



Published in final edited form as:

Am J Obstet Gynecol. 2013 May ; 208(5): 343–353. doi:10.1016/j.ajog.2012.11.013.

A systematic review of randomized trials assessing human papillomavirus testing in cervical cancer screening

Insiyyah Y. Patanwala, MD, Heidi M. Bauer, MD, MS, MPH, Justin Miyamoto, MD, Ina U. Park, MD, MS, Megan J. Huchko, MD, MPH, and Karen K. Smith-McCune, MD, PhD

University of California, San Francisco, School of Medicine (Drs Patanwala and Miyamoto) and University of California, San Francisco Departments of Obstetrics, Gynecology, and Reproductive Sciences (Drs Huchko and Smith-McCune), Family and Community Medicine (Dr Park), and Epidemiology and Biostatistics (Dr Bauer), San Francisco, CA; and Sexually Transmitted Disease Control Branch, Division of Communicable Disease Control, Center for Infectious Diseases, California Department of Public Health, Richmond, CA (Drs Bauer and Park).

Abstract

Our objective was to assess the sensitivity and specificity of human papillomavirus (HPV) testing for cervical cancer screening in randomized trials. We conducted a systematic literature search of the following databases: MEDLINE, CINAHL, EMBASE, and Cochrane. Eligible studies were randomized trials comparing HPV-based to cytology-based screening strategies, with disease status determined by colposcopy/biopsy for participants with positive results. Disease rates (cervical intraepithelial neoplasia [CIN]2 or greater and CIN3 or greater), sensitivity, and positive predictive value were abstracted or calculated from the articles. Six studies met inclusion criteria. Relative sensitivities for detecting CIN3 or greater of HPV testing-based strategies vs cytology ranged from 0.8 to 2.1. The main limitation of our study was that testing methodologies and screening/management protocols were highly variable across studies. Screening strategies in which a single initial HPV-positive test led to colposcopy were more sensitive than cytology but resulted in higher colposcopy rates. These results have implications for cotesting with HPV and cytology as recommended in the United States.

Keywords

cervical cytology; cervical intraepithelial neoplasia; colposcopy; Papanicolaou test

Cytology, the primary modality for cervical cancer screening in the United States, has resulted in significant declines in cervical cancer morbidity and mortality.¹ Nonetheless, cervical cytology has limitations, including a false-negative rate for cancer of at least 20%,² leading to a search for more sensitive screening strategies. Human papillomavirus (HPV) is the causal factor of cervical cancer, and 18 HPV types including HPV16 and HPV18 have been associated with invasive cancer.³ HPV DNA testing has therefore been proposed as an alternative or adjunct for cervical cancer screening, with advantages that the test is more objective and sensitive than cytology.

Most women acquire HPV soon after sexual debut and spontaneously clear the virus within 1-2 years after infection; only approximately 10% of women remain HPV positive 5 years after acquisition.⁴ Although incident HPV infection is common, the risk for cervical cancer is associated with persistent infection. Therefore, the prognostic value of a single positive HPV test in young women is limited.

Recent guidelines in the United States recommend that in women older than 30 years, cervical cancer screening can be performed with either cytology every 3 years or cytology plus HPV cotesting every 5 years.^{5,6} In cases in which the HPV test is positive and cytology normal, repeating both tests in 12 months is recommended, unless HPV16 or HPV18 is present.⁷ In addition to cotesting as used in the United States, other strategies under evaluation globally include HPV testing alone without cytology and HPV testing first followed by cytology triage for positive results.

Metaanalyses comprised largely of observational studies have shown that HPV testing-based strategies are more sensitive but less specific than cytology-based screening strategies.^{8,9} Although observational studies are useful in determining test accuracy, performance estimates from observational studies can be biased when evaluating 2 diagnostic tests simultaneously in a manner that differs from how the tests are used separately. Potential bias is minimized when sensitivity and specificity of different strategies (single or combined tests) are analyzed in randomized trials.¹⁰ More reliable comparisons of sensitivity and specificity can be made in the context of randomized clinical trials conducted over 1 or more screening episodes.¹¹

Our objective was to assess the sensitivity and specificity of HPV testing in randomized trials by summarizing data from randomized trials of various cervical cancer screening strategies that incorporate HPV testing.

