoy IH, Vereecken AJ, Bogers JJ. BD-ProExC as adjunct molecular marker for improved detection of CIN2+ after HPV primary screening. *Cancer Epidemiol Biomarkers & Prevention* 2011; **20**:628–37.

Ferreccio 2013 *{published data only}*
Ferreccio C, Barriga MI, Lagos M, Ibanez C, Poggi H, Gonzalez F, et al. Screening trial of human papillomavirus for early detection of cervical cancer in Santiago, Chile. *International Journal of Cancer* 2013;**132**:916-23.

Gravitt 2010 *{published data only}*
Gravitt PE, Paul P, Katki HA, Vendantham H, Ramakrishna G, Sudula M, et al. Effectiveness of VIA, Pap and HPV DNA testing in a cervical cancer screening program in a peri-urban community in Andra Pradesh, India. *PLOS one* 2010;**5**:e13711.

Hovland 2010 *{published data only}*
Hovland S, Arbyn M, Lie AK, Ryd W, Borge B, Berle EJ, et al. A comprehensive evaluation of the accuracy of cervical pre-cancer detection methods in a high-risk area in East Congo. *British Journal of Cancer* 2010;**102**:957–65.

Iftner 2015 *{published data only}*
Iftner T, Becker S, Neis KJ, Castanon A, Iftner A, Holz B, et al. Head-to-head comparison of the RNA-based Aptima human papillomavirus (HPV) assay and the DNA-based hybrid capture 2 HPV test in a routine screening population of women aged 30 to 60 years in Germany. *Journal of Clinical Microbiology* 2015;**53**:2509–16.

Kulasingam 2002 *{published data only}*
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.

Labani 2014 *{published data only}*
Labani S, Asthana S, Sodhani P, Gupta S, Bhambhani S, Pooja B, et al. Care HPV cervical cancer screening demonstration in a rural population of North India. *European Journal of Obstetrics & Gynecology and Reproductive Biology* 2014;**176**:75–9.

Li 2009 *{published data only}*
Li N, Shi J-F, Franceshi S, Zhang W-H, Dai M, Liu B, et al. Different cervical cancer screening approaches in a Chinese multicentre study. *British Journal of Cancer* 2009; **100**:532–7.

Mahmud 2012 *{published data only}*
Mahmud SM, Sangwa-Lugoma G, Nasr SH, Kayembe PK, Tozin RR, Drouin P, et al. Comparison of human papillomavirus testing and cytology for cervical cancer screening in a primary health care setting in the Democratic Republic of the Congo. *Gynecologic Oncology* 2012;**124**: 286–91.

McAdam 2010a *{published data only}*
McAdam M, Sakita J, Tarivonda L, Pang J, Frazer IH. Evaluation of a cervical cancer screening program based on HPV testing and LLETZ excision in a low resource setting. *PLoS ONE* 2010;**5**:e13266.

McAdam 2010b *{published data only}*
McAdam M, Sakita J, Tarivonda L, Pang J, Frazer IH. Evaluation of a cervical cancer screening program based on HPV testing and LLETZ excision in a low resource setting. *PLoS ONE* 2010;**5**:e13266.

Monsonego 2011 *{published data only}*
Monsonego J, Hudgens MG, Zerat L, Zerat J-C, Syrjanen 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. *International Journal of Cancer* 2011;**129**:691-701.

Moy 2010 *{published data only}*
Moy LM, Zhao F-H, Li L-Y, Ma J-F, Zhang Q-M, Chen F, et al. Human papillomavirus testing and cervical cytology in primary screening for cervical cancer among women in rural China: comparison of sensitivity, specificity, and frequency of referral. *International Journal of Cancer* 2010; **127**:646–56.

Naucler 2009 *{published data only}*
Naucler P, Ryd W, Tornberg S, Strand A, Wadell G, Elfgren K, et al. Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. *Journal of the National Cancer Institute* 2009; **101**:88–99.

Nieves 2013 *{published data only}*
Nieves L, Enerson CL, Belinson S, Brainard J, Chiesa-Vottero A, Nagore N, et al. Primary cervical cancer screening and triage using an mRNA human papillomavirus assay and visual inspection. *International Journal of Gynecological Cancer* 2013;**23**:513–8.

Pan 2003 *{published data only}*
Pan Q, Belinson JL, Li L, Pretorius RG, Qiao YL, Zhang WH, et al. A thin layer, liquid-based Pap test for mass screening in an area of China with a high incidence of cervical carcinoma. A cross-sectional, comparative study. *Acta Cytologica* 2003;**47**:45–50.

Paraskevaidis 2001 *{published data only}*
Paraskevaidis E, Malamou-Mitsi V, Koliopoulos G, Pappa L, Lolis E, Georgiou I, et al. Expanded cytological referral criteria for colposcopy in cervical screening: comparison with human papillomavirus testing. *Gynecologic Oncology* 2001;**82**:355–9.

Petry 2003 *{published data only}*
Petry KU, Menton S, Menton M, van Loenen-Frosch F, de Carvalho Gomes H, Holz B, et al. Inclusion of HPV testing in routine cervical cancer screening for women above 29 years in Germany: results for 8466 patients. *British Journal of Cancer* 2003;**88**:1570–7.

Qiao 2008 *{published data only}*
Qiao YI, Sellors JW, Eder PS, Bao Y-P, Lim JM, Zhao F-H, et al. A new HPV DNA test for cervical-cancer screening in developing regions: a cross-sectional study of clinical accuracy in rural China. *Lancet Oncology* 2008;**9**:929–36.

Ronco 2006 *{published data only}*
Ronco G, Segnan N, Giorgi-Rossi P, Zappa M, Casadei GP, Carozzi F, et al. Human papillomavirus testing and liquid based cytology: results at recruitment from the new technologies for cervical cancer randomized controlled trial. *Journal of the National Cancer Institute* 2006;**98**:765–74.

Salmeron 2003 *{published data only}*
Salmeron J, Lazcano-Ponce E, Lorincz A, Hernandez M, Hernandez P, Leyva A, et al. Comparison of HPV-based assays with Papanicolaou smears for cervical cancer screening in Morelos State, Mexico. *Cancer Causes & Control* 2003;**14**:505–12.

Sankaranarayanan 2004a *{published data only}*
Sankaranarayanan R, Chatterji R, Shastri SS, Wesley RS, Basu P, Mahe C, et al. Accuracy of human papillomavirus testing in primary screening of cervical neoplasia: results from a multicenter study in India. *International Journal of Cancer* 2004;**112**:341–7.

Sarian 2005 *{published data only}*
Sarian LO, Derchain SF, Naud P, Roteli-Martins C, Longatto-Filho A, Tatti S, et al. Evaluation of visual inspection with acetic acid (VIA), Lugol's Iodine (VILI), cervical cytology and HPV testing as cervical screening tools in Latin America. This report refers to partial results from the LAMS study. *Journal of Medical Screening* 2005;**12**: 142–9.

Schneider 2000 *{published data only}*
Schneider A, Hoyer H, Lotz B, Leistritza S, Kuhne-Heid R, Nindl I, et al. Screening for high-grade cervical intra-epithelial neoplasia and cancer by testing for high-risk HPV, routine cytology or colposcopy. *International Journal of Cancer* 2000;**89**:529–34.

Shipitsyna 2011 *{published data only}*
Shipitsyna E, Zolotoverkhaya E, Kuevda D, Nasonova V, Romanyuk T, Khachaturyan A, et al. Prevalence of high-risk human papillomavirus types and cervical squamous intraepithelial lesions in women over 30 years of age in St. Petersburg, Russia. *Cancer Epidemiology* 2011;**35**:160–4.

Syrjanen 2002 *{published and unpublished data}*
Syrjaenen S, Shabalova I. Human papillomavirus testing and conventional Pap smear cytology as optional screening tools of women at different risks for cervical cancer in the countries of the former Soviet Union. *Journal of Lower Genital Tract Disease* 2002;**6**:97–110.

Wu 2010 *{published data only}*
Wu R, Belinson SE, Du H, Na W, Qu X, Wu R, et al. Human papillomavirus messenger RNA assay for cervical cancer screening: the Shenzhen Cervical Cancer Screening Trial I. *International Journal of Gynecological Cancer* 2010; **20**:1411–4.

References to studies excluded from this review

An 2003 *{published data only}*
An HJ, Cho NH, Lee SY, Kim IH, Lee C, Mun MS, et al. Correlation of cervical carcinoma and precancerous lesions with human papillomavirus (HPV) genotypes detected with the HPV DNA Chip Microarray Method. *Cancer* 2003;**97**: 1672–80.

Belinson 2001 *{published data only}*
Belinson J, Qiao YL, Pretorius R, Zhang WH, Elson P, Li L, et al. Shanxi province cervical cancer screening study: a cross-sectional comparative trial of multiple techniques to detect cervical neoplasia. *Gynecologic Oncology* 2001;**83**: 439–44.

Belinson 2010a *{published data only}*
Belinson JL, Hu S, Niyazi M, Pretorius RG, Wang H, Wen C, et al. Prevalence of type-specific human papillomavirus in endocervical, upper and lower vaginal, perineal and vaginal self-collected specimens: implications for vaginal self-collection. *International Journal of Cancer* 2010;**127**: 1151–7.

Belinson 2011 *{published data only}*
Belinson JL, Wu R, Belinson SE, Qu X, Yang B, Du H, et al. A population-based clinical trial comparing endocervical high-risk HPV testing using hybrid capture 2 and Cervista From the SHENCCAST II Study. *American Journal of Clinical Pathology* 2011;**135**:790–5.

Belinson 2012 *{published data only}*
Belinson JL, Du H, Yang B, Wu R, Belinson SE, Qu X, et al. Improved sensitivity of vaginal self-collection and high-risk human papillomavirus testing. *International Journal of Cancer* 2012;**15**:1855–60.

Benevolo 2011a *{published data only}*
Benevolo M, Terrenato I, Mottolese M, Marandino F, Carosi M, Rollo F, et al. Diagnostic and prognostic validity of the human papillomavirus E6/E7 mRNA test in cervical cytological samples of HC2-positive patients. *Cancer Causes & Control* 2011;**22**:869-75.

Benevolo 2011b *{published data only}*
Benevolo M, Vocaturo A, Caraceni D, French D, Rosini S, Zappacosta R, et al. Sensitivity, specificity, and clinical value of human papillomavirus (HPV) E6/E7 mRNA assay as a triage test for cervical cytology and HPV DNA test. *Journal of Clinical Microbiology* 2011;**49**:2643–50.

Benoy 2011 *{published data only}*
Benoy IH, Vanden Broeck D, Ruymbeke MJ, Sahebali S, Arbyn M, Bogers JJ, et al. Prior knowledge of HPV status improves detection of CIN2+ by cytology screening. *American Journal of Obstetrics and Gynecology* 2011;**205**: 569.e1–7.

Castle 2010 *{published data only}*
Castle PE, Schiffman M, Wheeler CM, Wentzensen N, Gravitt PE. Impact of improved classification on the association of human papillomavirus with cervical precancer. *American Journal of Epidemiology* 2010;**171**:155-63.

Chao 2010 *{published data only}*
Chao A, Chang C-J, Lai C-H, Chao F-Y, Hsu Y-H, Chou H-H, et al. Incidence and outcome of acquisition of human papillomavirus infection in women with normal cytology-a population-based cohort study from Taiwan. *International Journal of Cancer* 2010;**126**:191-8.

Coquillard 2011 *{published data only}*
Coquillard G, Palao B, Patterson BK. Quantification of intracellular HPV E6/E7 mRNA expression increases the specificity and positive predictive value of cervical cancer screening compared to HPV DNA. *Gynecologic Oncology* 2011;**120**:89–93.

Costa 2000 *{published data only}*
Costa S, Sideri M, Syrjanen K, Terzano P, De Nuzzo M, De Simone P, et al. Combined Pap smear, cervicography and HPV DNA testing in the detection of cervical intraepithelial neoplasia and cancer. *Acta Cytologica* 2000;**44**:310–18.

Coste 2003 *{published data only}*
Coste J, Cochand-Priollet B, de Cremoux P, Le Galles C, Cartier I, Molinie V, et al. Cross sectional study of conventional cervical smear, monolayer cytology, and human papillomavirus DNA testing for cervical cancer screening. *BMJ* 2003;**326**:733–7.

Dai 2006 *{published data only}*
Dai M, Bao YP, Li N, Clifford GM, Vaccarella S, Snijders PJ, et al. Human papillomavirus infection in Shanxi Province, People's Republic of China: a population-based study. *British Journal of Cancer* 2006;**95**:96–101.

de Andrade 2011 *{published data only}*
de Andrade ACV, Luz PM, Velasque L, Veloso VG, Moreira RI, Russomano F, et al. Factors associated with colposcopy-histopathology confirmed cervical intraepithelial neoplasia among HIV infected women from Rio De Janeiro, Brazil. *PLoS ONE* 2011;**6(3)**:e18297.

Denny 2000 *{published data only}*
Denny L, Kuhn L, Pollack A, Wainwright H, Wright TC. Evaluation of alternative methods of cervical cancer screening for resource-poor settings. *Cancer* 2000;**89**: 826–33.

Depuydt 2012 *{published data only}*
Depuydt CE, Benoy IH, Beert JF, Criel AM, Bogers JJ, Arbyn M. Clinical validation of a type-specific real-time quantitative human papillomavirus PCR against the performance of hybrid capture 2 for the purpose of cervical cancer screening. *Journal of Clinical Microbiology* 2012;**50**: 4073–7.

De Vuyst 2005 *{published data only}*
De Vuyst H, Claeys P, Njiru S, Muchiri L, Steyaert S, De Sutter P, et al. Comparison of Pap smear, visual inspection with acetic acid, human papillomavirus DNA-PCR testing and cervicography. *International Journal of Gynecology and Obstetrics* 2005;**89**:120–6.

Diamantopoulou 2013 *{published data only}*
Diamantopoulou S, Spathis A, Chranioti A, Anninos D, Stamataki M, Chrelias C, et al. Liquid based cytology and HPV DNA testing in a Greek population compared to colposcopy and histology. *Clinical and Experimental Obstetrics & Gynecology* 2013;**40**:131–6.

Fereccio 2003 *{published data only}*
Ferreccio C, Bratti MC, Sherman ME, Herrero R, Wacholder S, Hildesheim A, et al. A comparison of single and combined visual, cytologic and virologic tests as screening strategies in a region at high risk of cervical cancer. *Cancer Epidemiology Biomarkers and Prevention* 2003;**12**: 815–23.

Huang 2010 *{published data only}*
Huang M-Z, Huang S, Li D-Q, Nie X-M, Li H-B, Jiang X-M. The study of the combination detection of HPV-DNA and p16INK4a in cervical lesions. *Medical Oncology* 2010; **28 Suppl 1**:S547–52.

Idelevich 2011 *{published data only}*
Idelevich P, Kristt D, Schechter E, Lew S, Elkeles A, Terkieltaub D, et al. Screening for cervical neoplasia: a community-based trial comparing Pap staining, human papilloma virus testing, and the new bi-functional Cell Detect stain. *Diagnostic Cytopathology* 2011;**40**:1054–61.

Junyangdikul 2013 *{published data only}*
Junyangdikul P, Tanchotsrinon W, Chansaenroj J, Nilyaimit P, Lursinsap C, Poovorawan Y. Clinical prediction based on HPV DNA testing by hybrid capture 2 (HC2) in combination with Liquid-based Cytology (LBC). *Asian Pacific Journal of Cancer Prevention* 2013;**14**:903–7.

Katki 2011 *{published data only}*
Katki HA, Kinney WK, Fetterman B, Lorey T, Poitras NE, Cheung L, et al. Cervical cancer risk for women undergoing concurrent testing for human papillomavirus and cervical cytology: a population-based study in routine clinical practice. *Lancet Oncology* 2011;**12**:663–72.

Kelesidis 2011 *{published data only}*
Kelesidis T, Aish L, Steller MA, Aish IS, Shen J, Foukas P, et al. Human papillomavirus (HPV) detection using in situ hybridization in histologic samples: correlations with cytologic changes and polymerase chain reaction HPV detection. *American Journal of Clinical Pathology* 2011;**136**: 119–27.

Kim 2013 *{published data only}*
Kim JH, Kim IW, Kim YW, Park DC, Kim YW, Lee KH, et al. Comparison of single-, double- and triple-combined testing, including Pap test, HPV DNA test and cervicography, as screening methods for the detection of uterine cervical cancer. *Oncology Reports* 2013;**29**:1645–51.

Kitchener 2011 *{published data only}*
Kitchener HC, Gilham C, Sargent A, Bailey A, Albrow R, Roberts C, et al. A comparison of HPV DNA testing and liquid based cytology over three rounds of primary cervical screening: extended follow up in the ARTISTIC trial. *European Journal of Cancer* 2011;**47**:864–71.

Kuhn 2000 *{published data only}*
Kuhn L, Denny L, Pollack A, Lorincz A, Richart RM, Wright TC. Human papillomavirus DNA testing for cervical cancer screening in low-resource settings. *Journal of the National Cancer Institute* 2000;**92**:818–25.

Li 2010a *{published data only}*

Li C, Wu M, Wang J, Zhang S, Zhu L, Pan J, et al. A population-based study on the risks of cervical lesion and human papillomavirus infection among women in Beijing, People's Republic of China. *Cancer Epidemiology, Biomarkers & Prevention* 2010;**19**:2655–64.

Li 2010b *{published data only}*

Li Y, Ye F, Lu W-G, Zeng W-J, Wei L-H, Xie X. Detection of human telomerase RNA gene in cervical cancer and precancerous lesions comparison with cytological and human papillomavirus DNA test findings. *International Journal of Gynecological Cancer* 2010;**20**:631–7.

Longatto-Filho 2012 *{published data only}*

Longatto-Filho A, Naud P, Derchain SF, Roteli-Martins C, Tatti S, Hammes LS, et al. Performance characteristics of Pap test, VIA, VILI, HR-HPV testing, cervicography, and colposcopy in diagnosis of significant cervical pathology. *Virchows Archiv* 2012;**460**:577–85.

Ma 2010 *{published data only}*

Ma L, Bian M-L, Cheng J-Y, Xiao W, Hao M, Zhu J, et al. Hybrid capture II for high-risk human papillomavirus DNA testing to detect cervical precancerous lesions: a qualitative and quantitative study. *Experimental and Therapeutic Medicine* 2010;**1**:193–8.

Masumoto 2003 *{published data only}*

Masumoto N, Fujii T, Ishikawa M, Mukai M, Saito M, Iwata T, et al. Papanicolaou tests and molecular analyses using new fluid-based specimen collection technology in 3000 Japanese women. *British Journal of Cancer* 2003;**88**: 1883–8.

Mesher 2010 *{published data only}*

Mesher D, Szarewski A, Cadman L, Cubie H, Kitchener H, Luesley D, et al. Long-term follow-up of cervical disease in women screened by cytology and HPV testing: results from the HART study. *British Journal of Cancer* 2010;**102**: 1405–10.

Monsonego 2012 *{published data only}*

Monsonego J, Hudgens MG, Zerat L, Zerat JC, Syrjänen K, Smith JS. 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). *Gynecologic Oncology* 2012; **125**:175–80.

Nieminen 2004 *{published data only}*

Nieminen P, Vuorma S, Viikki M, Hakama M, Anttila A. Comparison of HPV test versus conventional and automation-assisted Pap screening as potential screening tools for preventing cervical cancer. *BJOG* 2004;**111**: 842–8.

Oh 2001 *{published data only}*

Oh YL, Shin KJ, Han J, Kim DS. Significance of high-risk human papillomavirus detection by polymerase chain reaction in primary cervical cancer screening. *Cytopathology* 2001;**12**:75–83.

Ozcan ES 2011 *{published data only}*

Özcan ES, Taş kı n S, Ortaç F. High-risk human papilloma virus prevalence and its relation with abnormal cervical cytology among Turkish women. *Journal of Obstetrics and Gynaecology* 2011;**31**:656–8.

Quincy 2012 *{published data only}*

Quincy BL, Turbow DJ, Dabinett LN, Dillingham R, Monroe S. Diagnostic accuracy of self-collected human papillomavirus specimens as a primary screen for cervical cancer. *Journal of Obstetrics and Gynaecology* 2012;**32**: 795–9.

Ratnam 2000 *{published data only}*

Ratnam S, France EL, Ferenczy A. Human papillomavirus testing for primary screening of cervical cancer precursors. *Cancer Epidemiology Biomarkers and Prevention* 2000;**9**: 945–51.