Materials and methods

We conducted a systematic literature search in 2010 of the following electronic databases: MEDLINE, CINAHL, EMBASE, and the Cochrane Library. We used the following subsets of search terms combined by the word “and:” (papillomavirus/papilloma-viridae/papilloma virus/hpv) and (cytodiagnosis/cytolo*/pap smear/papanicolaou/colposcopy/cervical smear/uterine smear/cervicovagina smear/cervix uteri smear/endocervix smear) and (cervical cancer/cervical intraepithelial neoplasia/CIN/cervical dysplasia/cervical neoplasm/uterine cervical neoplasm/uterine cervix tumor/cervix cancer/uterine cervix carcinoma-in situ/uterine cervix dysplasia/cervix dysplasia/cervix neoplasm/uterine cervix cancer/mass screening/cancer screening). The following limit was placed on all searches to retrieve primarily randomized studies: “random*.” No language restrictions were included. The references cited in the articles selected for study inclusion were hand searched for additional citations. For each primary study, citations of relevant articles were searched for in MEDLINE, and a cited reference search was conducted in the Institute for Scientific Information Web of Knowledge electronic database. As a secondary analysis of published data, the study was exempt from institutional review board approval.

Study selection

We applied the following inclusion criteria: (1) the study was a randomized trial comparing HPV-based strategies to cytology-based strategies for primary cervical cancer screening and (2) disease status was determined by colposcopy/biopsy for study participants in whom treatment was warranted. Our study was limited to screening strategies and therefore did not assess the use of reflex HPV testing for a cytology result of atypical squamous cells of

undetermined significance (ASCUS). None of the studies in this review used genotyping assays recently approved by the US Food and Drug Administration.

Studies were selected with a 2-step method. First, resulting titles and abstracts from literature searches were analyzed, and citations that were likely to meet the aforementioned criteria were chosen. The full manuscripts of these citations were then evaluated to determine whether full inclusion criteria were met.

Each study was abstracted onto pre-tested data abstraction forms by at least 2 reviewers with a third reviewer for adjudication of discrepancies. In cases of multiple publications from a single study, the publication analyzing the most recent data set was used. If this was unclear, the study authors were contacted for clarification; repeat data abstraction was performed if indicated by the authors' response. In addition, although another study (the HPV FOCAL trial: a randomized trial of Human papillomavirus testing for cervical cancer screening)¹² met our inclusion criteria, colposcopy results were not published; therefore, it was excluded from the systematic review. The final results from 1 trial were published after the performance of the systematic literature review; results from this publication¹³ were abstracted and used in the analysis primarily rather than the interim results.¹⁴

Study classification

We devised a simplified classification system of different HPV testing strategies to determine whether differences in outcomes were related to testing strategy. These strategies, illustrated in Figure 1, included the following: (1) HPV testing alone (no cytology) with referral colposcopy for positive HPV test results, (2) HPV testing with cytology triage for positive HPV test results, (3) combination of cytology and HPV testing (cotesting) with an active response to positive HPV testing results (women with a positive initial HPV test were referred to colposcopy), and (4) cotesting with a passive response to positive HPV testing results (women with a positive HPV test but normal cytology underwent more frequent surveillance but were not referred to colposcopy based on the initial positive HPV test result).

Study quality

Quality was assessed using the following markers for internal validity: method of randomization, blinding, intention-to-treat analysis, and loss to follow-up. Blinding was assessed with a point given for each entity blinded: patient, clinician, pathologist, cytologist, and statistician.