Riethmuller 1999 *{published data only}*

Riethmuller D, Gay C, Bertrand X, Bettinger D, Schaal JP, Carbillet JP, et al. Genital human papillomavirus infection among women recruited for routine cervical cancer screening or for colposcopy determined by hybrid capture II and polymerase chain reaction. *Diagnostic Molecular Pathology* 1999;**8**:157–64.

Rijkaart 2012b *{published data only}*

Rijkaart DC, Berkhof J, Van Kemenade FJ, Coupe VM, Rozendaal L, Heideman DA, et al. HPV DNA testing in population-based cervical screening (VUSA-Screen study): results and implications. *British Journal of Cancer* 2012; **106**:975–81.

Sankaranarayanan 2004b *{published data only}*

Sankaranarayanan R, Thara S, Sharma A, Roy C, Shastri S, Mahé C, et al. Accuracy of conventional cytology: results from a multicentre screening study in India. *Journal of Medical Screening* 2004;**11**:77–84.

Sankaranarayanan 2005 *{published data only}*

Sankaranarayanan R, Nene BM, Dinshaw KA, Mahe C, Jayant K, Shastri SS, et al. A cluster randomized controlled trial of visual, cytology and human papillomavirus screening for cancer of the cervix in rural India. *International Journal of Cancer* 2005;**116**:617–23.

Schiffman 2000 *{published data only}*

Schiffman M, Herrero R, Hildesheim A, Sherman ME, Bratti M, Wacholder S, et al. HPV DNA screening in cervical cancer screening. Results from women in a high-risk province of Costa-Rica. *JAMA* 2000;**283**:87–93.

Shastri 2005 *{published data only}*

Shastri SS, Dinshaw K, Amin G, Goswami S, Patil S, Chinoy R, et al. Concurrent evaluation of visual, cytological and HPV testing as screening methods for the early detection of cervical neoplasia in Mumbai, India. *Bulletin of the World Health Organization* 2005;**83**:186–94.

Sherman 2003 *{published data only}*

Sherman ME, Lorincz AT, Scott D, Wacholder S, Castle PE, Glass AG, et al. Baseline cytology, human papillomavirus testing, and risk for cervical neoplasia: a 10-year cohort

analysis. *Journal of the National Cancer Institute* 2003;**95**: 46–52.

Siriaunkgul 2014 ***{published data only}***
Siriaunkgul S, Settakorn J, Sukpan K, Srisomboon J, Suprasert P, Kasatpibal N, et al. Population-based cervical cancer screening using high-risk HPV DNA test and liquid-based cytology in northern Thailand. *Asian Pacific Journal of Cancer Prevention* 2014;**15**:6837–42.

Surabhi 2011 ***{published data only}***
Surabhi K, Ragini M. Complementary procedures in cervical cancer screening in low resource settings. *The Journal of Obstetrics and Gynecology of India* 2011;**61**:436-8.

Wang 2013 ***{published data only}***
Wang JL, Yang YZ, Dong WW, Sun J, Tao HT, Li RX, et al. Application of human papillomavirus in screening for cervical cancer and precancerous lesions. *Asian Pacific Journal of Cancer Prevention* 2013;**14**:2979–82.

Womack 2000 ***{published data only}***
Womack SD, Chirenje ZM, Blumenthal PD, Gaffikin L, McGrath JA, Chipato T, et al. Evaluation of a human papillomavirus assay in cervical screening in Zimbabwe. *BJOG* 2000;**107**:33–8.

Zhao 2010 ***{published data only}***
Zhao F-H, Lin MJ, Chen F, Hu S-Y, Zhang R, Belinson JL, et al. Performance of high-risk human papillomavirus DNA testing as a primary screen for cervical cancer: a pooled analysis of individual patient data from 17 population-based studies from China. *Lancet Oncology* 2010;**11**:1160–71.

Additional references

Anttila 2004
Anttila A, Ronco G, Clifford G, Bray F, Hakama M, Arbyn M, et al. Cervical cancer screening programmes and policies in 18 European countries. *British Journal of Cancer* 2004; **91**:935–41.

Arbyn 2004
Arbyn M, Buntinx F, Van Ranst M, Paraskevaidis E, Martin-Hirsch P, Dillner J. Cytologic versus virologic triage of women with equivocal Pap smears: a meta-analysis of the accuracy to detect high-grade intraepithelial neoplasia. Journal of the National Cancer Institute 2004; Vol. 96: 280–93.

Arbyn 2006
Arbyn M, Sasieni P, Meijer CJ, Clavel C, Koliopoulos G, Dillner J. Chapter 9: Clinical applications of HPV testing: a summary of meta-analyses. *Vaccine* 2006;**24 Suppl 3**: 78–89.

Arbyn 2008
Arbyn M, Sankaranarayanan R, Muwonge R, Keita N, Dolo A, Mbalawa CG, et al. Pooled analysis of the accuracy of five cervical cancer screening tests assessed in eleven studies in Africa and India. *International Journal of Cancer* 2008; **123**:153–60.

Arbyn 2009
Arbyn M, Raifu AO, Bray F, Weiderpass E, Anttila A. Trends of cervical cancer mortality in the member states of the European Union. *European Journal of Cancer* 2009;**45 (15)**:2640–8.

Arbyn 2009a
Arbyn M, Ronco G, Cuzick J, Wentzensen N, Castle PE. How to evaluate emerging technologies in cervical cancer screening?. *International Journal of Cancer* 2009;**125**: 2489–96.

Arbyn 2010
Arbyn M, Anttila A, Jordan J, Ronco G, Schenck U, Segnan N, et al. European Guidelines for Quality Assurance in Cervical Cancer Screening. Second edition--summary document. *Annals of Oncology* 2010;**21**:448–58.

Arbyn 2011
Arbyn M, Castellsagué X, de Sanjosé S, Bruni L, Saraiya M, Bray F, et al. Worldwide burden of cervical cancer in 2008. *Annals of Oncology* 2011;**12**:2675–86.

Arbyn 2012
Arbyn M, Ronco G, Anttila A, Meijer CJ, Poljak M, Ogilvie G, et al. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. *Vaccine* 2012;**30 Suppl 5**:F88–99.

Bosch 2002
Bosch FX, Lorincz A, Munoz N, Meijer CJ, Shah KV. The causal relation between human papillomavirus and cervical cancer. *Journal of Clinical Pathology* 2002;**55**:244–65.

Breitenecker 2004
Breitenecker G, Dinges HP, Regitnig P, Wiener H, Vutuc C. Cytopathology in Austria. *Cytopathology* 2004;**15**:113–8.

Castle 2011b
Castle PE, Stoler MH, Wright TC Jr, Sharma A, Wright TL, Behrens CM. Performance of carcinogenic human papillomavirus (HPV) testing and HPV16 or HPV18 genotyping for cervical cancer screening of women aged 25 years and older: a subanalysis of the ATHENA study. *Lancet Oncology* 2011;**12**:880–90.

Chan 1999
Chan AB, Fox JD. NASBA and other transcription-based amplification methods for research and diagnostic microbiology. *Review Medical Microbiology* 1999;**10**: 184–96.

Chu 2006
Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *Journal of Clinical Epidemiology* 2006;**59** (12):1331–2.

Dijkstra 2013
Dijkstra MG, Van Niekerk D, Rijkaart DC, Van Kemenade FJ, Heideman DA, Snijders PJ, et al. Primary hrHPV DNA testing in cervical cancer screening: how to manage screen-positive women? A POBASCAM trial substudy. *Cancer Epidemiology, Biomarkers & Prevention* 2014;**23**:55–63.

Dilner 1999
Dilner J. The serological response to papillomaviruses. *Seminars in Cancer Biology* 1999;**9**:423–30.

IARC 2005
International Agency for research on Cancer. *Cervix Cancer Screening: IARC Handbooks of Cancer Prevention*. Vol. **10**, Lyon: IARC, 2005.

IARC 2007
IARC working group on the evaluation of carcinogenic risks to humans. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Volume 90: Human Papilloma viruses*. Lyon, France: IARC Press, 2007.

Jeronimo 2006
Jeronimo J, Schiffman M. Colposcopy at a crossroads. *American Journal of Obstetrics and Gynecology* 2006;**195**: 349–53.

Jordan 2008
Jordan J, Arbyn M, Martin-Hirsch P, Schenck U, Baldauf JJ, Da Silva D, et al. European guidelines for quality assurance in cervical cancer screening: recommendations for clinical management of abnormal cervical cytology, part 1. *Cytopathology* 2008;**19**:342–54.

Kjaer 2004
Kjaer S. Long term predictive value of HPV testing (the Danish prospective cohort study). 21st International Human Papillomavirus Conference. Mexico City, Feb 20–26, 2004:Abstract 339.

Laara 1987
Laara E, Day NE, Hakarna M. Trends in mortality from cervical cancer in the Nordic countries: association with organised screening programmes. *Lancet* 1987;**1**:1247–9.

Mandelblatt 2002
Mandelblatt JS, Lawrence WF, Womack SM, Jacobson D, Yi B, Hwang YT, et al. Benefits and costs of using HPV testing to screen for cervical cancer. *JAMA* 2002;**287**: 2372–81.

McCaffery 2004
McCaffery K, Waller J, Forrest S, Cadman L, Szarewski A, Wardle J. Testing positive for human papillomavirus in routine cervical screening: examination of psychosocial impact. *BJOG* 2004;**111**:1437–43.

McCredie 2008
McCredie MR, Sharples KJ, Paul C, Baranyai J, Medley G, Jones RW, et al. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. *Lancet Oncology* 2008;**9**:425–34.

McIndoe 1984
McIndoe WA, McLean MR, Jones RW, Mullins PR. The invasive potential of carcinoma in situ of the cervix. *Obstetrics and Gynecology* 1984;**64**:451–8.

Meera 2002
Meera E, Le Gales C, Cochand-Priollet B, Cartier I, de Cremoux P, Vacher-Lavenu MC, et al. Cost of screening for cancerous and precancerous lesions of the cervix. *Diagnostic Cytopathology* 2002;**27**:251–7.

Moher 2009
Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *PLoS Medicine* 2009;**6**(7):e1000097. [DOI: 10.1371/journal.pmed1000097

NHSCSP 2004
Luesley D, Leeson S. *NHSCSP Publication No. 20. Colposcopy and Programme Management. Guidelines for the NHS Cervical Screening Programme*. Sheffield: NHS Cancer Screening Programmes, 2004.

Papanicolaou 1941
Papanicolaou GN, Trauat HF. The diagnostic value of vaginal smears in carcinoma of the uterus. *American Journal of Obstetrics and Gynecology* 1941;**42**:193–206.

Poljak 2012
Poljak M, Cuzick J, Kocjan BJ, Iftner T, Dillner J, Arbyn M. Nucleic acid tests for the detection of alpha human papillomaviruses. *Vaccine* 2012;**30**(Suppl 5):F100–6.

Quinn 1999
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.

Reitsma 2005
Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005; **58**(10):982–90.

Reitsma 2009
Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Chapter 9: Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. The Cochrane Collaboration, 2009. Available from: srdta.cochrane.org/.

RevMan 2014 [Computer program]
Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager 5 (RevMan 5). Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Rijkaart 2012a
Rijkaart DC, Berkhof J, Van Kemenade FJ, Coupe VM, Hesselink AT, Rozendaal L, Heideman DA, Verheijen RH, Bulk S, Verweij WM, Snijders PJ, Meijer CJ. Evaluation of 14 triage strategies for HPV DNA-positive women in population-based cervical screening. *Int J Cancer* 2012;**130**: 602–10.

Ronco 2014
Ronco G, Dillner J, Elfström KM, Tunesi S, Snijders PJ, Arbyn M, Kitchener H, Segnan N, Gilham C, Giorgi-Rossi P, Berkhof J, Peto J, Meijer CJ, International HPV sc. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. *Lancet* 2014;**383**:524–32..

SAS 9.4 [Computer program]
SAS Institute Inc. SAS Institute Inc. SAS/ACCESS® 9.4 Interface to ADABAS. Reference. Cary, NC, 2013.

Saslow 2012
Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, Garcia FA, Moriarty AT, Waxman AG, Wilbur DC, Wentzensen N, Downs LS Jr, Spitzer M, Moscicki AB, Franco EL, Stoler MH, Schiffman M, Castle PE, Myers ER, American Cancer Society, American Society for Colposcopy and Cervical Pathology, American Society for Clinical Pathology. 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. *Am J Clin Pathol.* 2012;**137**:516–42.

Schatzkin 1987
Schatzkin A, Connor RJ, Taylor PR, Bunnag B. Comparing new and old screening tests when a reference procedure cannot be performed on all screenees. Example of automated cytometry for early detection of cervical cancer. *American Journal of Epidemiology* 1897;**125**:672–8.

Schiffman 2007
Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. *Lancet* 2007;**370**:890–907.

Sigurdsson 1999
Sigurdsson K. The Icelandic and Nordic cervical screening programs: trends in incidence and mortality rates through1995. *Acta Obstetricia et Gynecologica Scandinavica* 1999;**78**:478–85.

Solomon 2002
Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, et al. The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA* 2002;**287**: 2114–9.

Spence 2007
Spence AR, Goggin P, Franco EL. Process of care failures in invasive cervical cancers: A systematic review and meta-analysis. *Preventive Medicine* 2007;**45**:93–106.

STATA 14 [Computer program]
StataCorp. StataCorp. *Stata Statistical Software: Release 14*. College Station, TX. StataCorp LP, 2015.

Stoler 2001
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.

Takwoingi 2010
Twakoingi Y, Deeks JJ. MetaDAS: A SAS macro for meta-analysis of diagnostic accuracy studies. Quick reference and worked example.. Available from methods.cochrane.org/sdt/software-meta-analysis-dta-studies. 1.3 2010.

Walboomers 1999
Walboomers JMM, Jacobs MV, Manos MM, Bosch X, Kummer A, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *Journal of Pathology* 1999;**189**:12–9.

Wentzensen 2015
Wentzensen N, Walker JL, Gold MA, Smith KM, Zuna RE, Mathews C, et al. Multiple biopsies and detection of cervical cancer precursors at colposcopy. *Journal of Clinical Oncology* 2015;**33**:83–9.

Whiting 2003
Whiting P, Rutjes AWS, Reitsma JB, Bossuyt PMM, Kleijnen J. The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology* 2003;**3**:25.

Wilson 1968
Wilson JMG, Jungner G. Principles and practice of screening for disease. *Public Health Papers*. Vol. **34**, Geneva: World Health Organization, 1968.

Wright 2006
Wright TC Jr, Massad LS, Dunton CJ, Spitzer M, Wilkinson EJ, Solomon D. 2006 consensus guidelines for the management of women with abnormal cervical cancer screening tests. *American Journal of Obstetrics and Gynecology* 2007;**197**:346–55.

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Agorastos 2005

Clinical features and settings	Women >17 years old attending the outpatient clinics of six hospitals in Northern Greece for routine cervical screening. No history of hysterectomy or treatment for CIN
Participants	1296 women (mean age 43) from Greece
Study design	Cross-sectional study of women receiving both CC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsies. When colposcopy was normal biopsies were not taken. Colposcopy was performed in all screen-positives and in a random 5% of screen-negatives
Index and comparator tests	IT: HPV testing by PCR (PGMY09/PGMY11) for the detection of 27 HPV types (6, 11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 45, 51, 52, 53, 54, 55, 56, 57, 58, 59, 66, 68, 73, 82, 83, 84). Referred for colposcopy if positive CT: CC, referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women >17 years attending for routine screening
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	5% of screen-negatives also underwent colposcopy
Differential verification avoided? All tests	Yes	Same RS was applied in all cases
Incorporation avoided? All tests	Yes	Screening tests did not form part of the RS

Agorastos 2005 (*Continued*)

Reference standard results blinded? All tests	Yes	Pathologists were blinded to the screening tests
Index test results blinded? All tests	Yes	RS performed after the screening tests' interpretation
Relevant clinical information? All tests	Unclear	It is not clear what information was available to the cytologists
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

Agorastos 2015

Clinical features and settings	Women aged 25-55 attending routine cervical screening at the outpatient clinics of 9 Gynaecology Departments (2 in Athens, 4 in Thessaloniki, 1 in Larissa, 1 in Patras and 1 in Alexandroupolis) were asked to be enrolled in the study. Exclusion criteria were current pregnancy, current or previous history of CIN in the past 5 years, follow-up for cytological abnormalities and hysterectomy
Participants	4009 women attending for cervical screening in Greece. The mean age was 39.9 years
Study design	Cross-sectional study of women receiving both LBC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsies. When colposcopy was normal biopsies were not taken. Colposcopy was performed in all screen-positives and in a random 3% of screen-negatives
Index and comparator tests	IT: HPV testing by Cobas HPV test (Roche). Referred for colposcopy if positive CT: LBC. Referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women 25-55 years old

Agorastos 2015 (*Continued*)

Acceptable reference standard? All tests	Yes	Colposcopy and colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	3% of screen-negatives also underwent colposcopy
Differential verification avoided? All tests	Yes	Same RS was applied in all cases
Incorporation avoided? All tests	Yes	Screening tests did not form part of the RS
Reference standard results blinded? All tests	No	Pathologists were aware of the cytology and colposcopy result, but not of the HPV DNA test result
Index test results blinded? All tests	Yes	RS performed after the screening tests' interpretation
Relevant clinical information? All tests	Unclear	It is not clear what information was available to the cytologists
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

Belinson 2003

Clinical features and settings	Non-pregnant women 35-50 years old with no history of pelvic radiation or hysterectomy from villages in the Shanxi Province in China, were invited to participate
Participants	8497 women (mean age 40.9) from rural China
Study design	Cross-sectional study of women receiving both cytology, direct and self-HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy, if colposcopy was normal random biopsies were taken. Only for screen-positives
Index and comparator tests	IT: HPV testing by HC2 direct cervical sampling (Digene), positivity threshold at 1 pg/mL. Referred for colposcopy + biopsies if positive IT: HPV testing by HC2 self-vaginal sampling (Digene), positivity threshold at 1 pg/mL. Referred for colposcopy + biopsies if positive

Belinson 2003 (*Continued*)

	CT: LBC. Referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women 35-50 years old
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	3 months between self-sampling and direct-sampling
Partial verification avoided? All tests	No	Colposcopy only performed when the tests were positive
Differential verification avoided? All tests	Yes	The same RS was performed on all occasions
Incorporation avoided? All tests	Yes	The RS did not consist of cytology or HPV testing
Reference standard results blinded? All tests	Yes	Pathologists were blinded to the test results
Index test results blinded? All tests	Yes	The RS was applied after the screening tests
Relevant clinical information? All tests	Unclear	It is not clear what information was given to the cytologists
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

Belinson 2010

Clinical features and settings	1000 women were recruited by the Renmin Hospital in the Buyi-Miao Autonomous District (BMAD) of Guizhou Province, China. Women were excluded if they were pregnant, younger than 30, did not have an intact uterus, or had a history of pelvic irradiation or cervical cancer
Participants	979 women aged 30-54 examined in a colposcopy clinic in Guizhou, China
Study design	Cross-sectional study of women receiving both LBC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy, if colposcopy was normal random biopsies were taken. Only for screen-positives
Index and comparator tests	IT: HPV testing by SNIPER (Genetel Pharmaceuticals). Referred for colposcopy + biopsies if positive CT: LBC. Referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women aged 30-54 years
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Both tests performed at the same visit
Partial verification avoided? All tests	No	Only women with positive results were invited for colposcopy
Differential verification avoided? All tests	Yes	The same RS applied for all occasions
Incorporation avoided? All tests	Yes	The tests were not part of the RS
Reference standard results blinded? All tests	Unclear	It is not clear whether the colposcopists had knowledge of the test results
Index test results blinded? All tests	Yes	Colposcopy was performed after the test results were reported

Belinson 2010 (*Continued*)

Relevant clinical information? All tests	Unclear	It is not clear what information was given to the cytologists
Uninterpretable results reported? All tests	No	There is no mention of un interpretable results
Withdrawals explained? All tests	Yes	Of the 211 women asked to return, all but 21 or 90% of the women returned for colposcopy

Bigras 2005

Clinical features and settings	Women mainly 30 or older attending mainly private gynaecologists in Switzerland
Participants	13,842 women (mean age 44.4 years, range 17-93) from Switzerland
Study design	Cross-sectional study of women receiving both LBC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsies, if colposcopy was normal random biopsies were taken. Screen-positives and a random 5% sample of screen-negatives were referred
Index and comparator tests	IT: HPV testing by HC2, positivity threshold 1 pg/mL, referred for colposcopy if positive CT: LBC (Surepath), referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women > 17 years attending for routine screening
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	A random 5% sample of screen-negatives was also verified
Differential verification avoided? All tests	Yes	The same RS was used in all cases