Study quality was determined by using the Quality Assessment of Diagnostic Accuracy (QUADAS) questionnaire, an evidence-based tool that assesses the quality of diagnostic accuracy studies.¹⁵ Although a 14-point QUADAS tool is used primarily to assess the quality of observational studies rather than randomized trials of diagnostic tests, we used a modified 11-point assessment to provide additional information on study quality. The modified QUADAS tool covered the following topics: representativeness of the study population to a screening population, description of selection criteria, validity of reference standard, the delay between index and reference tests, the reproducibility of these tests, blinding, clinical information available to clinicians, and reporting of inadequate results and loss to follow-up. Based on the criteria used in previous studies, we chose a score of 50% or higher (at least 6 of 11 points) to denote a high-quality study.¹⁵

Values abstracted and computed

Disease outcome measures for the purposes of this study were cervical intraepithelial neoplasia (CIN)2 or worse (CIN2 or greater) and CIN3 or worse (CIN3 or greater). For each study, the rates of HPV test positivity/abnormal cytology rates or percent referred to colposcopy were either abstracted or calculated. When the positivity/abnormality rate was not recorded, the colposcopy referral rate was used instead because these values should theoretically be the same. The only exception was for passive-response studies in which a positive HPV test did not result in a referral to colposcopy. The only passive-response study that did not give a colposcopy referral rate was phase 1 of the New Technologies for Cervical Cancer Screening (NTCCS) trial in the 25-34 year age group.¹⁶ Hereafter, the term, test positivity rate, will signify HPV positivity rate, abnormal cytology rate, or colposcopy referral rate.

Using these values, the rates of CIN2 or greater and CIN3 or greater per positive test result were calculated by multiplying the test positivity rate by the total number of participants in each group and using that result as the denominator and the total number of CIN cases as the numerator. To calculate the rate of disease per person screened, the total number of participants was used as the denominator. Most of these data were not available for the second round of studies with multiple screening rounds; therefore, these measures were reported only for the first round of screening.

To assess the performance of the different strategies, the values for relative sensitivity, specificity, and relative positive predictive value (PPV) for HPV testing vs cytology-based strategies were directly abstracted from the publication if available. If not stated in the text, relative sensitivity was computed by taking the ratio of disease rate in the HPV based-testing group divided by the disease rate in the cytology-based testing group.

Comparisons of test specificity were reported for the A Randomised Trial of HPV Testing in Primary Cervical Screening (ARTISTIC) trial¹⁸ and the Finnish trial.¹⁶ The relative PPV was used as a marker for specificity and was calculated with the number of cases found divided by the number of participants who underwent colposcopy used as the denominator. The relative sensitivities for the Population Based Screening Study Amsterdam (POBASCAM) trial^{13,14} and relative sensitivities and relative PPVs for the India trial¹⁷ were calculated in the manner described in the previous text. The relative sensitivities and PPVs were calculated for the ARTISTIC trial using published sensitivities and PPVs.¹⁸

One potential harm associated with cervical cancer screening is the excess number of colposcopies and biopsies needed to detect 1 case of disease (number needed to colposcopy [NNC]). The NNC for CIN2 or greater and CIN3 or greater was calculated for round 1 as the reciprocal of the number of CIN cases per positive test result. The NNC was not calculated for the NTCCS phase 1 study¹⁶ in 25-34 year olds because the colposcopy rate was not given. The SwedeScreen study¹⁹ was also not included because overall sensitivities and PPVs were not published. The colposcopy rate used for the POBASCAM trial was reported in the earlier publication.¹⁴

Statistical analysis

Statistical methods used to evaluate sensitivities varied between individual studies and included a χ^2 test, a Fisher exact test, and Poisson regression analysis. For the studies in which we calculated relative sensitivities based on the published data, statistical comparison was not performed.

The NNC values were compared statistically by calculating the z-statistic on the SE of the difference between the NNC for the HPV arm and the cytology arm. A *P* value was then determined for this z-statistic, and the statistical significance was defined as $P < .05$.

Results

Study characteristics and quality

Six studies met inclusion criteria and were chosen after undergoing the evaluation process described in the previous text (Figure 2).^{13,14,16-20} Covariate characteristics such as study location, study size, and exclusion criteria are summarized in Table 1. The sample sizes ranged from 12,410 to 131,746 women; combined the studies included a total of 422,084 women. Four studies had data available from 2 screening rounds. A round is defined by a screening episode; in the second round, participants from the first round were rescreened. The time between screening intervals ranged from 3 to 5 years. Outcome measures ranged from CIN1 or greater to cancer death. Cancer is rare in countries with established screening programs and the numbers of cases were small in most studies.