Bigras 2005 (*Continued*)

Incorporation avoided? All tests	Yes	The screening tests did not form part of the RS
Reference standard results blinded? All tests	Unclear	It is unclear whether the pathologists had knowledge of the test results
Index test results blinded? All tests	Yes	The RS was applied after the tests were reported
Relevant clinical information? All tests	Unclear	It is unclear whether cytologists had knowledge of any clinical information
Uninterpretable results reported? All tests	No	Unsatifactory smears were not reported
Withdrawals explained? All tests	Yes	Were explained

Blumenthal 2001

Clinical features and settings	Women 25-55 years old attending primary care clinics in Zimbabwe were invited. No hysterectomy or previous diagnosis of cervical cancer
Participants	2073 women from Chitungwiza and the greater Harare area in Zimbabwe
Study design	Cross-sectional study of women receiving both conventional cytology, VIA and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopy with or without biopsies was performed on all women. If colposcopy was normal, biopsies were not taken
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold 1 pg/mL IT: VIA CT: CC. For the calculation of the accuracy indices the threshold of LSIL+ was used
Follow-up	
Notes	Some of the data were extracted from the publication by Womack 2000

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women 25-55 years old attending primary care clinics

Blumenthal 2001 (*Continued*)

Acceptable reference standard? All tests	Yes	Colposcopy with biopsy (no biopsy if colposcopy was normal)
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	All women underwent colposcopy
Differential verification avoided? All tests	Yes	The same RS was applied to all women
Incorporation avoided? All tests	Yes	The screening tests were not part of the RS
Reference standard results blinded? All tests	Yes	Colposcopists were not aware of the screening test results
Index test results blinded? All tests	Unclear	It is not clear whether cytologists were aware of the colposcopic diagnosis
Relevant clinical information? All tests	Unclear	It is not clear whether cytologists had the routine clinical information
Uninterpretable results reported? All tests	No	The numbers of inadequate smears and HPV tests are not given
Withdrawals explained? All tests	Yes	Withdrawals are explained

Cardenas-Turanzas 2008

Clinical features and settings	Women > 30 years old without prior cervical abnormalities, attending two cancer centres and one general hospital
Participants	835 women (mean age 46.7 years) undergoing routine screening in the USA and Canada
Study design	Cross-sectional study of women receiving both cytology and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy, if colposcopy was normal random biopsies were taken. All women were referred for colposcopy
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold at 1 pg/mL CT: CC For the calculation of the accuracy indices an abnormal result was considered any smear showing ASCUS+
Follow-up	

Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women > 30 years without prior cervical abnormalities
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies and random biopsies if colposcopy was negative
Acceptable delay between tests? All tests	Yes	Both tests were serially performed at the same visit
Partial verification avoided? All tests	Yes	All women were referred for colposcopy
Differential verification avoided? All tests	Yes	The same RS was applied in the case of positive cytology and persistent type-specific positive HPV test
Incorporation avoided? All tests	Yes	The RS was not composed of the index and comparator tests
Reference standard results blinded? All tests	Yes	The pathologists were not aware of the screening tests results
Index test results blinded? All tests	Yes	The personnel reporting the screening test results were not aware of the RS results
Relevant clinical information? All tests	Unclear	It is not clear whether clinical information was given to the cytologists
Uninterpretable results reported? All tests	Yes	All results including inadequate specimens were reported
Withdrawals explained? All tests	Yes	Withdrawals were explained

Castle 2011a

Clinical features and settings	Women presenting for routine cervical cancer screening were enrolled into the ATHENA study at 61 clinical centres in 23 US states. Eligible women were aged 21 years or older and were not pregnant. Eligible women had an intact uterus, had not received treatment for CIN with 12 months of enrolment, and had no present or planned participation in a clinical trial for HPV treatment. For this sub-analysis, the population was restricted to women aged 25 years and older
Participants	41,955 women aged 25 years or older (mean age 41.9) in the USA
Study design	Cross-sectional study of women receiving both cytology and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: Colposcopy with or without biopsies was performed on all women with a positive screening test and a random sample of women with negative tests
Index and comparator tests	IT: HPV testing by Cobas HPV test (Roche). Positivity was not a criterion for referral. Referred for colposcopy + biopsies if a first-generation HPV test (Amplicor or Linear Array) was positive. That left only 48/4275 Cobas-positive women without referral for colposcopy CT: LBC. Referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women aged ≥ 25 years attending routine screening
Acceptable reference standard? All tests	Yes	Colposcopy with or without biopsies
Acceptable delay between tests? All tests	Yes	All tests performed at the same visit
Partial verification avoided? All tests	Yes	A random sample of 1041 screen-negative women had colposcopy
Differential verification avoided? All tests	Yes	The same RS applied to all women
Incorporation avoided? All tests	Yes	The screening tests did not form part of the RS
Reference standard results blinded? All tests	Yes	Colposcopists and pathologists were masked to cytology and HPV test results

Castle 2011a (*Continued*)

Index test results blinded? All tests	Yes	Colposcopy was performed after the tests were reported
Relevant clinical information? All tests	Unclear	It is not clear what information was available to the cytologists
Uninterpretable results reported? All tests	Yes	1054 women had missing or invalid test results
Withdrawals explained? All tests	Yes	Withdrawals were explained

Clavel 2001

Clinical features and settings	Asymptomatic non-pregnant women 15-76 years old without recent cervical cytological abnormalities, or untreated cervical lesion in the last 2 years, or AIDS, attending a central urban hospital for routine screening
Participants	7932 women (median age 34) undergoing biennial or triennial routine screening, Rheims, France
Study design	Longitudinal study of women receiving both cytology and HPV testing
Target condition and reference standard(s)	TC: histologically-proven HSIL RS: colposcopically-directed biopsy or LEEP. Colposcopy only if no lesion was seen. Only for screen-positives
Index and comparator tests	IT: HPV testing by HC2 (DIgene), positivity threshold at 1 pg/mL. If cytology was negative, women with a positive HPV test were referred for the RS 6 months later if a second HPV test was also positive CT: CC in 2281 women, LBC (Thinprep) in 5651 women. Referred for RS if the result was ASCUS or worse
Follow-up	368/773 women with positive HPV test but negative cytology did not return for a second HPV test
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women 15-76 years old undergoing routine screening
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsy or LEEP

Clavel 2001 (*Continued*)

Acceptable delay between tests? All tests	Yes	Test serially performed at the same examination
Partial verification avoided? All tests	No	The RS was applied only if one of the test results were positive
Differential verification avoided? All tests	Yes	The RS was the same in all situations
Incorporation avoided? All tests	Yes	The RS was not composed of the index and comparator tests
Reference standard results blinded? All tests	No	Colposcopists were aware of the test results
Index test results blinded? All tests	Yes	The RS was applied after the screening tests were reported
Relevant clinical information? All tests	Unclear	There is not sufficient information
Uninterpretable results reported? All tests	No	The numbers of uninterpretable results that would be expected to have occurred were not given
Withdrawals explained? All tests	No	The withdrawals were not completely explained

Cuzick 1995

Clinical features and settings	Women attending a family planning clinic in London for routine smear. No history of CIN or abnormal smear in the last 3 years
Participants	1985 women (median age 29 years, 93% between 20 and 45) in London, UK
Study design	Cross-sectional study of women receiving both CC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy. If colposcopy was normal a biopsy might have not been taken. RS applied only to screen-positives
Index and comparator tests	IT: HPV testing by PCR for 4 high-risk types (16, 18, 31, 33). Referred for colposcopy if positive CT: CC referred for colposcopy if ASCUS+
Follow-up	
Notes	

Cuzick 1995 (*Continued*)

Table of Methodological Quality		
Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women of the appropriate age spectrum attending for routine cervical screening
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	No	Only screen-positives had colposcopy
Differential verification avoided? All tests	Yes	The same RS in all cases
Incorporation avoided? All tests	Yes	Screening tests not part of the RS
Reference standard results blinded? All tests	Yes	Pathologists were blinded to the results of the screening tests
Index test results blinded? All tests	Yes	RS performed after the interpretation of the screening tests
Relevant clinical information? All tests	Unclear	It is unclear what information was given to the cytologists
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

Cuzick 1999

Clinical features and settings	Women aged ≥ 35 years attending for a routine smear in general practitioner practices in the UK , no previous treatment, no cytologic abnormality in the last 3 years
Participants	2988 women (mean age 46) in the UK
Study design	Cross-sectional study of women receiving both CC and HPV testing

Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy, if colposcopy was normal a biopsy was not taken. Only screen-positives were referred for colposcopy
Index and comparator tests	IT: HPV testing by PCR (MY09/11) for the detection of 10 high-risk types (16, 18, 31, 33, 35, 45, 51, 52, 56, 58). Referred for colposcopy if positive CT: CC, referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women > 34 years attending for routine screening
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsy
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	No	Only screen-positives underwent colposcopy
Differential verification avoided? All tests	Yes	The same RS was applied to all screen-positives
Incorporation avoided? All tests	Yes	The screening tests were not part of the RS
Reference standard results blinded? All tests	Yes	Pathologists were blinded to the results of the screening tests
Index test results blinded? All tests	Yes	The RS was performed after the screening tests were reported
Relevant clinical information? All tests	Unclear	It is not clear whether cytologists were given the routine clinical information
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

Cuzick 2003

Clinical features and settings	Women 30-60 years old attending for routine cervical screening were recruited from 161 family practices in the UK. No treatment for CIN, no abnormal smear in the last 3 years
Participants	10,358 women (mean age 42 years) from the UK
Study design	Women received both CC and HPV testing. Then an RCT on the management of women with minor abnormalities (borderline smears and HPV positives with negative smears) was conducted. Women were randomised to either surveillance at 6-12 months or immediate colposcopy
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsies, if colposcopy was negative biopsies were not taken. All screen-positives eventually underwent colposcopy although some with a 12-month delay (the ones with minor abnormalities randomised to surveillance). A random 5% sample of screen-negatives also underwent colposcopy
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold 1 pg/mL, referred for colposcopy if positive CT: CC, referred for colposcopy if ASCUS+
Follow-up	Women with ASCUS or HPV-positive with negative smears were randomised either to immediate colposcopy or to surveillance at 6 and 12 months with colposcopy performed at the end of the 12 months
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women 30-60 years old attending for routine screening
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	No	In 296 women colposcopy was performed 12 months after the screening tests
Partial verification avoided? All tests	Yes	Colposcopy was also performed in 5% of screen-negatives
Differential verification avoided? All tests	Yes	Same RS in all cases
Incorporation avoided? All tests	Yes	Screening tests not part of the RS

Cuzick 2003 (*Continued*)

Reference standard results blinded? All tests	No	Colposcopists and pathologists were aware of the results
Index test results blinded? All tests	Yes	Screening tests done before the RS
Relevant clinical information? All tests	Yes	Cytologists received routine clinical information
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

de Cremoux 2003

Clinical features and settings	Women older than 18, not pregnant, without a recent (< 1 year) history of surgery or laser treatment of the cervix, whose cervix was visible by the physician, attending for cervical smear in a French university hospital or private practices
Participants	1757 women (mean age 33.3) in France
Study design	Cross-sectional study of women receiving CC LBC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy, if colposcopy was normal, biopsies were not taken. Colposcopy was performed in all women
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold 1 pg/mL CT: CC, positivity threshold ASCUS+ CT: LBC, positivity threshold ASCUS+
Follow-up	
Notes	Some cytological data are taken from the article Coste 2003

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women > 18 years attending for routine screening
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsy

de Cremoux 2003 (*Continued*)

Acceptable delay between tests? All tests	Yes	Performed at the same time
Partial verification avoided? All tests	Yes	RS applied to all women
Differential verification avoided? All tests	Yes	The same RS in all cases
Incorporation avoided? All tests	Yes	Screening tests not part of the RS
Reference standard results blinded? All tests	Yes	Pathologists were blinded to the screening tests
Index test results blinded? All tests	Yes	Cytologist were blinded to other results
Relevant clinical information? All tests	Unclear	It is not clear what information was available to the cytologists
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Unclear	It is not clear whether there were any withdrawals

Depuydt 2011

Clinical features and settings	Women undergoing routine screening in 9 gynaecological practices in Flanders (Belgium) . Exclusion criteria included pregnancy and history of cervical disease
Participants	3126 women with a median age of 42.7 years (range 18.0-84.3) in Flanders, Belgium
Study design	Cross-sectional study of women receiving LBC, HPV testing and BD ProExC ICC staining
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ and CIN 3+ RS: colposcopically-directed biopsy, if colposcopy was normal, biopsies were not taken. Colposcopy was performed in all women
Index and comparator tests	IT: HPV testing by PCR for the detection of 13 high-risk types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68) IT: BD ProExC immunocytochemistry CT: LBC, positivity threshold ASCUS+
Follow-up	Women were followed up for the detection of CIN 2+ for a further period of 24 months

Depuydt 2011 (*Continued*)

Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women attending for routine screening
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	Colposcopy performed in all women
Differential verification avoided? All tests	Yes	The same RS applied in all situations
Incorporation avoided? All tests	Yes	The tests did not form part of the RS
Reference standard results blinded? All tests	No	In the initial colposcopy yes but in the subsequent one the colposcopist was aware of the results
Index test results blinded? All tests	Unclear	Unclear whether cytologists were aware of the colposcopy results
Relevant clinical information? All tests	Unclear	It is unclear what information was available to the cytologists
Uninterpretable results reported? All tests	Yes	None had an inadequate smear
Withdrawals explained? All tests	Unclear	It does not seem as if there have been any withdrawals from follow-up

Ferreccio 2013

Clinical features and settings	Women residing in Santiago, Chile, were invited to participate through an outreach campaign, excluding women who were pregnant, hysterectomised or virgins
Participants	8407 women from Santiago, Chile (mean age 42.2 years)
Study design	Cross-sectional study of women receiving CC, HPV testing

Ferreccio 2013 (*Continued*)

Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ and CIN 3+ RS: colposcopically-directed biopsy. Colposcopy was also performed in a random sample of screen-negatives
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold 1 pg/mL CT: CC, positivity threshold ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women aged 25-64 years
Acceptable reference standard? All tests	Yes	Colposcopy and colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	Colposcopy was also performed in sample of high-risk (VIA positive) screen-negatives
Differential verification avoided? All tests	Yes	The same RS applied in all situations
Incorporation avoided? All tests	Yes	The tests did not form part of the RS
Reference standard results blinded? All tests	No	Pathologists were blind to the HPV test result, but not necessarily to the Pap test result
Index test results blinded? All tests	Yes	Screening tests done before the RS
Relevant clinical information? All tests	Unclear	It is unclear what information was available to the cytologists
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

Gravitt 2010

Clinical features and settings	Women > 25 years old mentally competent with an intact uterus were invited from 42 villages in a peri-urban rural community
Participants	2331 women (mean age 37) from Andhra Pradesh, India
Study design	Cross-sectional study of women receiving CC, VIA and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ or CIN 3+ RS: colposcopically-directed biopsy, colposcopy only if no lesion was seen. Only for screen-positives
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold at 1 pg/mL. Referred for colposcopy if positive IT: VIA, referred for colposcopy if positive CT: CC. Referred for colposcopy if ASCUS+
Follow-up	
Notes	A random 20% sample of screen-negative women also underwent colposcopy. For the calculation of the accuracy indices in this meta-analysis only the adjusted-for verification bias estimates are used

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women > 25 years
Acceptable reference standard? All tests	Yes	Colposcopy with biopsies if indicated
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	A random sample of screen-negative women also underwent colposcopy
Differential verification avoided? All tests	Yes	The same RS applied in all occasions
Incorporation avoided? All tests	Yes	The RS did not include the screening tests
Reference standard results blinded? All tests	Yes	Colposcopists were not aware of screening test results
Index test results blinded? All tests	Yes	Screening tests were performed before the RS

Gravitt 2010 (*Continued*)

Relevant clinical information? All tests	Unclear	It is not clear whether cytologists had the routine information
Uninterpretable results reported? All tests	No	Inadequate or unsatisfactory specimens were not reported
Withdrawals explained? All tests	Yes	Withdrawals were explained

Hovland 2010

Clinical features and settings	Women attending 3 gynaecological clinics in Bukavu, Democratic Republic of Congo, during November and December 2003 were recruited for the study. Exclusion criteria were pregnancy, severe gynaecological bleeding, previous hysterectomy and age < 25 or > 60 years
Participants	343 women between 25 and 60 years of age (median: 37 years) in DR Congo
Study design	Cross-sectional study of women receiving conventional and LBC, HPV DNA testing and HPV E6/E7 mRNA testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy, if no lesion was seen a random biopsy was taken. Colposcopy was performed in all women
Index and comparator tests	IT: HPV testing by PCR (GP5+/6+),for the detection of 14 high-risk types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68) IT: E6/E7 mRNA testing (NASBA) for the detection of 5 high-risk types (16, 18, 31, 33, 45) or 9 high-risk types (16, 18, 31, 33, 45, 35, 51, 52, 58) CT: CC. CT: LBC
Follow-up	
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Unclear	Women aged 25-60 years attending gynaecological clinics. The reason was not specified
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies

Hovland 2010 (*Continued*)

Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	Colposcopy performed in all women
Differential verification avoided? All tests	Yes	The same RS applied in all situations
Incorporation avoided? All tests	Yes	The tests did not form part of the RS
Reference standard results blinded? All tests	Unclear	It is not clear whether the pathologist had knowledge of the test results
Index test results blinded? All tests	Unclear	It is not clear whether cytologists had knowledge of the pathology results
Relevant clinical information? All tests	Unclear	It is not clear what information was given to the cytologists
Uninterpretable results reported? All tests	Yes	Nine Pap (2.6%) and 14 liquid-based smears (4.1%) were assessed as unsatisfactory
Withdrawals explained? All tests	Yes	Histology was unsatisfactory in 30 cases (8.7%), and these cases were left out of the overall statistical calculations

Iftner 2015

Clinical features and settings	Women 30-60 years of age who were undergoing routine cervical screening at 3 German centres, in Tübingen, Saarbrücken, and Freiburg, were invited to participate in the study. Exclusion criteria were hysterectomy or destructive therapy of the cervix, pregnancy, an abnormal cytological result within the past 6 months, HIV infection, and organ transplantation
Participants	9451 women attending routine screening
Study design	Cross-sectional study of women receiving conventional and LBC, HPV DNA testing and HPV E6/E7 mRNA testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ or CIN 3+ RS: colposcopy with colposcopically-directed biopsy (if required) for screen-positives and 3.6% of screen-negatives

Iftner 2015 (*Continued*)

Index and comparator tests	IT: HPV testing by HC2 (1 pg/mL), referred for colposcopy if positive IT: E6/E7 mRNA testing by Aptima HPV assay, referred for colposcopy if positive CT: LBC, referred for colposcopy if LSIL+
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women 30-60 years of age who were undergoing routine cervical screening
Acceptable reference standard? All tests	Yes	Colposcopy with biopsies if indicated
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	A random sample of screen negative women also underwent colposcopy
Differential verification avoided? All tests	Yes	The same RS applied in all situations
Incorporation avoided? All tests	Yes	The tests did not form part of the RS
Reference standard results blinded? All tests	Yes	All LBC-positive samples and samples with abnormal histological findings were collected by the respective clinical departments, and a blinded review was performed by independent external experts, for quality control
Index test results blinded? All tests	Yes	Screening tests were performed before the RS
Relevant clinical information? All tests	Unclear	It is not clear what information was given to the cytologists
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

Kulasingam 2002

Clinical features and settings	Women aged 18-50 years without hysterectomy, chronic immune suppression or treatment for CIN, presenting for annual examinations at planned parenthood clinics
Participants	4075 women (mean age 25) in Washington State, USA
Study design	Cross-sectional study of women receiving LBC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 3+ RS: colposcopically-directed biopsy, if colposcopy was normal random biopsies were taken. Only for screen-positives and 7% of screen-negatives
Index and comparator tests	IT: HPV testing by PCR (MY09, MY11, HMB01) for the detection of 18 high-risk HPV types (16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 55, 56, 58, 59, 68, 73, 82, 84). Referred for colposcopy and biopsies if positive IT: HPV testing by HC2 (Digene) only for the last 1150 women. Positivity threshold at 1 pg/mL. Referred for colposcopy + biopsies if positive CT: LBC. Referred for colposcopy if ASCUS+
Follow-up	
Notes	7% of women with negative results also underwent colposcopy. For the calculation of the accuracy indices in this meta-analysis only the corrected estimates (adjusted for verification bias and loss to follow-up) are used

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women 18 to 50 years old attending for annual routine examinations
Acceptable reference standard? All tests	Yes	Colposcopy with biopsies
Acceptable delay between tests? All tests	Yes	Both tests performed at the same visit
Partial verification avoided? All tests	Yes	A random sample of women (7%) with negative tests also received colposcopy
Differential verification avoided? All tests	Yes	The RS was the same for all women
Incorporation avoided? All tests	Yes	Cytology and HPV testing were not included in the RS
Reference standard results blinded? All tests	Yes	Pathologists had no knowledge of clinical data

Kulasingam 2002 (*Continued*)

Index test results blinded? All tests	Yes	The RS was applied after the screening tests were reported
Relevant clinical information? All tests	No	Cytologists had no clinical information
Uninterpretable results reported? All tests	Yes	Inadequate or unsatisfactory results were reported
Withdrawals explained? All tests	Yes	Withdrawals were explained

Labani 2014

Clinical features and settings	All ever-married women aged 30 to 59 years were targeted for screening. Women who had undergone a total hysterectomy, or who had been diagnosed with cancer or precancer, were excluded from the study. Menstruating women were excluded temporarily. Pregnant women were eligible to participate in the study 12 weeks after the end of their pregnancy
Participants	5032 women from Uttar Pradesh, India. The mean age of all women screened was 37.9 (SD 7.5) years
Study design	Cross-sectional study of women receiving CC, clinician-collected HPV testing, self-collected HPV testing and VIA
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ or CIN 3+ RS: colposcopically-directed biopsy, colposcopy only if no lesion was seen. Only for screen-positives
Index and comparator tests	IT: HPV testing by care HPV, positivity threshold at 1 pg/mL. Referred for colposcopy if positive IT: HPV self-testing by vaginal care HPV, positivity threshold at 1 pg/mL. Referred for colposcopy if positive IT: VIA, referred for colposcopy if positive CT: CC. Referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women aged 30-59

Labani 2014 (*Continued*)

Acceptable reference standard? All tests	Yes	Colposcopy with or without directed biopsy
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	No	Only screen-positives had colposcopy
Differential verification avoided? All tests	Yes	The same RS applied in all cases
Incorporation avoided? All tests	Yes	Cytology and HPV testing were not included in the RS
Reference standard results blinded? All tests	Unclear	Relevant information was not given
Index test results blinded? All tests	Unclear	Relevant information was not given
Relevant clinical information? All tests	Unclear	Unclear what information was given to the cytologists
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

Li 2009

Clinical features and settings	Women aged 15-69, mentally and physically competent, married, non-pregnant without a hysterectomy were contacted at home by village doctors
Participants	2562 women from three provinces (Shanxi, Lianoning, Guangdong) in China
Study design	Cross-sectional study of women receiving LBC, HPV testing, VIA, screening colposcopy and fluorescence spectroscopy
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy, if colposcopy was normal random biopsies were taken. Only for screen-positives
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold at 1 pg/mL. Referred for colposcopy + biopsies if positive IT: VIA, Rreferred for colposcopy + biopsies if positive IT: fluorescence spectroscopy, referred for colposcopy + biopsies if positive

	IT: screening colposcopy, referred for colposcopy + biopsies if positive CT: LBC (AutoCyte). Referred for colposcopy + biopsies if LSIL+
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women 15-59 years old
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies and random biopsies if colposcopy was normal
Acceptable delay between tests? All tests	Yes	Both tests taken at the same visit
Partial verification avoided? All tests	Yes	Women received the RS not only if they had positive cytology or HPV test, but also if they had positive VIA, spectroscopy or screening colposcopy
Differential verification avoided? All tests	Yes	The same RS was applied in the case of positive cytology and positive HPV test
Incorporation avoided? All tests	Yes	The RS was not composed of the LBC and HPV tests
Reference standard results blinded? All tests	No	Doctors performing the final colposcopy were aware of screening results
Index test results blinded? All tests	Yes	The RS was applied after the screening tests were reported
Relevant clinical information? All tests	Unclear	It is not clear whether clinical information was given to the cytologists
Uninterpretable results reported? All tests	Yes	All results including inadequate specimens were reported
Withdrawals explained? All tests	Yes	Withdrawals were explained

Mahmud 2012

Clinical features and settings	Women residing in a suburb of Kinshasa, Democratic Republic of Congo. Women were eligible if they were ≥ 30 years and had an intact uterus but were not pregnant
Participants	1528 women in DR Congo
Study design	Cross-sectional study of women receiving CC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy, in 20% of women where colposcopy was normal random biopsies were taken. Colposcopy was performed in all women
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold at 1 pg/mL IT: HPV testing by HC2+4 (Digene), positivity threshold at 1 pg/mL CT: CC
Follow-up	
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Unscreened women 30 or older
Acceptable reference standard? All tests	Yes	Colposcopy and directed biopsy
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	Colposcopy was performed in all women
Differential verification avoided? All tests	Yes	The same RS was used in all situations
Incorporation avoided? All tests	Yes	The screening tests did not form part of the RS
Reference standard results blinded? All tests	Yes	Pathologists and colposcopists were not aware of the screening test results
Index test results blinded? All tests	Unclear	Cytopathologists were blinded to the results of the colposcopy and the HPV tests, but unclear if blinded to results of the pathology

Mahmud 2012 (*Continued*)

Relevant clinical information? All tests	Unclear	It is not clear what information was given to the cytologists
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

McAdam 2010a

Clinical features and settings	A pilot study recruited women aged 30-50 years in 2006 by poster and flier advertisement, radio publicity, and nurse "awareness" visits to villages round Port Vila, Efate Island, Vanuatu. Women with a history of gynaecological surgery were excluded
Participants	499 apparently healthy Ni-Vanuatu women (mean age 39.3 years)
Study design	Cross-sectional study of women receiving CC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy. Colposcopy was performed in all women
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold at 1 pg/mL IT: VIA IT: VILI CT: LBC
Follow-up	
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Apparently healthy women 30-50 years old
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsy
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	Colposcopy was performed in all women

McAdam 2010a (*Continued*)

Differential verification avoided? All tests	Yes	The same RS applied in all situations
Incorporation avoided? All tests	Yes	The screening tests did not form part of the RS
Reference standard results blinded? All tests	Yes	The colposcopists and pathologists were not aware of the screening test results
Index test results blinded? All tests	Yes	All cytology and histology examinations were blinded to the clinical and HPV findings
Relevant clinical information? All tests	Unclear	It is not clear what information was given to the cytologists
Uninterpretable results reported? All tests	Yes	Rates of unsatisfactory results were given
Withdrawals explained? All tests	Unclear	Not clear whether all women with an indication for LLETZ had the procedure

McAdam 2010b

Clinical features and settings	A pilot study recruited women aged 30-50 years in 2006 by poster and flier advertisement, radio publicity, and nurse "awareness" visits to villages round Port Vila, Efate Island, Vanuatu. Women with a history of gynaecological surgery were excluded
Participants	512 apparently healthy Ni-Vanuatu women (mean age 38.36 SD 5.6)
Study design	Cross-sectional study of women receiving CC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: LLETZ only in screen-positive women
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold at 1 pg/mL CT: LBC
Follow-up	
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Apparently healthy women 30-50 years old

McAdam 2010b (*Continued*)

Acceptable reference standard? All tests	Yes	LLETZ
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	No	LLETZ was performed only in screen-positives
Differential verification avoided? All tests	Unclear	The same RS applied in all situations
Incorporation avoided? All tests	Unclear	The screening tests did not form part of the RS
Reference standard results blinded? All tests	Unclear	The colposcopists and pathologists were not aware of the screening test results
Index test results blinded? All tests	Unclear	All cytology and histology examinations were blinded to the clinical and HPV findings
Relevant clinical information? All tests	Unclear	It is not clear what information was given to the cytologists
Uninterpretable results reported? All tests	Unclear	Rates of unsatisfactory results were given
Withdrawals explained? All tests	Yes	Number of women who did not consent to LLEZT was given

Monsonego 2011

Clinical features and settings	From April 2008-February 2009, women aged 20-65 years who were seen for their annual exam in 17 private gynaecology practices in Paris, France, were invited to participate in this voluntary screening. Women were not eligible if they had undergone total hysterectomy, were pregnant or had an abnormal cytology in the past 6 months
Participants	4429 women in Paris, France
Study design	Cross-sectional study of women receiving CC, HPV DNA testing and HPV mRNA testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsies, for screen-positives and a random sample 14% of screen-negatives. If colposcopy was negative random biopsies were taken

Monsonego 2011 (*Continued*)

Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold at 1 pg/mL. Referred for colposcopy if positive IT: HPV mRNA testing by Aptima (Gen-Probe). Referred for colposcopy if positive CT: LBC. Referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women 20-65 years old attending routine screening
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	Colposcopy was performed on a random sample of screen-negatives
Differential verification avoided? All tests	Yes	The same RS applied in all situations
Incorporation avoided? All tests	Yes	The screening tests did not form part of the RS
Reference standard results blinded? All tests	No	Histopathologists were not blinded to cytology results
Index test results blinded? All tests	Yes	Colposcopy was performed after the tests were reported
Relevant clinical information? All tests	Unclear	It is unclear what information was given to the pathologists
Uninterpretable results reported? All tests	Yes	Reported
Withdrawals explained? All tests	Yes	Explained

Moy 2010

Clinical features and settings	Screening activities were conducted in 4 women's and children's hospitals in 3 provinces (Shanxi, Jiangxi, and Gansu) in China. From 2003-2005, women who were 30-49 years of age were eligible. In 2006, women who were 30-54 years of age were eligible. For all screening years, women were eligible if they were married or reported previous sexual activity; had no clinical suspicion of pregnancy (last menstrual period began < 5 weeks previously in non-menopausal women); were able to give informed consent; had no reported history of CIN, cancer of cervix, or hysterectomy; had no debilitating disease (physically unable to undergo study procedures); and had no reported history of cervical cancer screening
Participants	9057 women (mean age 39) in China
Study design	Cross-sectional study of women receiving LBC, HPV testing, VIA and VILI
Target condition and reference standard(s)	TC: histologically-confirmed CIN 3+ RS: colposcopically-directed biopsy. The criteria for referral varied by year. In 2003 and 2005, if a woman was VIA- or VILI-positive, she was referred for colposcopy. In 2004 and 2006, all women had colposcopy regardless of the results of VIA and VILI. Directed biopsy was performed on any visible lesion. If a woman was VIA- or VILI-negative, but with either a Pap test of ASC-H, AGUS, LSIL or higher, or positive for HR-HPV DNA by HC2 testing, she was recalled after 2 weeks for colposcopy and received four-quadrant biopsy
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold at 1 pg/mL IT: VIA IT: VILI CT: LBC. Positivity threshold LSIL+
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women 30-54 without prior screening
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	Colposcopy was not limited to screen-positives

Moy 2010 (*Continued*)

Differential verification avoided? All tests	Yes	The same RS applied in all situations
Incorporation avoided? All tests	Yes	The tests did not form part of the RS
Reference standard results blinded? All tests	Yes	The pathologists had no knowledge of the test results
Index test results blinded? All tests	Yes	Colposcopy was performed after the test results were reported
Relevant clinical information? All tests	Unclear	It is not clear what information was given to the cytologists
Uninterpretable results reported? All tests	Yes	Reported
Withdrawals explained? All tests	Unclear	It is not clear whether all women that should have had colposcopy attended

Naucler 2009

Clinical features and settings	Women aged 32-38 attending the Swedish Cervical Cancer Screening Programme
Participants	6257 women attending cervical screening in 5 Swedish cities (Stockholm, Uppsala, Malmo, Umea, Gothenburg)
Study design	RCT of HPV testing and CC versus CC alone. Only the cross-sectional results of the first screening round, from the experimental arm only were included in this meta-analysis
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy, if colposcopy was normal, random biopsies were taken
Index and comparator tests	IT: HPV testing by PCR (GP5+, GP6+) for the detection of 14 high-risk HPV types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68). If cytology was negative, women with a positive HPV test were referred for the RS if a second HPV test 12 months later was also type-specific positive CT: CC. For the calculation of the accuracy indices an abnormal result was considered any smear showing ASCUS+. However in 4 cities the option of a repeat smear was given after a result of ASCUS
Follow-up	73 of 328 women who were HPV-positive and CC-negative in the first exam did not return for a second exam one year later

Notes	Apart from the histology specimens taken inside the protocol colposcopy, the study had access to histology specimen taken outside the protocol through the national pathology registry

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	No	Only women aged 32-38 years were included
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies and random biopsies if colposcopy was normal
Acceptable delay between tests? All tests	Yes	Both tests taken at the same visit
Partial verification avoided? All tests	No	Only women with positive tests were referred for colposcopy
Differential verification avoided? All tests	Yes	The same RS was applied in the case of positive cytology and persistent type-specific positive HPV test
Incorporation avoided? All tests	Yes	The RS was not composed of the index and comparator tests
Reference standard results blinded? All tests	Yes	The women and the clinical personnel were not aware of the screening test results
Index test results blinded? All tests	Yes	The RS was applied after the screening tests were reported
Relevant clinical information? All tests	Unclear	It is not clear whether clinical information was given to the cytologists
Uninterpretable results reported? All tests	Yes	All results including inadequate specimens were reported
Withdrawals explained? All tests	Yes	Withdrawals were explained

Nieves 2013

Clinical features and settings	The study was conducted in rural Mexico. Women aged 30-50 years, non-pregnant, with no history of hysterectomy or pelvic irradiation and varied histories of screening, participated
Participants	2049 women in rural Mexico, median age 39.2 years (range, 30-50 years)
Study design	Cross-sectional study of women receiving LBC, HPV DNA testing, HPV mRNA testing, self-HPV DNA testing, self-HPV mRNA testing and VIA
Target condition and reference standard(s)	TC: histologically-confirmed CIN 3+ RS: colposcopically-directed biopsy for all women, if colposcopy was normal random biopsies were taken. Colposcopy was performed to screen-positives only
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold at 1 pg/mL. Referred for colposcopy if positive IT: HPV mRNA testing by Aptima (Gen-Probe). Referred for colposcopy if positive IT: self-HPV testing by HC2 (Digene), positivity threshold at 1 pg/mL. Referred for colposcopy if positive IT: self-HPV mRNA testing by Aptima (Gen-Probe). Referred for colposcopy if positive IT: VIA CT: LBC. Referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women 30-50 years
Acceptable reference standard? All tests	Yes	Colposcopy and biopsies
Acceptable delay between tests? All tests	Yes	Performed on the same visit
Partial verification avoided? All tests	No	Only screen-positives had colposcopy
Differential verification avoided? All tests	Yes	The same RS applied in all situations
Incorporation avoided? All tests	Yes	The tests were not part of the RS
Reference standard results blinded? All tests	Unclear	Not clear whether pathologists had access to test results

Nieves 2013 (*Continued*)

Index test results blinded? All tests	Yes	Technicians and cytologists were not aware of the other test results
Relevant clinical information? All tests	Unclear	Not clear what information was given to cytologists
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

Pan 2003

Clinical features and settings	Previously unscreened women 35-45 years old, with no history of pelvic radiation or hysterectomy residing in the Shanxi Province in China
Participants	1993 women (mean age 39.1) from rural China
Study design	Cross-sectional study of women receiving self-HPV testing, LBC, direct HPV testing and VIA
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: Colposcopically-directed biopsy for all women, if colposcopy was normal random biopsies were taken. Colposcopy was performed on all women
Index and comparator tests	IT: HPV testing by HC2 direct sampling (Digene), positivity threshold at 1 pg/mL IT: HPV testing by HC2 self sampling (Digene), positivity threshold at 1 pg/mL CT: LBC. For the calculation of the accuracy indices an abnormal result was considered any smear showing ASCUS+
Follow-up	
Notes	Some data were obtained from the publication by Belinson 2001

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	No	Very limited age spectrum of previously unscreened women
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit

Pan 2003 (*Continued*)

Partial verification avoided? All tests	Yes	The RS was applied to all women
Differential verification avoided? All tests	Yes	The same RS was performed on all occasions
Incorporation avoided? All tests	Yes	The screening tests were not part of the RS
Reference standard results blinded? All tests	Yes	Pathologists had no knowledge of the screening tests
Index test results blinded? All tests	Yes	Cytologists were not aware of the final diagnosis
Relevant clinical information? All tests	Unclear	It is not clear whether cytologists had the routine clinical information
Uninterpretable results reported? All tests	Yes	The numbers of inadequate smears were given
Withdrawals explained? All tests	Yes	It does not seem as if there were any withdrawals

Paraskevaidis 2001

Clinical features and settings	Women 17-79 years old without prior history of cervical pathology attending the outpatient clinics of a university hospital for routine screening
Participants	977 women (mean age 38) in Ioannina, Greece
Study design	Cross-sectional study of women receiving both CC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy. Colposcopy only if no lesion was seen. Only for screen-positives
Index and comparator tests	IT: HPV testing by PCR (MY09, MY11) for the detection of 11 high-risk types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58). Referred for colposcopy if positive CT: CC. Referred for colposcopy if reactive cellular changes+
Follow-up	
Notes	For the calculation of CC accuracy indices the thresholds of ASCUS+ and LSIL+ were used in this meta-analysis. Since women with reactive cellular changes also underwent colposcopy, this limits verification bias

Table of Methodological Quality

Paraskevaidis 2001 (*Continued*)

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women 17-79 years old attending routine screening
Acceptable reference standard? All tests	Yes	Colposcopy with biopsies if necessary
Acceptable delay between tests? All tests	Yes	Both tests performed at the same visit
Partial verification avoided? All tests	Yes	Women with less than ASCUS were also referred for the RS
Differential verification avoided? All tests	Yes	The same RS was used in all occasions of positive screening test
Incorporation avoided? All tests	Yes	The RS was not composed of the screening tests
Reference standard results blinded? All tests	No	Colposcopists were aware of the results
Index test results blinded? All tests	Yes	RS performed after the screening tests were reported
Relevant clinical information? All tests	Unclear	It is not clear whether relevant clinical information was revealed to the cytologists
Uninterpretable results reported? All tests	No	Inadequate and unsatisfactory results were not reported
Withdrawals explained? All tests	Yes	Withdrawals were explained

Petry 2003

Clinical features and settings	Women > 30 years old attending urban, suburban or rural office-based gynaecology practices in Hannover and Tuebingen for routine screening. No hysterectomy, not pregnant, no history of atypical cytology or CIN in the last year
Participants	8101 women (mean age 42.7) from Germany
Study design	Cross-sectional study of women receiving both CC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsies. If colposcopy was negative a biopsy might have not been taken. Screen-positives and a random 3.4% sample of screen-negatives underwent

Petry 2003 (*Continued*)

	colposcopy
Index and comparator tests	IT: HPV testing by HC2 (DIgene), positivity threshold 1 pg/mL, referred for colposcopy if positive CT: CC, referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women > 30 years old attending for routine screening
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	3.4% of screen-negatives were also verified
Differential verification avoided? All tests	Yes	The same RS was applied in each case
Incorporation avoided? All tests	Yes	Screening tests were not part of the RS
Reference standard results blinded? All tests	Unclear	It is not clear whether colposcopists and pathologists had knowledge of the screening test results
Index test results blinded? All tests	Yes	RS was performed after the screening tests were reported
Relevant clinical information? All tests	Yes	Cytologists were given routine clinical information
Uninterpretable results reported? All tests	Unclear	Unsatisfactory smears were reported as Pap IIw, which also included ASCUS. There was no mention of unsatisfactory HPV testing specimens
Withdrawals explained? All tests	Yes	Were explained

Qiao 2008

Clinical features and settings	Women aged 30-54 living in rural villages in Shanxi Province, China. Non pregnant, no history of CIN, pelvic radiation or hysterectomy
Participants	2530 women (mean age 43.4) from rural China
Study design	Cross-sectional study of women receiving self-HPV testing, direct-HPV testing, LBC and VIA
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy for all women. Colposcopy only if no lesion was seen
Index and comparator tests	IT: HPV testing by HC2 (Digene). For the calculation of the accuracy indices the test was considered positive at the threshold of 1 pg/mL IT: HPV testing by Care HPV test (self-sampling). For the calculation of the accuracy indices the test was considered positive at the threshold of 1 pg/mL IT: HPV testing by Care HPV test (directed sampling). For the calculation of the accuracy indices the test was considered positive at the threshold of 1 pg/mL CT: LBC (Surepath). For the calculation of the accuracy indices the test was considered positive at the threshold of LSIL+
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women aged 30-54
Acceptable reference standard? All tests	Yes	Colposcopy with directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	All women underwent colposcopy and biopsies
Differential verification avoided? All tests	Yes	The same RS was applied in all circumstances
Incorporation avoided? All tests	Yes	The screening tests were not part of the RS
Reference standard results blinded? All tests	Unclear	It is not clear whether pathologists were aware of the screening test results