The NTCCS study¹⁶ evaluated 3 different HPV testing-based methodologies compared with cytology in separate populations imbedded in 1 large study. Phase 1 evaluated a combination of HPV and cytology with differing responses to the HPV test based on age: younger women received passive follow-up, whereas older women received active follow-up (Figure 1). Phase 2 enrolled additional participants to evaluate HPV testing alone compared with cytology in all ages. The different trials were analyzed individually. Although the NTCCS trials categorized results based on age, most studies did not do this for all of their outcomes.

The studies differed in important methodological aspects such as HPV testing methods and thresholds for colposcopy (Table 1). The POBASCAM^{13,14} and SwedeScreen¹⁹ studies used general primer pair GP5-GP6 for HPV DNA polymerase chain reaction testing, whereas the other studies used the Hybrid Capture 2 (hc2) high-risk HPV test (Qiagen, Gaithersburg, MD).

Some studies used liquid based-cytology (LBC), whereas others used conventional cytology. Phase 1 of the NTCCS¹⁶ study used conventional cytology in the control arm and LBC combined with HPV testing in the experimental arm. Also, NTCCS¹⁶ used HPV testing and/or cytology in round 1 but only cytology in round 2 for all participants, meaning that the randomized screening strategy was not continued into round 2. Similarly, POBASCAM^{13,14} performed both HPV testing and cytology in all participants in round 2. Most importantly, the cytological and temporal thresholds for colposcopy referral varied between studies.

Two studies used ASCUS, 1 used low-grade squamous intraepithelial lesion (LSIL), and 3 used high-grade squamous intraepithelial lesion (HSIL) as the threshold for colposcopy. One study with passive response retested at 6 and 18 months, and 3 studies with passive response retested at 12 months (Table 1).

The presence of study quality measures is summarized in Table 2. Studies ranged in duration from 6.5 to 8 years. All studies had 2 rounds over this time period with the exception of the India study.¹⁷ The second round of the Finnish study²⁰ is still in progress. Most studies were of good quality, scoring 9-10 of 11 with the QUADAS tool. Although still of high quality, the Finnish study²⁰ had the lowest QUADAS score (6 of 11) because participant selection criteria were not clearly defined, the delay between positive index test and colposcopy was not described, colposcopy methods were not described in detail, and unsatisfactory test results and loss to follow-up were not reported.

Given that the HPV testing strategy, the HPV testing method, the cytology threshold for colposcopic referral, the type of cytology (LBC vs conventional), and age were heterogeneous in the included studies, a formal meta-analysis could not be conducted.

Accuracy of cytology vs HPV testing

To compare the numbers of interventions required to find disease, the test positivity rates and the rates of disease per woman screened and per woman who tested positive in round 1 are presented in Table 3. The test positivity for the HPV-based strategies ranged from 1.2% to 13.1% and the test positivity for the cytology-based strategies ranged from 1.2% to 7.0%. The wide range in positivity rates for HPV-based strategies were most likely related to population differences in HPV prevalence and different responses to positive HPV tests.

The range in cytology-based test positivity generally varied according to the cutoff threshold for colposcopy. For example, POBASCAM^{13,14} had a low abnormal cytology rate of 1.3%, but the threshold for colposcopy was a diagnosis of HSIL or worse. Higher rates of abnormal cytology (3.1%) were observed in NTCCS¹⁶ using a colposcopy threshold of ASCUS or worse. However, this pattern did not hold true in the ARTISTIC trial,¹⁸ in which the test positivity rate for cytology was 5.2% despite the colposcopy threshold being HSIL or worse.

Rates of disease per woman screened also varied widely. The rate of CIN2 or greater per woman screened ranged from 0.4–2.5% in the HPV testing groups and from 0.3–2.2% for cytology alone, whereas the rate of CIN3 or greater ranged from 0.1% to 1.3% in both the HPV-based testing groups and cytology groups.