Qiao 2008 (*Continued*)

Index test results blinded? All tests	Yes	Cytologists were not aware of the histology results
Relevant clinical information? All tests	Unclear	It is not clear whether cytologists had the routine clinical information
Uninterpretable results reported? All tests	Yes	Unsatisfactory results were reported
Withdrawals explained? All tests	Yes	Withdrawals were explained

Ronco 2006

Clinical features and settings	Women > 35 years attending routine cervical screening without hysterectomy, without treatment for CIN in the last 5 years, non pregnant
Participants	16,706 women (median age 45) in Italy
Study design	RCT of HPV testing and LBC versus CC. Only the cross-sectional results of the first screening round, from the experimental arm only were included in this meta-analysis
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy only for screen-positives. Colposcopy only if no lesion was seen
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold at 1 pg/mL or 2 pg/mL. Referred for colposcopy if positive CT: LBC (Thinprep). Referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women >35 years attending routine screening
Acceptable reference standard? All tests	Yes	Colposcopy with biopsy if required
Acceptable delay between tests? All tests	Yes	At the same visit

Ronco 2006 (*Continued*)

Partial verification avoided? All tests	No	RS not applied if both tests were negative
Differential verification avoided? All tests	Yes	The RS was the same in all situations
Incorporation avoided? All tests	Yes	The RS was not composed of the index and comparator tests
Reference standard results blinded? All tests	No	Colposcopists were aware of the test results
Index test results blinded? All tests	Yes	The RS was applied after the screening tests were reported
Relevant clinical information? All tests	Unclear	There was not sufficient information
Uninterpretable results reported? All tests	Yes	All results were reported
Withdrawals explained? All tests	Yes	Withdrawals were explained

Salmeron 2003

Clinical features and settings	Women attending the cervical cancer screening services in Morelos state, Mexico. Non pregnant, no hysterectomy, without history of CIN 2+
Participants	7732 women (mean age 42.5) in Morelos, Mexico
Study design	Cross-sectional study of women receiving both CC, self-collected HPV testing and direct HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsy only for screen-positives
Index and comparator tests	IT: HPV testing by HC2 (Digene) direct sampling, positivity threshold 1 pg/mL, referred for colposcopy if positive IT: HPV testing by HC2 (Digene) self-sampling, positivity threshold 1 pg/mL, referred for colposcopy if positive CT: CC, referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Salmeron 2003 (*Continued*)

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women attending a cervical cancer screening programme
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	No	Only screen-positives received the RS
Differential verification avoided? All tests	Yes	Same RS in all cases
Incorporation avoided? All tests	Yes	Screening tests not part of the RS
Reference standard results blinded? All tests	No	Colposcopists were aware of screening tests
Index test results blinded? All tests	Yes	RS performed after screening tests were reported
Relevant clinical information? All tests	Unclear	It is not clear what information was available to the cytologists
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

Sankaranarayanan 2004a

Clinical features and settings	Opportunistic recruitment of healthy asymptomatic women aged 25-65, with an intact uterus and no previous history of cervical neoplasia from three different locations in India. None had been previously screened
Participants	11,518 women from Kolkata, Muumbai and Trivandum, India
Study design	Cross-sectional study of women receiving both CC, HPV testing, VIA and VILI
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsies for all women, if colposcopy was negative biopsies were not taken

Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold 1 pg/mL IT: VIA IT: VILI CT: CC, for the calculation of the accuracy indices the threshold of LSIL+ was used
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women aged 25-65 without history of cervical neoplasia
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsy
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	Yes	All women received the RS
Differential verification avoided? All tests	Yes	The same RS was applied to all cases
Incorporation avoided? All tests	Yes	The screening tests were not part of the RS
Reference standard results blinded? All tests	Yes	Colposcopists were not aware of screening test results
Index test results blinded? All tests	Yes	Laboratory personnel were not aware of RS results
Relevant clinical information? All tests	Unclear	It is not clear what information was given to the cytologists
Uninterpretable results reported? All tests	Yes	Were reported
Withdrawals explained? All tests	Yes	Were explained

Sarian 2005

Clinical features and settings	Women aged 18-60, with an intact uterus, no history of abnormal Pap test in the last year, not under treatment for genital warts, not immunosuppressed, were invited by the local health units to attend for screening in Brazil and Argentina
Participants	10,138 women (mean age 37.9) from the cities of Campinas, Sao Paolo, Porto Alegre (Brazil) and Buenos Aires (Argentina)
Study design	Cross-sectional study of women receiving CC, HPV testing, VIA and VILI
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsies. If colposcopy was negative biopsies were not taken unless the smear was HSIL. Colposcopy was performed in screen-positives and in a random 5% of screen-negatives
Index and comparator tests	IT: HPV testing by HC2 (Digene), referred for colposcopy if > 1 pg/mL IT: VIA, referred for colposcopy if positive IT: VILI, referred for colposcopy if positive CT: CC, referred for colposcopy if LSIL+
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women 18-60 years old
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Screening tests performed at the same visit, colposcopy 45 days later
Partial verification avoided? All tests	Yes	5% random sample of screen-negatives had colposcopy plus the VIA and VILI positives
Differential verification avoided? All tests	Yes	The same RS in all occasions
Incorporation avoided? All tests	Yes	Screening tests not part of the RS
Reference standard results blinded? All tests	No	Colposcopists and pathologists were aware of test results

Sarian 2005 (*Continued*)

Index test results blinded? All tests	Yes	RS performed later
Relevant clinical information? All tests	Unclear	It is not clear what information was available to the cytologists
Uninterpretable results reported? All tests	No	Not reported
Withdrawals explained? All tests	No	Not explained

Schneider 2000

Clinical features and settings	Women 18-70 years old visiting the offices of 10 private gynaecologists in East Thirungia, Germany for screening. Non-pregnant, no history of cervical conisation, no hysterectomy or CIN, no atypical cytology in the last year
Participants	4761 women (median age 35) from Germany
Study design	Cross-sectional study of women receiving both CC, HPV testing and screening colposcopy
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsies. If colposcopy was negative random biopsies were taken. Only for screen-positive women
Index and comparator tests	IT: HPV testing by PCR (GP) for the detection of 14 high-risk types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 68). Referred for colposcopy if positive CT: CC. Referred for colposcopy if LSIL+
Follow-up	A second screening round was done for women negative for all three tests 4-8 months later
Notes	

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Women 18-70 years old
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	No	8 months elapsed between screening tests and verification

Schneider 2000 (*Continued*)

Partial verification avoided? All tests	Yes	The presence of a third screening test (colposcopy) and the referral of the screen-positives for verification limits the problem of partial verification
Differential verification avoided? All tests	Yes	The same RS was used for all tests
Incorporation avoided? All tests	Yes	The screening tests did not form part of the RS
Reference standard results blinded? All tests	Yes	Pathologists were not aware of the screening test results
Index test results blinded? All tests	Yes	The RS was performed later
Relevant clinical information? All tests	Unclear	It is not clear whether cytologists were given the routine clinical information
Uninterpretable results reported? All tests	Yes	The numbers were given
Withdrawals explained? All tests	Yes	The reasons for withdrawals were explained

Shipitsyna 2011

Clinical features and settings	Consecutive women aged 30-65 years, not pregnant and with no history of treatment for CIN grade 2 and higher (CIN 2+), receiving routine gynaecological care at the outpatient department of the D.O. Ott Research Institute of Obstetrics and Gynaecology, St. Petersburg, Russia from June 2008-April 2009 were enrolled
Participants	823 women (mean age 39.5 + 8.4 years) in St. Petersburg, Russia
Study design	Cross-sectional study of women receiving CC and HPV testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsies only for screen-positives
Index and comparator tests	IT: HPV testing by HC2 (Digene), referred for colposcopy if > 1 pg/mL CT: CC, referred for colposcopy if LSIL+
Follow-up	
Notes	

Table of Methodological Quality

Shipitsyna 2011 (*Continued*)

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women 30-65 years old
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	No	Only screen-positives were verified
Differential verification avoided? All tests	Yes	The same RS applied in all situations
Incorporation avoided? All tests	Yes	The screening tests did not form part of the RS
Reference standard results blinded? All tests	Unclear	It is not clear what information was available to the pathologists and colposcopists
Index test results blinded? All tests	Yes	Colposcopy was performed after the tests were reported
Relevant clinical information? All tests	Unclear	It is not clear what information was available to the cytologists
Uninterpretable results reported? All tests	Yes	In six (0.7%) women, samples were initially graded unsatisfactory for cytological assessment. Those women were called for repeated cytology and were tested negative
Withdrawals explained? All tests	Yes	44 women did not return for colposcopy

Syrjanen 2002

Clinical features and settings	Consecutive women attending 6 different outpatient clinics in 3 New Independent States of the former Soviet Union. The study focused on 3 target populations 1. women participating locally organised screening programmes for cervical cancer, 2. those attending regular gynaecology clinics for various indications, 3. those attending STD clinics
Participants	3175 women (mean age 32.7) from Belarus, Russia and Latvia
Study design	Cross-sectional study of women receiving both CC and HPV testing

Syrjanen 2002 (*Continued*)

Target condition and reference standard(s)	TC: histologically-confirmed CIN 3+ RS: colposcopically-directed biopsies, if colposcopy was normal a biopsy might have not been taken. Only for screen-positives
Index and comparator tests	IT: HPV testing by HC2, positivity threshold 1 pg/mL, referred for colposcopy if positive CT: CC, referred for colposcopy if LSIL+
Follow-up	
Notes	For the calculation of the accuracy indices the results from primary screening and not from re-screening were used

Table of Methodological Quality

Item	**Authors' judgement**	**Description**
Representative spectrum? All tests	Yes	Most of the women were attending a local cervical screening programme
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	No	Only screen-positives were verified
Differential verification avoided? All tests	Yes	The same RS was performed in all cases
Incorporation avoided? All tests	Yes	Screening tests were not part of the RS
Reference standard results blinded? All tests	Unclear	It is not clear what information was available to the colposcopists and pathologists
Index test results blinded? All tests	Yes	The RS was performed after the screening tests were reported
Relevant clinical information? All tests	Unclear	It is not clear what information was available to the cytologists
Uninterpretable results reported? All tests	No	Not reported
Withdrawals explained? All tests	Yes	Were explained

Wu 2010

Clinical features and settings	The Shenzhen Cervical Cancer Screening Trial I (SHENCCAST I) took place in the city of Shenzhen, Guangdong Province, in southern China. Women were eligible if they were 25 to 59 years of age, were non pregnant, had had no cervical cancer screening for at least 3 years, had no prior hysterectomy, and had no prior pelvic radiation
Participants	2098 women in southern China (mean age 35)
Study design	Cross-sectional study of women receiving CC, HPV DNA testing and HPV mRNA testing
Target condition and reference standard(s)	TC: histologically-confirmed CIN 2+ RS: colposcopically-directed biopsies. If colposcopy was negative random biopsies were taken. Only for screen-positives
Index and comparator tests	IT: HPV testing by HC2 (Digene), positivity threshold at 1 pg/mL. Referred for colposcopy if positive IT: HPV mRNA testing by Aptima (Gen-Probe). Referred for colposcopy if positive CT: LBC. Referred for colposcopy if ASCUS+
Follow-up	
Notes	

Table of Methodological Quality

Item	Authors' judgement	Description
Representative spectrum? All tests	Yes	Women aged 25-59
Acceptable reference standard? All tests	Yes	Colposcopically-directed biopsies
Acceptable delay between tests? All tests	Yes	Performed at the same visit
Partial verification avoided? All tests	No	Only screen-positives were referred
Differential verification avoided? All tests	Yes	The same RS applied in all situations
Incorporation avoided? All tests	Yes	The screening tests did not form part of the RS
Reference standard results blinded? All tests	Yes	Pathologists were blinded to HPV test results
Index test results blinded? All tests	Yes	Cytologists were blinded to the pathology and the HPV tests

Wu 2010 (*Continued*)

Relevant clinical information? All tests	Unclear	It is not clear what information was available to the pathologists
Uninterpretable results reported? All tests	No	Were there any invalid results?
Withdrawals explained? All tests	Yes	95 women who were requested to return for colposcopy based on their test results did not return

ASCUS: atypical squamous cells of undetermined significance; **CC**: conventional cytology; **CIN**: cervical intraepithelial neoplasia; **CT**: comparator test; **HC2**: hybrid capture 2; **HPV**: human papillomavirus; **HSIL**: high-grade squamous intraepithelial lesion; **IT**: index test; **LBC**: liquid-based cytology; **LEEP**: loop electro-excision procedure; **LLETZ** - large loop excision of the transformation zone; **LSIL**: low-grade squamous intraepithelial lesion; **Pap**: Papanicolaou; **PCR**: polymerase chain reactions; **RS**: reference standard; **RCT**: randomised controlled trial; **TC**: target condition; **VIA**: visual inspection with acetic acid; **VILI**: visual inspection with Lugol's iodine

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
An 2003	The reference standard was biopsy without colposcopy and the criteria for the reference standard application were undetermined
Belinson 2001	Presents data on the same participants as the study by Pan 2003
Belinson 2010a	Data on cytology were not given
Belinson 2011	Data on cytology were not given
Belinson 2012	Data on cytology were not given
Benevolo 2011a	Retrospective study on selected population with a positive HC2 test
Benevolo 2011b	Retrospective study on selected population consisting mainly of abnormal smears
Benoy 2011	The outcome was CIN 2+ detection within 24 months of the screening round
Castle 2010	The population consisted of women with abnormal (ASCUS) smears
Chao 2010	A longitudinal study of baseline HPV and cytology-negative women
Coquillard 2011	The study sample was a mix of low and high risk populations

(*Continued*)

Costa 2000	HC1 was used for HPV testing which is no longer used in clinical practice
Coste 2003	Contained only data on cytology. Referred to the included study de Cremoux 2003
Dai 2006	HPV testing not a criterion for reference standard application
de Andrade 2011	Population consisted of HIV-infected women
De Vuyst 2005	105 of 653 women included in this study were being investigated because of an abnormal Pap smear
Denny 2000	HC1 was used for HPV testing which is no longer used in clinical practice
Depuydt 2012	Accuracy given for cumulative diagnosis of CIN 2+ over the follow-up period
Diamantopoulou 2013	No appropriate reference standard
Fereccio 2003	Refers to the study by Schiffman 2000, which was also excluded because the HPV testing was not a criterion for the gold standard application
Huang 2010	Selected population with high rate of CIN2+
Idelevich 2011	Selected population with high rates of CIN2+
Junyangdikul 2013	No appropriate reference standard
Katki 2011	A retrospective study. Has as outcome measure the cumulative incidence of CIN2+ over 5 years. Data to calculate cross-sectional accuracy immediately after screening were not given. Data on how many women had colposcopy were not given
Kelesidis 2011	The sample was not representative of the population attending routine screening (archived cytological samples of women with archived histological samples)
Kim 2013	Histology given as normal or abnormal (CIN1+)
Kitchener 2011	HC2 was not a criterion for immediate colposcopy referral. A second positive test one year later was required
Kuhn 2000	HC1 was used for HPV testing which is no longer used in clinical practice
Li 2010a	HPV testing was not a criterion for colposcopy referral
Li 2010b	Not general population (30% had CIN 2+)
Longatto-Filho 2012	Refers to the study Sarian 2005
Ma 2010	Cytology results were not given

(*Continued*)

Masumoto 2003	HPV testing was performed only on 477/3000 women. The population included women under investigation for abnormal smears
Mesher 2010	Provides the longitudinal data of the Cuzick 2003 study
Monsonego 2012	Same population as in the Monsonego 2011 study
Nieminen 2004	A positive HPV test was not a criterion for the application of the reference standard. The sample was not representative of the general population
Oh 2001	Reference standard (biopsy) not applied in all screen positives. The criteria for the application of the reference standard were unclear. Colposcopy was not a part of the reference standard
Ozcan ES 2011	Unclear what the referral criteria for colposcopy were
Quincy 2012	Not appropriate gold standard
Ratnam 2000	69% of women received HCI as HPV testing and 31% HCII. HPV testing results were not presented separately for the two methods
Riethmuller 1999	The target condition was not CIN but HPV infection. CIN rates were not reported. 130/596 women were referred because of an abnormal smear
Rijkaart 2012b	Not all screen-positives would be offered verification (ie HPV-positive/cyto-negative and low-grade cyto/HPV neg)
Sankaranarayanan 2004b	Presented data on the same population as the study Sankaranarayanan 2004a, which is included in the analysis
Sankaranarayanan 2005	RCT where women received either VIA or cytology or HPV testing
Schiffman 2000	A positive HPV test was not a criterion for the application of the reference standard
Shastri 2005	Presented data on a subgroup of the population of the study Sankaranarayanan 2004a which is included in the analysis
Sherman 2003	It is a longitudinal study of the risk of CIN 3 10 years after a baseline Pap smear and HPV test. It is not a cross-sectional comparison of diagnostic accuracy. HPV testing was not a criterion for reference standard application
Siriaunkgul 2014	HPV test positivity was not a criterion for reference standard application
Surabhi 2011	HPV testing was only done on a selected subgroup of the population
Wang 2013	Severe selection bias
Womack 2000	Presented data on the same population as the study by Blumenthal 2001

(*Continued*)

Zhao 2010	A pooled analysis of individual participant data. The published studies that have contributed their participants to this paper are already included in our meta-analysis. However this paper also included data from unpublished studies

ASCUS: atypical squamous cells of undetermined significance; **CIN**: cervical intraepithelial neoplasia; **HC2**: hybrid capture 2; **HPV**: human papillomavirus;**Pap**: Papanicolaou

DATA

Presented below are all the data for all of the tests entered into the review.