Table 4 tabulates the relative sensitivities and relative PPVs for each round of each study, with cytology alone as the reference standard. In general, HPV testing was more sensitive for CIN2 or greater and CIN3 or greater than cytology in the first round of screening and less sensitive in the second round of screening.

The overall relative sensitivities of HPV-based strategies for detecting CIN3 or greater ranged from 0.9 to 2.1. For the 4 studies that used a strategy with a passive response to a positive HPV test result, the overall relative sensitivities for CIN3 or greater ranged from 0.9 to 1.1. For the NTCCS studies¹⁶ with an active response to a single positive HPV result, the relative sensitivities for CIN3 or greater ranged from 1.6 to 2.1 and were statistically significant. Thus, higher sensitivities for HPV testing-based strategies were observed only for strategies that incorporated immediate referral to colposcopy based on a single initial positive HPV test.

Regarding cancer outcomes, the NTCCS trial showed significantly decreased detection of invasive cancer during the second round of screening and overall after 2 rounds of screening in the HPV group.¹⁶ The India trial showed no difference in cancer detection between the HPV-only testing group and the cytology-only testing group; however, the incidence rates of stage II or higher cervical cancer and death from cervical cancer were higher in the cytology group than in the HPV testing group.¹⁷ The SwedeScreen trial described the numbers and histology of the cancers detected in the intervention and control groups; the total numbers were small and the authors did not perform a statistical comparison.¹⁹ The POBASCAM study reported that the number of cancers detected in the second round of screening was significantly lower in the intervention group than the control group, but the difference was not statistically significant overall after 2 rounds of screening.¹³

Only 2 studies reported specificities. The ARTISTIC trial reported that the specificity for detecting CIN2 or greater during round 1 was lower with cytology with HPV cotesting

compared with cytology alone.¹⁸ Conversely, the Finnish trial reported a higher specificity when comparing HPV testing with cytology triage with cytology alone.²⁰

Because specificities were not given for most studies, the relative PPV for HPV-based strategies compared with cytology-based strategies was used as a marker for specificity. Most studies did not report relative PPVs for round 2 or overall for both rounds. For round 1, the relative PPVs of HPV-based strategies compared with cytology-based strategies for detection of CIN2 or greater ranged from 0.4 to 1.3 and of CIN3 or greater ranged from 0.2 to 1.2. In the studies in which the threshold for colposcopy was ASCUS or greater, the relative PPV ranged from 0.2 to 0.9, whereas the relative PPV for the ARTISTIC trial¹⁸ was 1.0 in which only women with HSIL or worse were referred to colposcopy.

Numbers of colposcopies needed to detect 1 case of disease

To compare the potential harms of the various strategies, we calculated the numbers of women who would need to undergo colposcopy to detect 1 case of disease from round 1 of each study (CIN2 or greater and CIN3 or greater in Figures 3 and 4, respectively). For detection of CIN2 or greater, significantly higher numbers of colposcopies were needed in the HPV arm in the India trial,¹⁷ the NTCCS phase 1 trial in women 35-60 years of age,¹⁶ and the ARTISTIC trial,¹⁸ whereas significantly higher numbers of colposcopies were needed in the cytology arm in the Finnish trial²⁰ (Figure 3). For detection of CIN3 or greater, significantly higher numbers of colposcopies were needed in the HPV arm in the NTCCS phase 1 trial in women 35-60 years of age¹⁶ and in the ARTISTIC trial¹⁸ (Figure 4). For the India study,¹⁷ which reported cancer rather than CIN3 or greater as an outcome, the NNC values were 27.7 and 14.8 for the HPV-based and cytology-based groups, respectively, to find 1 case of cancer; this difference was statistically significant ($P < .0001$).