Tests. Data tables by test

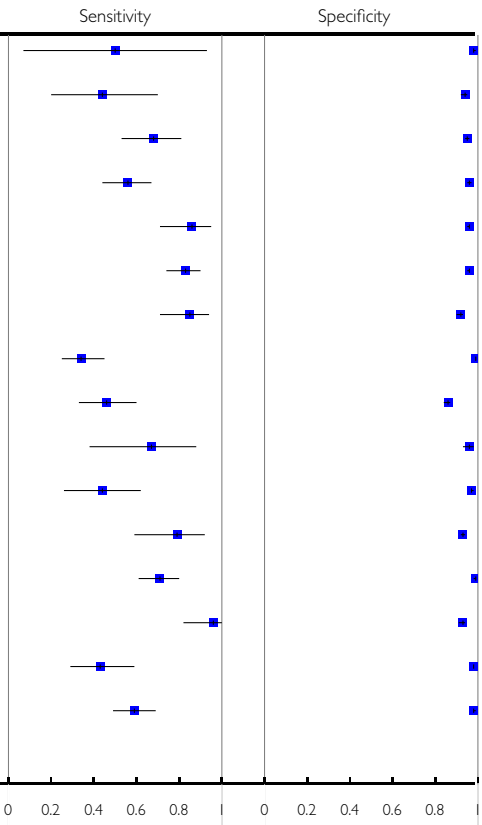
Test	No. of studies	No. of participants
1 CC (ASCUS+) for CIN2+	16	61099
2 CC (ASCUS+) for CIN3+	9	51857
3 CC (LSIL+) for CIN2+	9	41494
4 CC (LSIL+) for CIN3+	5	35648
5 LBC (ASCUS+) for CIN2+	15	82003
6 LBC (ASCUS+) for CIN3+	13	71919
7 LBC (LSIL+) for CIN2+	10	33519
8 LBC (LSIL+) for CIN3+	5	21166
9 HC2 (1pg/mL) for CIN2+	25	138230
10 HC2 (1 pg/mL) for CIN3+	19	120380
11 HC2 (2 pg/mL) for CIN2+	2	26768
12 HC2 (2 pg/mL) for CIN3+	2	26768
13 PCR (13 hr types or more) for CIN2+	6	16343
14 PCR (13 hr types or more) for CIN3+	4	14048
15 PCR (10-11 hr types) for CIN2+	2	3965
16 PCR (10-11 hr types) for CIN3+	1	2988
17 Aptima for CIN2+	3	15895
18 Aptima for CIN3+	4	17944
19 PCR (4 hr types) for CIN2+	1	1985
20 Care HPV test (0.5 pg/ml) for CIN2+	2	7044
21 Care HPV test (0.5 pg/ml) for CIN3+	2	7046
22 Cobas for CIN2+	2	11666
23 Cobas for CIN3+	2	11666
24 NASBA (5 types) for CIN2+	1	313
25 NASBA (9 types) for CIN2+	1	313
26 HC2+4 (1 pg/ml) for CIN2+	1	1352
27 HC2+4 (1 pg/ml) for CIN3+	1	1352
28 HC2 (1pg/mL) for CIN2+ no verification bias	12	53013
29 CC or LBC (ASCUS+) for CIN2+ no verification bias	8	31341
30 HC2 (1pg/mL) for CIN2+ women >30	13	69334
31 self HPV test for CIN2+	4	23474

Test 1. CC (ASCUS+) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 1 CC (ASCUS+) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Agorastos 2005	2	20	2	1272	0.50 [0.07, 0.93]	0.98 [0.98, 0.99]
Cardenas-Turanzas 2008	7	52	9	767	0.44 [0.20, 0.70]	0.94 [0.92, 0.95]
Clavel 2001	32	104	15	2130	0.68 [0.53, 0.81]	0.95 [0.94, 0.96]
Cuzick 1995	45	83	36	1821	0.56 [0.44, 0.67]	0.96 [0.95, 0.97]
Cuzick 1999	36	130	6	2816	0.86 [0.71, 0.95]	0.96 [0.95, 0.96]
Cuzick 2003	75	427	15	9841	0.83 [0.74, 0.90]	0.96 [0.95, 0.96]
de Cremoux 2003	35	141	6	1573	0.85 [0.71, 0.94]	0.92 [0.90, 0.93]
Ferreccio 2013	33	99	63	8017	0.34 [0.25, 0.45]	0.99 [0.99, 0.99]
Gravitt 2010	26	313	30	1862	0.46 [0.33, 0.60]	0.86 [0.84, 0.87]
Hovland 2010	10	11	5	279	0.67 [0.38, 0.88]	0.96 [0.93, 0.98]
Labani 2014	14	118	18	4506	0.44 [0.26, 0.62]	0.97 [0.97, 0.98]
Mahmud 2012	22	94	6	1264	0.79 [0.59, 0.92]	0.93 [0.92, 0.94]
Naucler 2009	62	84	25	6023	0.71 [0.61, 0.80]	0.99 [0.98, 0.99]
Paraskevaidis 2001	27	71	1	878	0.96 [0.82, 1.00]	0.93 [0.91, 0.94]
Petry 2003	20	156	26	7706	0.43 [0.29, 0.59]	0.98 [0.98, 0.98]
Salmeron 2003	60	127	41	7504	0.59 [0.49, 0.69]	0.98 [0.98, 0.99]

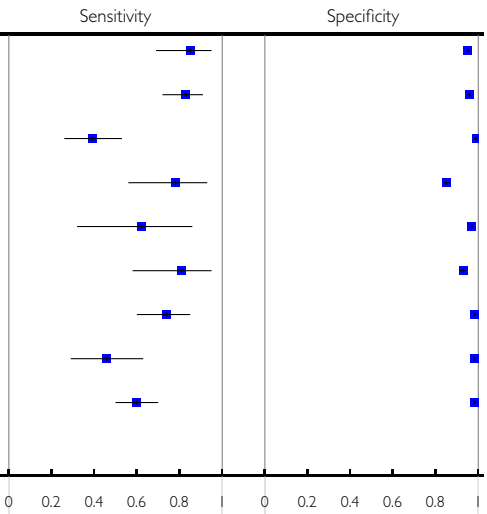


Test 2. CC (ASCUS+) for CIN3+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 2 CC (ASCUS+) for CIN3+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Cuzick 1999	29	137	5	2817	0.85 [0.69, 0.95]	0.95 [0.95, 0.96]
Cuzick 2003	57	445	12	9844	0.83 [0.72, 0.91]	0.96 [0.95, 0.96]
Ferreccio 2013	21	111	33	8047	0.39 [0.26, 0.53]	0.99 [0.98, 0.99]
Gravitt 2010	18	321	5	1887	0.78 [0.56, 0.93]	0.85 [0.84, 0.87]
Labani 2014	8	124	5	4511	0.62 [0.32, 0.86]	0.97 [0.97, 0.98]
Mahmud 2012	17	99	4	1266	0.81 [0.58, 0.95]	0.93 [0.91, 0.94]
Naucler 2009	37	109	13	6035	0.74 [0.60, 0.85]	0.98 [0.98, 0.99]
Petry 2003	17	159	20	7712	0.46 [0.29, 0.63]	0.98 [0.98, 0.98]
Salmeron 2003	56	131	37	7708	0.60 [0.50, 0.70]	0.98 [0.98, 0.99]

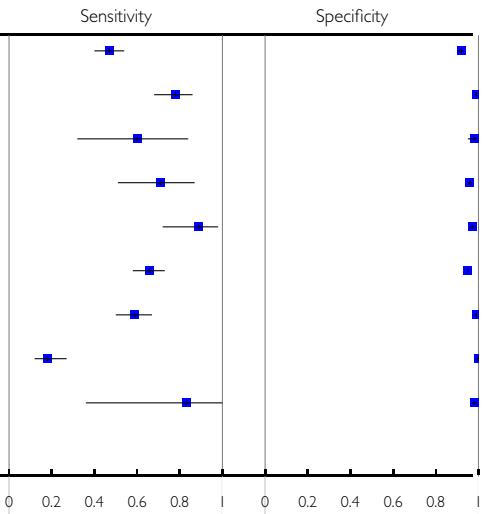


Test 3. CC (LSIL+) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 3 CC (LSIL+) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Blumenthal 2001	97	169	111	1822	0.47 [0.40, 0.54]	0.92 [0.90, 0.93]
Cuzick 2003	70	143	20	10125	0.78 [0.68, 0.86]	0.99 [0.98, 0.99]
Hovland 2010	9	7	6	283	0.60 [0.32, 0.84]	0.98 [0.95, 0.99]
Mahmud 2012	20	56	8	1302	0.71 [0.51, 0.87]	0.96 [0.95, 0.97]
Paraskevaidis 2001	25	24	3	925	0.89 [0.72, 0.98]	0.97 [0.96, 0.98]
Sankaranarayanan 2004a	109	516	57	9909	0.66 [0.58, 0.73]	0.95 [0.95, 0.95]
Sarian 2005	86	139	61	9852	0.59 [0.50, 0.67]	0.99 [0.98, 0.99]
Schneider 2000	21	21	93	4626	0.18 [0.12, 0.27]	1.00 [0.99, 1.00]
Shipitsyna 2011	5	13	1	760	0.83 [0.36, 1.00]	0.98 [0.97, 0.99]

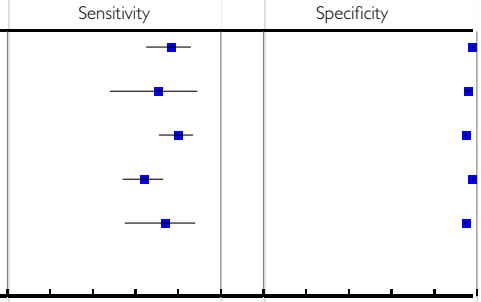


Test 4. CC (LSIL+) for CIN3+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 4 CC (LSIL+) for CIN3+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Cuzick 2003	53	160	16	10129	0.77 [0.65, 0.86]	0.98 [0.98, 0.99]
Mahmud 2012	15	61	6	1304	0.71 [0.48, 0.89]	0.96 [0.94, 0.97]
Sankaranarayanan 2004a	81	544	20	9946	0.80 [0.71, 0.87]	0.95 [0.94, 0.95]
Sarian 2005	64	161	36	9877	0.64 [0.54, 0.73]	0.98 [0.98, 0.99]
Syrjanen 2002	23	167	8	2977	0.74 [0.55, 0.88]	0.95 [0.94, 0.95]

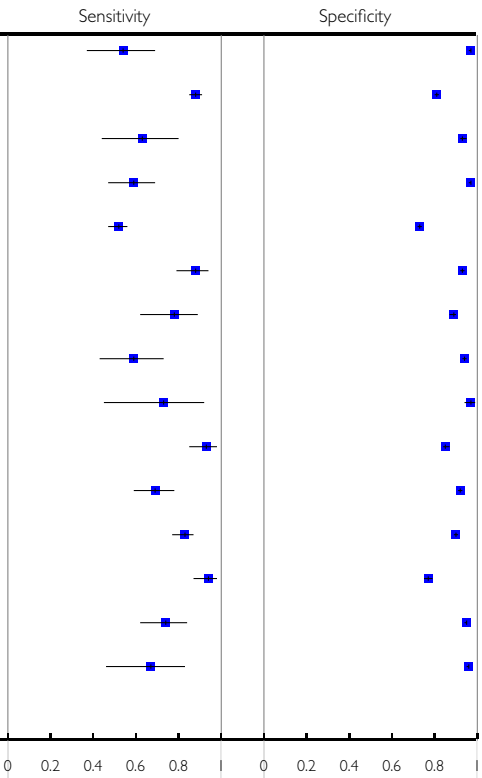


Test 5. LBC (ASCUS+) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 5 LBC (ASCUS+) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Agorastos 2015	22	119	19	3683	0.54 [0.37, 0.69]	0.97 [0.96, 0.97]
Belinson 2003	331	1523	44	6599	0.88 [0.85, 0.91]	0.81 [0.80, 0.82]
Belinson 2010	19	63	11	886	0.63 [0.44, 0.80]	0.93 [0.92, 0.95]
Bigras 2005	48	445	34	13315	0.59 [0.47, 0.69]	0.97 [0.96, 0.97]
Castle 2011a	222	1964	209	5428	0.52 [0.47, 0.56]	0.73 [0.72, 0.74]
Clavel 2001	72	385	10	5184	0.88 [0.79, 0.94]	0.93 [0.92, 0.94]
de Cremoux 2003	32	189	9	1529	0.78 [0.62, 0.89]	0.89 [0.87, 0.90]
Depuydt 2011	27	160	19	2699	0.59 [0.43, 0.73]	0.94 [0.93, 0.95]
Hovland 2010	11	9	4	277	0.73 [0.45, 0.92]	0.97 [0.94, 0.99]
Li 2009	69	370	5	2118	0.93 [0.85, 0.98]	0.85 [0.84, 0.87]
Monsonego 2011	70	355	31	3973	0.69 [0.59, 0.78]	0.92 [0.91, 0.93]
Moy 2010	175	860	37	7816	0.83 [0.77, 0.87]	0.90 [0.89, 0.91]
Pan 2003	81	432	5	1475	0.94 [0.87, 0.98]	0.77 [0.75, 0.79]
Ronco 2006	54	850	19	15593	0.74 [0.62, 0.84]	0.95 [0.94, 0.95]
Wu 2010	18	89	9	1899	0.67 [0.46, 0.83]	0.96 [0.95, 0.96]

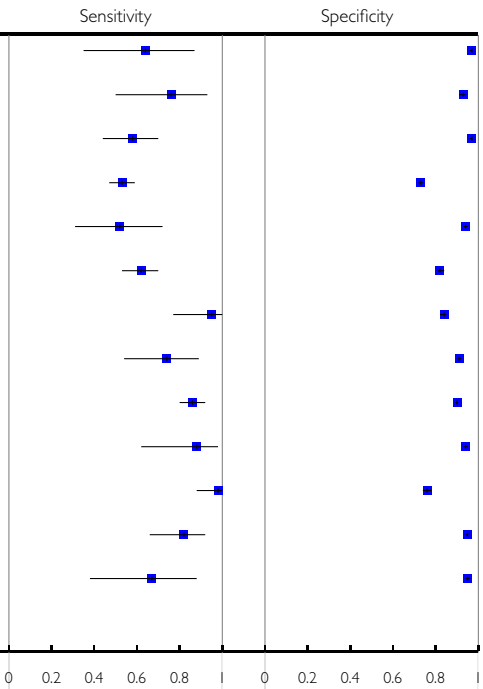


Test 6. LBC (ASCUS+) for CIN3+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 6 LBC (ASCUS+) for CIN3+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Agorastos 2015	9	132	5	3697	0.64 [0.35, 0.87]	0.97 [0.96, 0.97]
Belinson 2010	13	69	4	893	0.76 [0.50, 0.93]	0.93 [0.91, 0.94]
Bigras 2005	34	459	25	13324	0.58 [0.44, 0.70]	0.97 [0.96, 0.97]
Castle 2011a	146	2040	128	5509	0.53 [0.47, 0.59]	0.73 [0.72, 0.74]
Depuydt 2011	13	174	12	2706	0.52 [0.31, 0.72]	0.94 [0.93, 0.95]
Kulasingam 2002	80	694	50	3251	0.62 [0.53, 0.70]	0.82 [0.81, 0.84]
Li 2009	21	418	1	2122	0.95 [0.77, 1.00]	0.84 [0.82, 0.85]
Monsonego 2011	20	405	7	3997	0.74 [0.54, 0.89]	0.91 [0.90, 0.92]
Moy 2010	121	914	19	7834	0.86 [0.80, 0.92]	0.90 [0.89, 0.90]
Nieves 2013	14	119	2	1914	0.88 [0.62, 0.98]	0.94 [0.93, 0.95]
Pan 2003	42	471	1	1479	0.98 [0.88, 1.00]	0.76 [0.74, 0.78]
Ronco 2006	31	873	7	15605	0.82 [0.66, 0.92]	0.95 [0.94, 0.95]
Wu 2010	10	97	5	1903	0.67 [0.38, 0.88]	0.95 [0.94, 0.96]

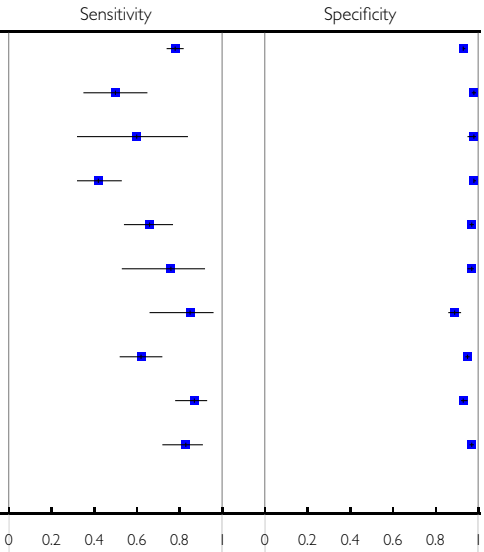


Test 7. LBC (LSIL+) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 7 LBC (LSIL+) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Belinson 2003	294	555	81	7567	0.78 [0.74, 0.82]	0.93 [0.93, 0.94]
Depuydt 2011	23	56	23	2803	0.50 [0.35, 0.65]	0.98 [0.97, 0.99]
Hovland 2010	9	7	6	279	0.60 [0.32, 0.84]	0.98 [0.95, 0.99]
Iftner 2015	38	156	52	9205	0.42 [0.32, 0.53]	0.98 [0.98, 0.99]
Li 2009	49	74	25	2414	0.66 [0.54, 0.77]	0.97 [0.96, 0.98]
McAdam 2010a	16	14	5	446	0.76 [0.53, 0.92]	0.97 [0.95, 0.98]
McAdam 2010b	23	51	4	434	0.85 [0.66, 0.96]	0.89 [0.86, 0.92]
Monsonego 2011	63	215	38	4113	0.62 [0.52, 0.72]	0.95 [0.94, 0.96]
Pan 2003	75	124	11	1783	0.87 [0.78, 0.93]	0.93 [0.92, 0.95]
Qiao 2008	58	69	12	2249	0.83 [0.72, 0.91]	0.97 [0.96, 0.98]

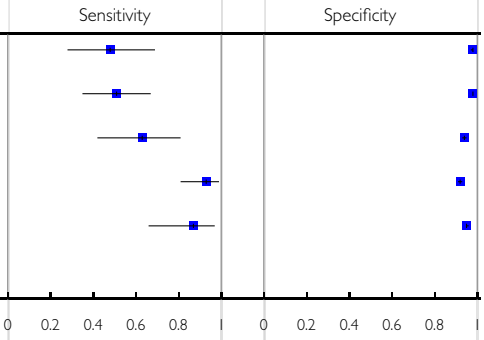


Test 8. LBC (LSIL+) for CIN3+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 8 LBC (LSIL+) for CIN3+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Depuydt 2011	12	67	13	2813	0.48 [0.28, 0.69]	0.98 [0.97, 0.98]
Iftner 2015	22	172	21	9236	0.51 [0.35, 0.67]	0.98 [0.98, 0.98]
Monsonego 2011	17	261	10	4141	0.63 [0.42, 0.81]	0.94 [0.93, 0.95]
Pan 2003	40	159	3	1791	0.93 [0.81, 0.99]	0.92 [0.91, 0.93]
Qiao 2008	20	107	3	2258	0.87 [0.66, 0.97]	0.95 [0.95, 0.96]

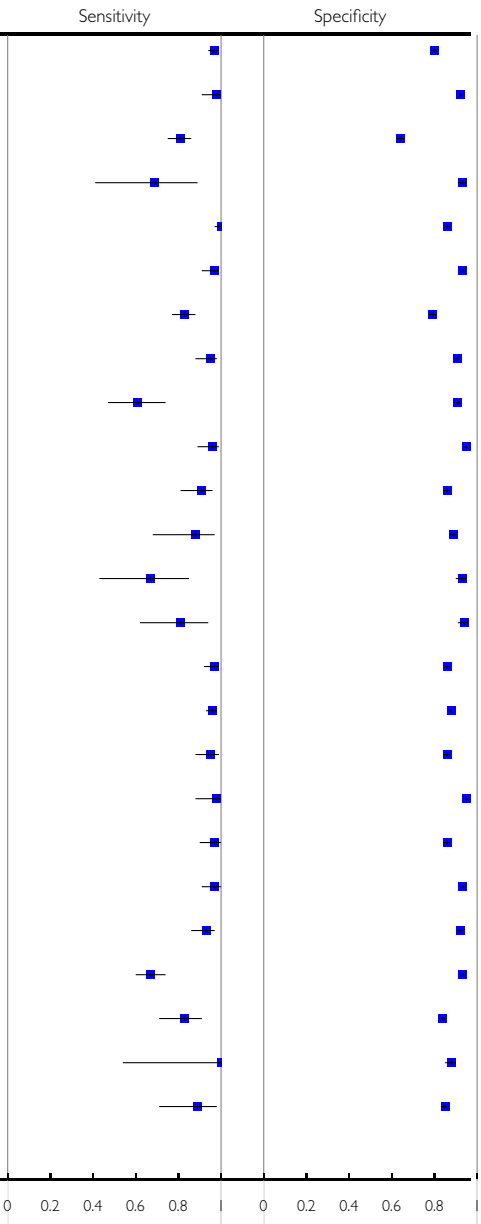


Test 9. HC2 (1pg/mL) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 9 HC2 (1pg/mL) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Belinson 2003	363	1652	12	6470	0.97 [0.94, 0.98]	0.80 [0.79, 0.81]
Bigras 2005	80	1063	2	12697	0.98 [0.91, 1.00]	0.92 [0.92, 0.93]
Blumenthal 2001	168	721	40	1270	0.81 [0.75, 0.86]	0.64 [0.62, 0.66]
Cardenas-Turanzas 2008	11	55	5	764	0.69 [0.41, 0.89]	0.93 [0.91, 0.95]
Clavel 2001	129	1085	0	6718	1.00 [0.97, 1.00]	0.86 [0.85, 0.87]
Cuzick 2003	87	697	3	9571	0.97 [0.91, 0.99]	0.93 [0.93, 0.94]
de Cremoux 2003	150	331	31	1273	0.83 [0.77, 0.88]	0.79 [0.77, 0.81]
Ferreccio 2013	91	742	5	7374	0.95 [0.88, 0.98]	0.91 [0.90, 0.91]
Gravitt 2010	34	200	22	1975	0.61 [0.47, 0.74]	0.91 [0.90, 0.92]
Iftner 2015	86	494	4	8867	0.96 [0.89, 0.99]	0.95 [0.94, 0.95]
Li 2009	67	351	7	2137	0.91 [0.81, 0.96]	0.86 [0.84, 0.87]
Mahmud 2012	21	148	3	1180	0.88 [0.68, 0.97]	0.89 [0.87, 0.90]
McAdam 2010a	14	34	7	436	0.67 [0.43, 0.85]	0.93 [0.90, 0.95]
McAdam 2010b	22	31	5	454	0.81 [0.62, 0.94]	0.94 [0.91, 0.96]
Monsonego 2011	98	595	3	3733	0.97 [0.92, 0.99]	0.86 [0.85, 0.87]
Moy 2010	204	1037	8	7268	0.96 [0.93, 0.98]	0.88 [0.87, 0.88]
Pan 2003	79	248	4	1505	0.95 [0.88, 0.99]	0.86 [0.84, 0.87]
Petry 2003	45	369	1	7493	0.98 [0.88, 1.00]	0.95 [0.95, 0.96]
Qiao 2008	68	333	2	1985	0.97 [0.90, 1.00]	0.86 [0.84, 0.87]
Ronco 2006	73	1112	2	15223	0.97 [0.91, 1.00]	0.93 [0.93, 0.94]
Salmeron 2003	94	626	7	7205	0.93 [0.86, 0.97]	0.92 [0.91, 0.93]
Sankaranarayanan 2004a	120	750	59	10589	0.67 [0.60, 0.74]	0.93 [0.93, 0.94]
Sarian 2005	52	665	11	3467	0.83 [0.71, 0.91]	0.84 [0.83, 0.85]
Shipitsyna 2011	6	101	0	716	1.00 [0.54, 1.00]	0.88 [0.85, 0.90]
Wu 2010	24	306	3	1682	0.89 [0.71, 0.98]	0.85 [0.83, 0.86]