Comment

This systematic review synthesizes and compares results from all of the randomized trials with published results comparing cytology with HPV-based testing for cervical cancer screening. Although cervical cancer morbidity and mortality are the most relevant outcomes, cancer is rare in countries with established screening programs. Given that CIN2 reverts in up to 40% of cases,²¹ detection of CIN2 may result in overtreatment and morbidity in reproductive-aged women; therefore, CIN3 or greater is likely the most clinically relevant outcome in cervical cancer screening trials. One consistent finding from all the studies is that CIN3 or greater is an uncommon outcome, which necessitates a careful assessment of tradeoffs between sensitivity and specificity. The low rates of CIN3 or greater are of additional significance, given that the intervals between screening rounds in the studies (3-5 years) are greater than the screening interval of 1-2 years previously used in the United States until the recent introduction of new guidelines.^{5,6}

The results presented here indicate that in randomized trials, HPV testing is significantly more sensitive for CIN3 or greater during the first round of screening for some, but not all, strategies. The fact that sensitivity of HPV testing is lower than that of cytology in the second round suggests that as more disease is detected and treated in the first round, there will be less disease to detect subsequently. After completion of both screening rounds, sensitivity for detection of CIN3 or greater was significantly increased only with strategies involving an active response to a positive HPV test (ie, immediate colposcopy).^{16,17} This approach, however, also increases false-positive results as indicated by the relative PPV for detection of CIN3 or greater of 0.50 for women aged 35-60 years in the NTCCS trial.¹⁶

The strategy that has been adopted in the United States is to add HPV testing to cytology for women over age 30 years, and to respond passively to HPV-positive results if the cytology

is normal, rather than to perform immediate colposcopy.^{5,6} Four trials in our analysis used a similar passive response with cotesting.^{13,16,18,19} None of these trials demonstrated an overall significantly increased relative sensitivity for CIN3 or greater, and only 1 trial demonstrated a significantly increased relative sensitivity for CIN2 or greater.

Another important metric is the number of colposcopies needed to find a case of disease. In the ARTISTIC trial, significantly higher numbers of colposcopies were needed in the HPV cotesting arm for detection of CIN3 or greater in round 1, but this did not result in increased sensitivity. The calculated number of colposcopies likely underestimates the true number of colposcopies because it reflects the initial test positivity rate and does not include subsequent colposcopies resulting from increased surveillance in women who remain HPV positive over time. Thus, it will be important to monitor the impact of cotesting on overall colposcopy rates as the new screening guidelines are adopted.

There are many strengths of this review. All studies were high-quality randomized trials with large study populations and provide longitudinal data about disease detection over multiple rounds of screening. This review also provides a novel method to estimate and compare the impact of different screening strategies through the calculation of the number of colposcopies needed to detect a single case of CIN2 or greater or CIN3 or greater.

This review also has limitations. We chose to review randomized controlled trials based on the strength of their study methodology and, furthermore, assessed study quality using a validated tool. However, there were weaknesses in individual studies that fall outside the parameters measured in this tool. For example, 3 of the trials used HPV testing on women under the age of 30 years (Table 1), an age group in which HPV testing is less specific because of the higher prevalence of transient HPV infections. The rate of loss to follow-up was high in some of the studies (Table 2). As discussed in the *Results* section, in 2 of the trials, the randomized screening strategy was not continued into the second round of screening.^{13,14,16}

Additional limitations include the fact that results on relative specificity of HPV testing compared with cytology were not reported in most of the trials, limiting the ability to compare test performance. The varied strategies for incorporating HPV testing and the differing thresholds for colposcopy used by the studies presented significant challenges in comparing the performance of the screening strategies across studies. The studies were performed before the widespread uptake of HPV vaccination, and test performance is likely to change as the prevalence of HPV decreases. The number of screening rounds studied in these trials to date is insufficient to assess the impact of HPV testing over a woman's lifetime. Further data will be forthcoming from additional rounds of screening from some of these trials. In addition, all of these studies were completed outside the United States and followed their respective country's screening guidelines, all of which differed from US guidelines.

This review highlights the need for a longitudinal randomized trial performed in the United States over multiple screening rounds of 3-5 year intervals in which relative specificities and rates of colposcopy are reported. In addition, triage of women with HPV-positive results directly to colposcopy should be considered in any prospective US trial.

In summary, this systematic review indicates that after 2 rounds of screening, HPV-testing based screening strategies are more sensitive than cytology for the detection of CIN3 or greater only when referral to colposcopy follows a single positive HPV test. This strategy results in more colposcopies needed to detect a single case of CIN3 or greater or cancer, especially in women over 35 years of age. Because CIN3 and cervical cancer are rare in well-screened populations, the impact on increased disease detection needs to be balanced

with the impact on cost, numbers of colposcopies, and morbidity associated with potential overtreatment.

Acknowledgments

We thank Dr John Boscardin, Associate Professor of Medicine and Biostatistics, University of California, San Francisco, School of Medicine, for assistance in statistical analysis.

This study was supported the University of California, San Francisco, Dean's Quarterly Research Fellowship (I.Y.P.).

REFERENCES

1. US National Institutes of Health. National Cancer Institute. [Accessed June 21, 2011] Surveillance Epidemiology and End Results. SEER stat fact sheets: cervix uteri. Available at: <http://seer.cancer.gov/statfacts/html/cervix.html>
2. Gay JD, Donaldson LD, Goellner JR. False-negative results in cervical cytologic studies. *Acta Cytol.* 1985; 29:1043–6. [PubMed: 3866457]
3. Munoz N, Bosch FX, de Sanjosé S, et al. human papillomavirus types associated with cervical cancer. *N Engl J Med.* 2003; 348:518–27. [PubMed: 12571259]
4. Moscicki AB, Shiboski S, Broering J, et al. DNA testing in adolescent and young women. *J Pediatr.* 1998; 132:277–84. [PubMed: 9506641]
5. Moyer VA, on behalf of the US Preventative Services Task Force. US Preventive Services Task Force Recommendation Statement. *Ann Intern Med.* 2012; 156:880–91. W312. [PubMed: 22711081]
6. Saslow D, Solomon D, Lawson HW, et al. Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. *CA Cancer J Clin.* 2012; 62:147–72. [PubMed: 22422631]
7. American Society for Colposcopy and Cervical Pathology. [Accessed June 21, 2011] HPV genotyping clinical update, 2009. Available at: <http://www.asccp.org/ConsensusGuidelines/HPVGenotypingClinicalUpdate/tabid/5963/Default.aspx>
8. 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):S3/78–89. [PubMed: 16950021]
9. Koliopoulos G, Arbyn M, Martin-Hirsch P, Kyrgiou M, Prendiville W, Paraskeva E. Diagnostic accuracy of human papillomavirus testing in primary cervical screening: a systematic review and meta-analysis of non-randomized studies. *Gynecol Oncol.* 2007; 104:232–46. [PubMed: 17084886]
10. Macaskill P, Walter S, Irwig L, Franco E. diagnostic performance when combining two diagnostic tests. *Stat Methods.* 2002; 21:2527.
11. International Agency for Research on Cancer. IARC Handbooks of cancer prevention, vol. 10: cervix cancer screening. Chapter 4: efficacy of screening. Available at: <http://www.iarc.fr/en/publications/pdfs-online/prev/handbook10/handbook10-chap4.pdf> [Accessed Feb. 2, 2012]
12. Ogilvie GS, van Niekerk DJ, Krajden M, et al. A randomized controlled trial of human papillomavirus (HPV) testing for cervical cancer screening: trial design and preliminary results (HPV FOCAL trial). *BMC Cancer.* 2010; 10:111. [PubMed: 20334685]
13. Rijkaart D, Berkhof J, Rozendaal L, et al. Human papillomavirus testing for the detection of high-grade cervical intraepithelial neoplasia and cancer: final results of the POBASCAM randomised controlled trial. *Lancet Oncol.* 2012; 13:78–88. [PubMed: 22177579]
14. Bulkman NW, et al. up of a randomised controlled implementation trial. *Lancet.* 2007; 370:1764–72. [PubMed: 17919718]
15. Roposch A, Moreau NM, Uleryk E, Doria AS. of reporting of diagnostic accuracy for US. *Radiology.* 2006; 241:854–60. [PubMed: 17053199]
16. Ronco G, Giorgi-Rossi P, Carozzi F, et al. cervical intraepithelial neoplasia: a randomised controlled trial. *Lancet Oncol.* 2010; 11:249–57. [PubMed: 20089449]