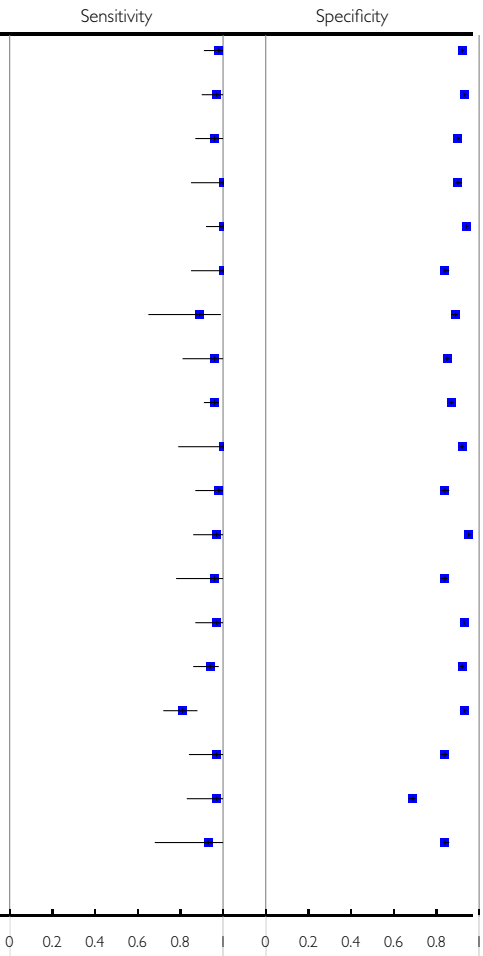


Test 10. HC2 (1 pg/mL) for CIN3+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 10 HC2 (1 pg/mL) for CIN3+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Bigras 2005	58	1085	1	12698	0.98 [0.91, 1.00]	0.92 [0.92, 0.93]
Cuzick 2003	67	717	2	9572	0.97 [0.90, 1.00]	0.93 [0.93, 0.94]
Ferreccio 2013	52	781	2	7377	0.96 [0.87, 1.00]	0.90 [0.90, 0.91]
Gravitt 2010	23	211	0	1997	1.00 [0.85, 1.00]	0.90 [0.89, 0.92]
Iftner 2015	43	537	0	8871	1.00 [0.92, 1.00]	0.94 [0.94, 0.95]
Li 2009	22	396	0	2144	1.00 [0.85, 1.00]	0.84 [0.83, 0.86]
Mahmud 2012	16	153	2	1181	0.89 [0.65, 0.99]	0.89 [0.87, 0.90]
Monsonego 2011	26	667	1	3735	0.96 [0.81, 1.00]	0.85 [0.84, 0.86]
Moy 2010	134	1107	6	7270	0.96 [0.91, 0.98]	0.87 [0.86, 0.88]
Nieves 2013	16	158	0	1875	1.00 [0.79, 1.00]	0.92 [0.91, 0.93]
Pan 2003	40	287	1	1508	0.98 [0.87, 1.00]	0.84 [0.82, 0.86]
Petry 2003	36	378	1	7493	0.97 [0.86, 1.00]	0.95 [0.95, 0.96]
Qiao 2008	22	379	1	1986	0.96 [0.78, 1.00]	0.84 [0.82, 0.85]
Ronco 2006	38	1147	1	15224	0.97 [0.87, 1.00]	0.93 [0.93, 0.93]
Salmeron 2003	87	633	6	7206	0.94 [0.86, 0.98]	0.92 [0.91, 0.93]
Sankaranarayanan 2004a	87	783	21	10627	0.81 [0.72, 0.88]	0.93 [0.93, 0.94]
Sarian 2005	31	686	1	3477	0.97 [0.84, 1.00]	0.84 [0.82, 0.85]
Syrjanen 2002	29	984	1	2161	0.97 [0.83, 1.00]	0.69 [0.67, 0.70]
Wu 2010	14	316	1	1684	0.93 [0.68, 1.00]	0.84 [0.83, 0.86]

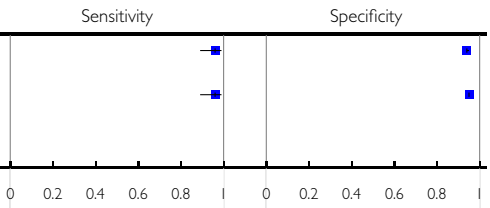


Test 11. HC2 (2 pg/mL) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 11 HC2 (2 pg/mL) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Cuzick 2003	86	589	4	9679	0.96 [0.89, 0.99]	0.94 [0.94, 0.95]
Ronco 2006	72	836	3	15499	0.96 [0.89, 0.99]	0.95 [0.95, 0.95]

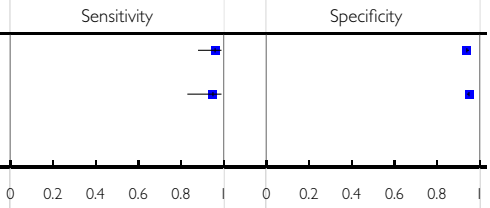


Test 12. HC2 (2 pg/mL) for CIN3+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 12 HC2 (2 pg/mL) for CIN3+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Cuzick 2003	66	609	3	9680	0.96 [0.88, 0.99]	0.94 [0.94, 0.95]
Ronco 2006	37	871	2	15500	0.95 [0.83, 0.99]	0.95 [0.94, 0.95]

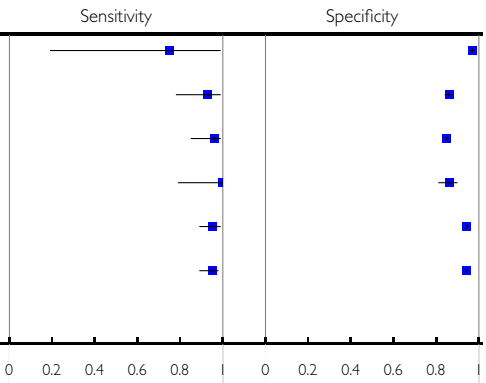


Test 13. PCR (13 hr types or more) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 13 PCR (13 hr types or more) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Agorastos 2005	3	34	1	1258	0.75 [0.19, 0.99]	0.97 [0.96, 0.98]
Belinson 2010	28	133	2	816	0.93 [0.78, 0.99]	0.86 [0.84, 0.88]
Depuydt 2011	44	429	2	2430	0.96 [0.85, 0.99]	0.85 [0.84, 0.86]
Hovland 2010	16	42	0	255	1.00 [0.79, 1.00]	0.86 [0.81, 0.90]
Naucler 2009	83	350	4	5652	0.95 [0.89, 0.99]	0.94 [0.94, 0.95]
Schneider 2000	108	263	6	4384	0.95 [0.89, 0.98]	0.94 [0.94, 0.95]

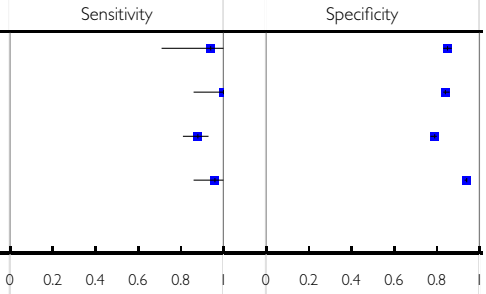


Test 14. PCR (13 hr types or more) for CIN3+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 14 PCR (13 hr types or more) for CIN3+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Belinson 2010	16	145	1	817	0.94 [0.71, 1.00]	0.85 [0.83, 0.87]
Depuydt 2011	25	448	0	2432	1.00 [0.86, 1.00]	0.84 [0.83, 0.86]
Kulasingam 2002	118	836	16	3105	0.88 [0.81, 0.93]	0.79 [0.77, 0.80]
Naucler 2009	48	385	2	5654	0.96 [0.86, 1.00]	0.94 [0.93, 0.94]

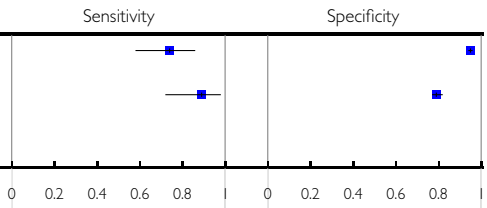


Test 15. PCR (10-11 hr types) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 15 PCR (10-11 hr types) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Cuzick 1999	31	146	11	2800	0.74 [0.58, 0.86]	0.95 [0.94, 0.96]
Paraskevaidis 2001	25	195	3	754	0.89 [0.72, 0.98]	0.79 [0.77, 0.82]

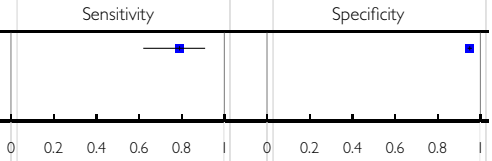


Test 16. PCR (10-11 hr types) for CIN3+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 16 PCR (10-11 hr types) for CIN3+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Cuzick 1999	27	150	7	2804	0.79 [0.62, 0.91]	0.95 [0.94, 0.96]

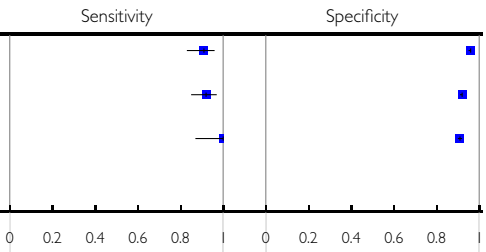


Test 17. Aptima for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 17 Aptima for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Iftner 2015	82	382	8	8979	0.91 [0.83, 0.96]	0.96 [0.95, 0.96]
Monsonego 2011	93	363	8	3965	0.92 [0.85, 0.97]	0.92 [0.91, 0.92]
Wu 2010	27	174	0	1814	1.00 [0.87, 1.00]	0.91 [0.90, 0.92]

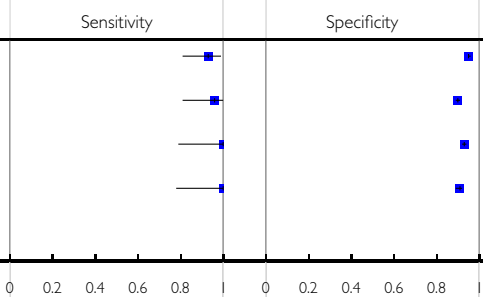


Test 18. Aptima for CIN3+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 18 Aptima for CIN3+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Iftner 2015	40	424	3	8984	0.93 [0.81, 0.99]	0.95 [0.95, 0.96]
Monsonego 2011	26	430	1	3972	0.96 [0.81, 1.00]	0.90 [0.89, 0.91]
Nieves 2013	16	133	0	1900	1.00 [0.79, 1.00]	0.93 [0.92, 0.94]
Wu 2010	15	186	0	1814	1.00 [0.78, 1.00]	0.91 [0.89, 0.92]

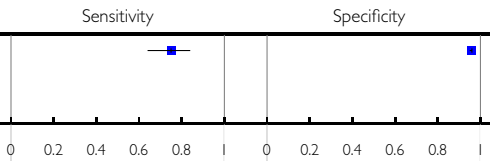


Test 19. PCR (4 hr types) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 19 PCR (4 hr types) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Cuzick 1995	61	85	20	1819	0.75 [0.64, 0.84]	0.96 [0.95, 0.96]

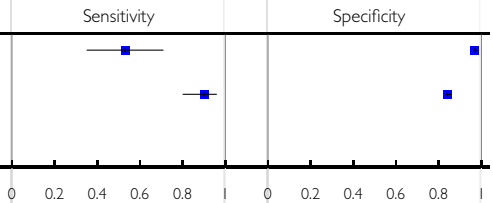


Test 20. Care HPV test (0.5 pg/ml) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 20 Care HPV test (0.5 pg/ml) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Labani 2014	17	120	15	4504	0.53 [0.35, 0.71]	0.97 [0.97, 0.98]
Qiao 2008	63	367	7	1951	0.90 [0.80, 0.96]	0.84 [0.83, 0.86]

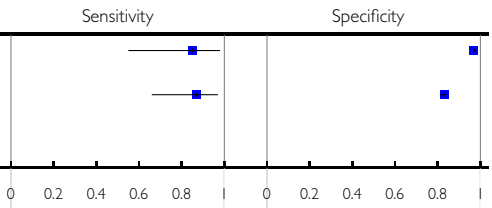


Test 21. Care HPV test (0.5 pg/ml) for CIN3+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 21 Care HPV test (0.5 pg/ml) for CIN3+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Labani 2014	11	126	2	4519	0.85 [0.55, 0.98]	0.97 [0.97, 0.98]
Qiao 2008	20	410	3	1955	0.87 [0.66, 0.97]	0.83 [0.81, 0.84]

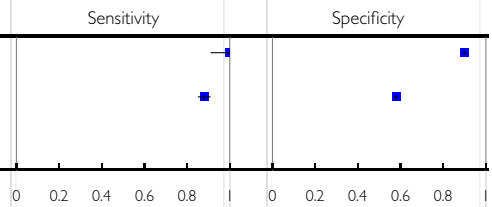


Test 22. Cobas for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 22 Cobas for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Agorastos 2015	41	368	0	3434	1.00 [0.91, 1.00]	0.90 [0.89, 0.91]
Castle 2011a	380	3122	51	4270	0.88 [0.85, 0.91]	0.58 [0.57, 0.59]

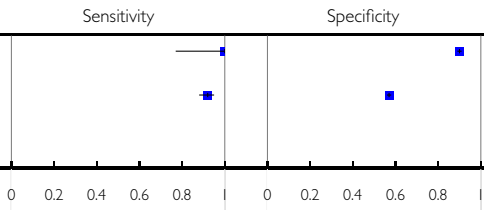


Test 23. Cobas for CIN3+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 23 Cobas for CIN3+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Agorastos 2015	14	395	0	3434	1.00 [0.77, 1.00]	0.90 [0.89, 0.91]
Castle 2011a	252	3250	22	4299	0.92 [0.88, 0.95]	0.57 [0.56, 0.58]

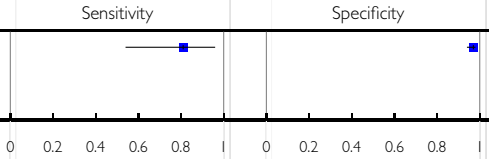


Test 24. NASBA (5 types) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 24 NASBA (5 types) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Hovland 2010	13	10	3	287	0.81 [0.54, 0.96]	0.97 [0.94, 0.98]

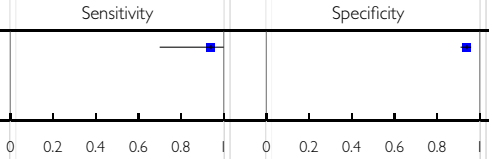


Test 25. NASBA (9 types) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 25 NASBA (9 types) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Hovland 2010	15	18	1	279	0.94 [0.70, 1.00]	0.94 [0.91, 0.96]

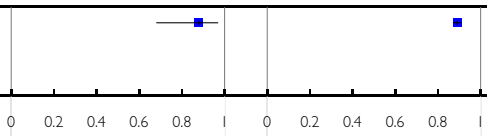


Test 26. HC2+4 (1 pg/ml) for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 26 HC2+4 (1 pg/ml) for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Mahmud 2012	21	147	3	1181	0.88 [0.68, 0.97]	0.89 [0.87, 0.91]		

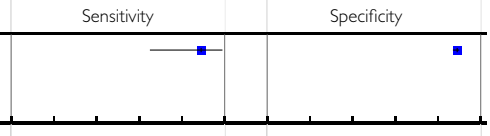


Test 27. HC2+4 (1 pg/ml) for CIN3+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 27 HC2+4 (1 pg/ml) for CIN3+

Study	TP	FP	FN	TN	Sensitivity	Specificity	Sensitivity	Specificity
Mahmud 2012	16	152	2	1182	0.89 [0.65, 0.99]	0.89 [0.87, 0.90]		

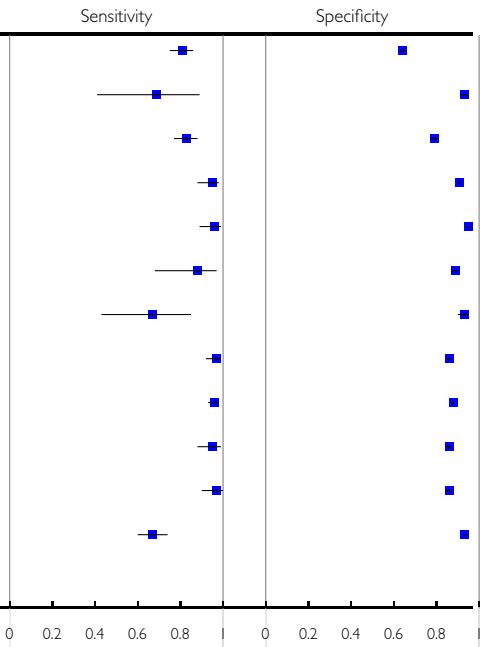


Test 28. HC2 (1pg/mL) for CIN2+ no verification bias.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 28 HC2 (1pg/mL) for CIN2+ no verification bias

Study	TP	FP	FN	TN	Sensitivity	Specificity
Blumenthal 2001	168	721	40	1270	0.81 [0.75, 0.86]	0.64 [0.62, 0.66]
Cardenas-Turanzas 2008	11	55	5	764	0.69 [0.41, 0.89]	0.93 [0.91, 0.95]
de Cremoux 2003	150	331	31	1273	0.83 [0.77, 0.88]	0.79 [0.77, 0.81]
Ferreccio 2013	91	742	5	7374	0.95 [0.88, 0.98]	0.91 [0.90, 0.91]
Iftner 2015	86	494	4	8867	0.96 [0.89, 0.99]	0.95 [0.94, 0.95]
Mahmud 2012	21	148	3	1180	0.88 [0.68, 0.97]	0.89 [0.87, 0.90]
McAdam 2010a	14	34	7	436	0.67 [0.43, 0.85]	0.93 [0.90, 0.95]
Monsonego 2011	98	595	3	3733	0.97 [0.92, 0.99]	0.86 [0.85, 0.87]
Moy 2010	204	1037	8	7268	0.96 [0.93, 0.98]	0.88 [0.87, 0.88]
Pan 2003	79	248	4	1505	0.95 [0.88, 0.99]	0.86 [0.84, 0.87]
Qiao 2008	68	333	2	1985	0.97 [0.90, 1.00]	0.86 [0.84, 0.87]
Sankaranarayanan 2004a	120	750	59	10589	0.67 [0.60, 0.74]	0.93 [0.93, 0.94]

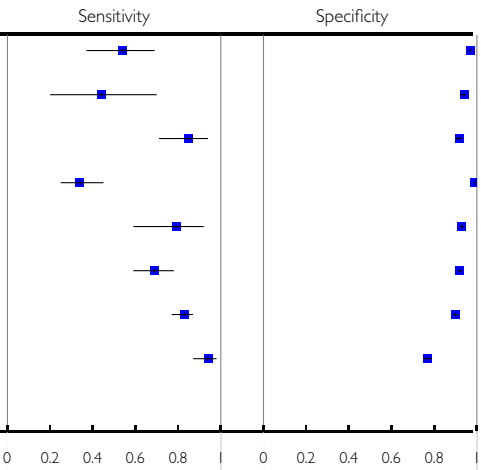


Test 29. CC or LBC (ASCUS+) for CIN2+ no verification bias.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 29 CC or LBC (ASCUS+) for CIN2+ no verification bias

Study	TP	FP	FN	TN	Sensitivity	Specificity
Agorastos 2015	22	119	19	3683	0.54 [0.37, 0.69]	0.97 [0.96, 0.97]
Cardenas-Turanzas 2008	7	52	9	767	0.44 [0.20, 0.70]	0.94 [0.92, 0.95]
de Cremoux 2003	35	141	6	1573	0.85 [0.71, 0.94]	0.92 [0.90, 0.93]
Ferreccio 2013	33	99	63	8017	0.34 [0.25, 0.45]	0.99 [0.99, 0.99]
Mahmud 2012	22	94	6	1264	0.79 [0.59, 0.92]	0.93 [0.92, 0.94]
Monsonego 2011	70	355	31	3973	0.69 [0.59, 0.78]	0.92 [0.91, 0.93]
Moy 2010	175	860	37	7816	0.83 [0.77, 0.87]	0.90 [0.89, 0.91]
Pan 2003	81	432	5	1475	0.94 [0.87, 0.98]	0.77 [0.75, 0.79]

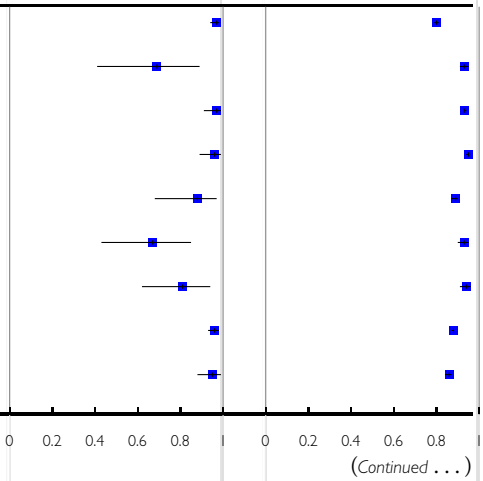


Test 30. HC2 (1pg/mL) for CIN2+ women >30.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 30 HC2 (1pg/mL) for CIN2+ women >30

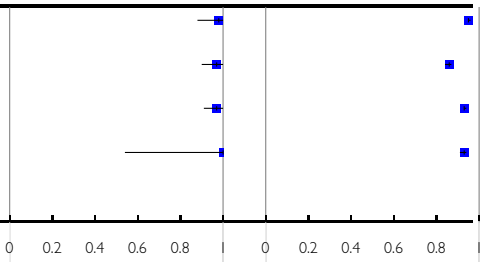
Study	TP	FP	FN	TN	Sensitivity	Specificity
Belinson 2003	363	1652	12	6470	0.97 [0.94, 0.98]	0.80 [0.79, 0.81]
Cardenas-Turanzas 2008	11	55	5	764	0.69 [0.41, 0.89]	0.93 [0.91, 0.95]
Cuzick 2003	87	697	3	9571	0.97 [0.91, 0.99]	0.93 [0.93, 0.94]
Iftner 2015	86	494	4	8867	0.96 [0.89, 0.99]	0.95 [0.94, 0.95]
Mahmud 2012	21	148	3	1180	0.88 [0.68, 0.97]	0.89 [0.87, 0.90]
McAdam 2010a	14	34	7	436	0.67 [0.43, 0.85]	0.93 [0.90, 0.95]
McAdam 2010b	22	31	5	454	0.81 [0.62, 0.94]	0.94 [0.91, 0.96]
Moy 2010	204	1037	8	7268	0.96 [0.93, 0.98]	0.88 [0.87, 0.88]
Pan 2003	79	248	4	1505	0.95 [0.88, 0.99]	0.86 [0.84, 0.87]



(*Continued . . .*)

(. . . *Continued*)

Study	TP	FP	FN	TN	Sensitivity	Specificity
Petry 2003	45	369	1	7493	0.98 [0.88, 1.00]	0.95 [0.95, 0.96]
Qiao 2008	68	333	2	1985	0.97 [0.90, 1.00]	0.86 [0.84, 0.87]
Ronco 2006	73	1112	2	15223	0.97 [0.91, 1.00]	0.93 [0.93, 0.94]
Shipitsyna 2011	6	57	0	716	1.00 [0.54, 1.00]	0.93 [0.91, 0.94]

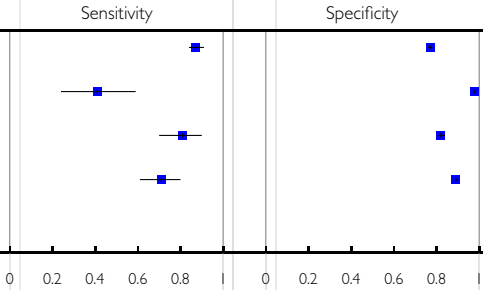


Test 31. self HPV test for CIN2+.

Review: Cytology versus HPV testing for cervical cancer screening in the general population

Test: 31 self HPV test for CIN2+

Study	TP	FP	FN	TN	Sensitivity	Specificity
Belinson 2003	328	1850	47	6272	0.87 [0.84, 0.91]	0.77 [0.76, 0.78]
Labani 2014	13	98	19	4527	0.41 [0.24, 0.59]	0.98 [0.97, 0.98]
Qiao 2008	57	409	13	1909	0.81 [0.70, 0.90]	0.82 [0.81, 0.84]
Salmeron 2003	72	823	29	7008	0.71 [0.61, 0.80]	0.89 [0.89, 0.90]



ADDITIONAL TABLES

Table 1. Pooled diagnostic accuracy of tests

Test	Disease threshold	studies	Pooled sensitivity (95% CI)	Pooled specificity (95% CI)
CC (ASCUS+)	CIN 2+	16	65.87% (54.94 to 75.33)	96.28% (94.72 to 97.39)
LBC (ASCUS+)	CIN 2+	15	75.51% (66.57 to 82.68)	91.85% (88.43 to 94.32)
CC (LSIL+)	CIN 2+	9	62.84% (46.79-76.50)	97.73% (96.09-98.70)

Table 1. Pooled diagnostic accuracy of tests (*Continued*)

LBC (LSIL+)	CIN 2+	10	70.33% (59.73 to 79.11)	96.20% (94.57 to 97.36)
HC2 (1 pg/mL)	CIN 2+	25	92.60% (99.45 to 95.30)	89.30% (87.03 to 91.20)
PCR (> 12 types)	CIN 2+	6	95.13% (89.50 to 97.84)	91.89% (83.79 to 96.13)
APTIMA	CIN 2+	3	92.66% (31.77 to 99.71)	93.31% (47.30 to 99.54)
CC (ASCUS+)	CIN 3+	9	70.27% (57.87 to 80.30)	96.67% (94.56 to 98.00)
LBC (ASCUS+)	CIN 3+	13	75.97% (64.72 to 84.49)	91.19% (87.21 to 94.01)
CC (LSIL+)	CIN 3+	5	74.43% (67.81 to 80.10)	96.86% (94.87 to 98.10)
LBC (LSIL+)	CIN 3+	5	71.91% (51.68 to 86.00)	96.05% (93.53 to 97.60)
HC2 (1 pg/mL)	CIN 3+	19	96.50% (94.00 to 97.90)	89.20% (86.70 to 91.30)
PCR (> 12 types)	CIN 3+	4	93.57% (69.90 to 98.91)	86.49% (68.16 to 95.04)
APTIMA	CIN 3+	4	96.04% (72.91 to 99.54)	92.80% (86.15 to 96.39)

Tests with fewer than three studies are not included in the table.

Table 2. Test comparisons

Comparison	Disease threshold	Relative sensitivity (95% CI)	Relative specificity (95% CI)	Studies	Analysis number
HC2 vs CC (ASCUS+)	CIN 2+	1.52 (1.24 to 1.86)	0.94 (0.92 to 0.96)	9	1
HC2 vs CC (ASCUS+)	CIN 3+	1.46 (1.12 to 1.91)	0.95 (0.93 to 0.9)	6	2
PCR (> 12 types) vs CC (ASCUS+)	CIN 2+	1.37 (0.58 to 3.21)	0.95 (0.76 to 1.19)	3	5
HC2 vs CC (LSIL+)	CIN 2+	1.28 (1.15 to 1.41)	0.91 (0.87 to 0.95)	6	7
HC2 vs CC (LSIL+)	CIN 3+	1.22 (1.12 to 1.32)	0.91 (0.87 to 0.95)	5	8
HC2 vs LBC (ASCUS+)	CIN 2+	1.18 (1.10 to 1.26)	0.96 (0.95 to 0.97)	10	11
HC2 vs LBC (ASCUS+)	CIN 3+	1.17 (1.05 to 1.30)	0.96 (0.95 to 0.98)	8	12

Table 2. Test comparisons (*Continued*)

PCR (> 12 types) vs LBC (ASCUS+)	CIN 2+	1.53 (0.53 to 4.44)	0.90 (0.89 to 0.92)	3	15
PCR (> 12 types) vs LBC (ASCUS+)	CIN 3+	1.47 (0.64 to 3.35)	0.94 (0.8 to 1.09)	3	16
HC2 vs LBC (LSIL+)	CIN 2+	1.35 (1.19 to 1.53)	0.92 (0.89 to 0.95)	8	17
HC2 vs LBC (LSIL+)	CIN 3+	1.30 (0.86 to 1.96)	0.92 (0.8 to 1.00)	4	18
APTIMA vs LBC (ASCUS+)	CIN 3+	1.30 (0.49 to 3.41)	0.98 (0.93 to 1.04)	3	22

Comparisons with fewer than three studies are not included in the table

Table 3. Variation in the accuracy of HC2 by covariates

Comparison	Studies	Disease threshold	Relative sensitivity (95% CI)	Relative specificity (95% CI)
Age > 30 vs any age	17 vs 20	CIN 2+	1.13 (1.03 to 1.25)	1.01 (0.98 to 1.04)
	13 vs 14	CIN 3+	1.10 (1.02 to 1.19)	1.04 (1.00 to 1.08)
Increased vs low risk of verification bias	17 vs 20	CIN 2+	1.05 (0.95 to 1.16)	1.00 (0.97 to 1.04)
	12 vs 15	CIN 3+	1.09 (1.01 to 1.18)	1.00 (0.96 to 1.05)
High-income vs middle-/low-income countries	21 vs 16	CIN 2+	1.01 (0.91 to 1.12)	1.03 (1.00 to 1.07)
	13 vs 14	CIN 3+	0.94 (0.87 to 1.02)	1.01 (0.96 to 1.05)

Assessed by bivariate random-effects meta-analysis including one covariate each time.

APPENDICES

Appendix 1. Search strategy

MEDLINE (Pubmed):

((Uterine Cervical Neoplasms [MeSH Terms] OR Uterine Cervical Dysplasia [MeSH Terms] OR Cervical Intraepithelial Neoplasia [MeSH Terms] OR ((cervix [tw] OR cervical [tw] OR cervico* [tw]) AND (cancer* [tw] OR carcinoma OR adenocarcinoma OR neoplas* [tw] OR dysplas* [tw] OR dyskaryos* [tw] OR squamous [tw] OR CIN [tw] OR CINII* [tw] OR CIN2* [tw] OR CINIII* [tw] OR CIN3* [tw] OR SIL [tw] OR HSIL [tw] OR H-SIL [tw] OR LSIL [tw] OR L-SIL [tw] OR ASCUS [tw] OR AS-CUS [tw])))
AND
(papillomaviridae [MeSH:NoExp] OR alphapapillomavirus [MeSH Terms] OR "DNA, viral" [MeSH Terms] OR Papillomavirus Infections [MeSH Terms] OR Tumor Virus Infections [MeSH Terms] OR "Cervix Uteri/virology" [MeSH Terms] OR HPV [tw] OR "human papillomavirus" [tw] OR papillomaviridae [tw] OR PCR OR "hybrid capture*" [tw] OR HC2 [tw] OR HCII [tw] OR "HC 2" [tw] OR "HC II" [tw] OR ((viral [tw] OR virolog* [tw]) AND (DNA [tw])))
AND
(Vaginal smears [MeSH Terms] OR Cytodiagnosis [MeSH Terms] OR Cell Transformation, Viral [MeSH Terms] OR Cytopathogenic Effect, Viral [MeSH Terms] OR ((pap [tw] OR papanicolaou [tw] OR vagina* [tw] OR cervical [tw] OR cervix [tw] OR cervico* [tw] OR cytolog* [tw]) AND (smear* OR test [tw] OR tests [tw] OR testing [tw] OR tested [tw] OR swab* OR scrap*))))

Embase (Ovid):

1. exp Uterine Cervical Neoplasms/
2. exp Uterine Cervical Dysplasia/
3. exp Cervical Intraepithelial Neoplasia/
4. (cervi$ adj3 (cancer or carcinoma or adenocarcinoma or neoplasm$ or dysplas$ or dyskaryo$ or or squamous or CIN$ or HSIL or LSIL or ASCUS) mp.
5. or/1-4
6. Papillomaviridae/
7. exp. alphapapillomavirus
8. exp."DNA, viral"
9. exp. Papillomavirus Infections
10. exp. Tumor Virus Infections
11. exp."Cervix Uteri/virology"
12. HPV mp.
13. "human papillomavirus" mp.
14. papillomaviridae mp.
15. PCR mp.
16. "hybrid capture$" mp.
17. HC2 mp.
18. HCII mp.
19. "HC 2" mp.
20. "HC II" mp.
21. ((viral or virology$) adj3 DNA) mp.
22. or/6-21
23. exp Vaginal smears/
24. exp. Cytodiagnosis/
25. exp. Cell Transformation, Viral/
26. exp. Cytopathogenic Effect, Viral/
27. ((pap or papanicolaou or vagina$ or cervical or cervix or cervico$ or cytology$) adj3 (smear$ or test or tests or testing or tested or swab$ or scrap$)) mp.
28. or/ 23-27
29. 5 and 22 and 28

Appendix 2. Quality Assessment of Diagnostic test Accuracy Studies (QUADAS) items

- Item 1 of the QUADAS tool (representative spectrum of participants) will be scored as 'yes' if the tests are done on a representative population of women attending cervical cancer screening within the age range 20-70 not being followed up for cervical abnormalities. The item will be scored as 'no' if the majority of the population is outside this range and 'unclear' if there is not sufficient information. Studies with a case-control design would score 'no' but they are excluded anyway from this review.
- Item 2 of the QUADAS tool (selection criteria clearly described) will not be used as suggested in Chapter 9 of the *Cochrane Handbook for Diagnostic Test Accuracy Reviews*Reitsma 2009.
- Item 3 of the QUADAS tool (acceptable reference standard) will be scored as 'yes' if the reference standard used is colposcopy with directed biopsies as minimum, or better yet by histological examination of the whole excised transformation zone. The item will be scored 'no' if the reference standard used is colposcopy alone without histology, and 'unclear' if there is not sufficient information.
- Item 4 of the QUADAS tool (acceptable delay between tests) will be scored as 'yes' if the total interval between cytology, human papillomavirus (HPV) testing and verification with the reference standard was less than 12 weeks, as the status of the condition is unlikely to change within this time period. The item will be scored 'no' if the interval was equal to or more than 12 weeks, and 'unclear' if there is not sufficient information.
- Item 5 of the QUADAS tool (partial verification avoided) will be scored as 'yes' if all women or at least a random sample of all women tested with cytology or HPV testing had disease status verification by the reference standard (colposcopy with directed biopsies) or when all women being positive for at least one screen test were verified together with a random sample of women being negative for all screen tests . The item will be scored 'no' if the selection of women who will receive verification is influenced by the results of the screening tests (i.e. if all screen-positives are verified and not all screen-negatives). The item will be scored as 'unclear' if there is not sufficient information.
- Item 6 of the QUADAS tool (differential verification avoided) will be scored as 'yes' if all women who had disease status verification, had this done by the same method. The item will be scored 'no' if the method of verification differed between groups of participants, and 'unclear' if there is not sufficient information.
- Item 7 of the QUADAS tool (incorporation avoided) will be scored as 'yes' if the reference standard used for disease status verification is not composed in any part by cervical cytology or HPV testing. A reference standard such as colposcopy would score 'yes'. If cytology is used as a reference standard the item will be scored as 'no', and 'unclear' if there is not sufficient information
- Item 8 of the QUADAS tool (sufficient index test description) will not be used as suggested in Chapter 9 of the *Cochrane Handbook for Diagnostic Test Accuracy Reviews*Reitsma 2009.
- Item 9 of the QUADAS tool (sufficient reference standard description) will not be used as suggested in Chapter 9 of the *Cochrane Handbook for Diagnostic Test Accuracy Reviews*Reitsma 2009.
- Item 10 of the QUADAS tool (index test results blinded) will be scored as 'yes' if the cytologists and the technicians interpreting the Pap smear and the HPV test were not aware of the colposcopy/biopsy results. The item will be scored 'no' if they were made aware of the reference standard results prior to the interpretation of the screening tests, and 'unclear' if there is not sufficient information given in the text.
- Item 11 of the QUADAS tool (reference standard results blinded) will be scored as 'yes' if the colposcopists and the pathologists were not aware of the cytology and HPV test results when interpreting the results of the reference standard. The item will be scored 'no' if either was aware of the screening test results (which is the case in clinical practice), and 'unclear' if there is not sufficient information in the text.
- Item 12 of the QUADAS tool (relevant clinical information) will be scored as 'yes' if the cytologist was aware of the woman's basic history (age, symptoms, previous cervical surgery). If the cytologist was not aware the item will be scored 'no', and if this information is not given in the text it will be scored 'unclear'.
- Item 13 of the QUADAS tool (un-interpretable results reported) will be scored as 'yes' if the numbers of inadequate cytology and HPV test results are given. It will be scored 'no' if the numbers of inadequate tests are not given, and 'unclear' if it is not certain whether all test results have been reported.
- Item 14 of the QUADAS tool (withdrawals explained) will be scored as 'yes' if it is clear what happened to all participants who entered the study, including the withdrawals. The item will be scored 'no' if it is not explained why no outcome could be obtained for some women, and if it is not clear whether all participants who entered the study were accounted for it will be scored 'unclear'.

CONTRIBUTIONS OF AUTHORS

G Koliopoulos: article search, data extraction, assessment of methodological quality, data analysis, drafting the review

V Nyawira Nyaga: data analysis

A Bryant: data analysis, drafting the review

N Santesso: article search, commented critically, drafted plain language summary and SoF table

P Martin-Hirsch: assessment of methodological quality, drafting the review

RA Mustafa: data analysis

H Schünemann: commented critically

E Paraskevaidis: conception and design, commented critically

M Arbyn: data extraction, data analysis, drafting the review

DECLARATIONS OF INTEREST

G Koliopoulos: no conflict of interest and no financial support for the development of this review

V Nyawira Nyaga: no conflict of interest and no financial support for the development of this review

A Bryant: no conflict of interest and no financial support for the development of this review

N Santesso: no conflict of interest and no financial support for the development of this review

P Martin-Hirsch: no conflict of interest and no financial support for the development of this review

RA Mustafa: no conflict of interest and no financial support for the development of this review

H Schünemann: no conflict of interest and no financial support for the development of this review

E Paraskevaidis: no conflict of interest and no financial support for the development of this review

M Arbyn: is supported by 1) European Commission (DG Sanco, Luxembourg) through the ECCG project (European Cooperation on Development and Implementation of Cancer Screening and Prevention Guidelines, IARC, Lyon, France) and the CoheaHr Project [603019], co-ordinated by the Free University Medical Center of Amsterdam; 2) Belgian Foundation Against Cancer, Brussels, Belgium; 3) Belgian Cancer Centre, Brussels, Belgium; and 4) IWT (Institute for the Promotion of Innovation by Science and Technology in Flanders, project number 060081), Brussels, Belgium.

SOURCES OF SUPPORT

Internal sources

- None, Other.

External sources

- Cochrane Gynaecological, Neuro-oncology and Orphan Cancers, Bath, UK.
- Department of Health, UK.
- NHS Cochrane Collaboration programme Grant Scheme CPG-506

European Commission (DG Sanco, Luxembourg) through the ECCG project, Luxembourg.
Belgian Foundation Against Cancer, Brussels, Belgium.
IWT (Institute for the Promotion of Innovation by Science and Technology in Flanders), Brussels, Belgium.
project number 060081
FP7 Programme of DG Research of the European Commisssion (though the COHEAHR Network, Grant nr 603019), Belgium.
Financial support of M Arbyn

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

The final review included only studies written in the English language.

INDEX TERMS

Medical Subject Headings (MeSH)

Cervical Intraepithelial Neoplasia [*diagnosis; pathology; virology]; Early Detection of Cancer [methods]; Papillomavirus Infections [*diagnosis]; Polymerase Chain Reaction; Precancerous Conditions [* diagnosis; pathology; virology]; Sensitivity and Specificity; Uterine Cervical Neoplasms [*diagnosis; pathology; virology]; Vaginal Smears [methods]

MeSH check words

Adult; Aged; Female; Humans; Middle Aged