17. Sankaranarayanan R, Nene BM, Shastri SS, et al. screening for cervical cancer in rural India. *N Engl J Med.* 2009; 360:1385–94. [PubMed: 19339719]
18. Kitchener HC, Almonte M, Thomson C, et al. cervical screening (ARTISTIC): a randomised controlled trial. *Lancet Oncol.* 2009; 10:672–82. [PubMed: 19540162]
19. Naucler P, Ryd W, Törnberg S, et al. Papanicolaou tests to screen for cervical cancer. *N Engl J Med.* 2007; 357:1589–97. [PubMed: 17942872]
20. Leinonen M, Nieminen P, Kotaniemi-Talonen L, et al. vs conventional cytology in a randomized setting. *J Natl Cancer Inst.* 2009; 101:1612–23. [PubMed: 19903804]
21. Moscicki AB, Ma Y, Wibbelsman C, et al. neoplasia 2 in adolescents and young women. *Obstet Gynecol.* 2010; 116:1373–80. [PubMed: 21099605]

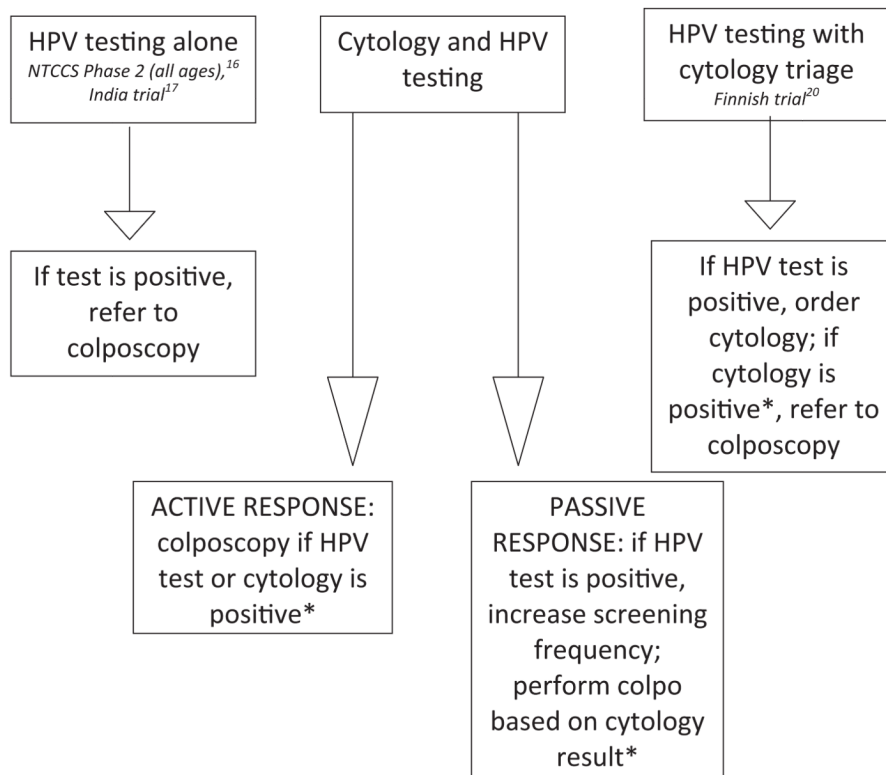


FIGURE 1. Strategies for incorporating HPV testing into cervical cancer screening
 Studies used different strategies for incorporating HPV testing into cervical cancer screening and for responding to positive HPV test results.
HPV, human papillomavirus.
 *Refer to Table 1 for colposcopy thresholds for positive cytology results.
 Patanwala. *HPV testing in cervical cancer screening. Am J Obstet Gynecol* 2013.

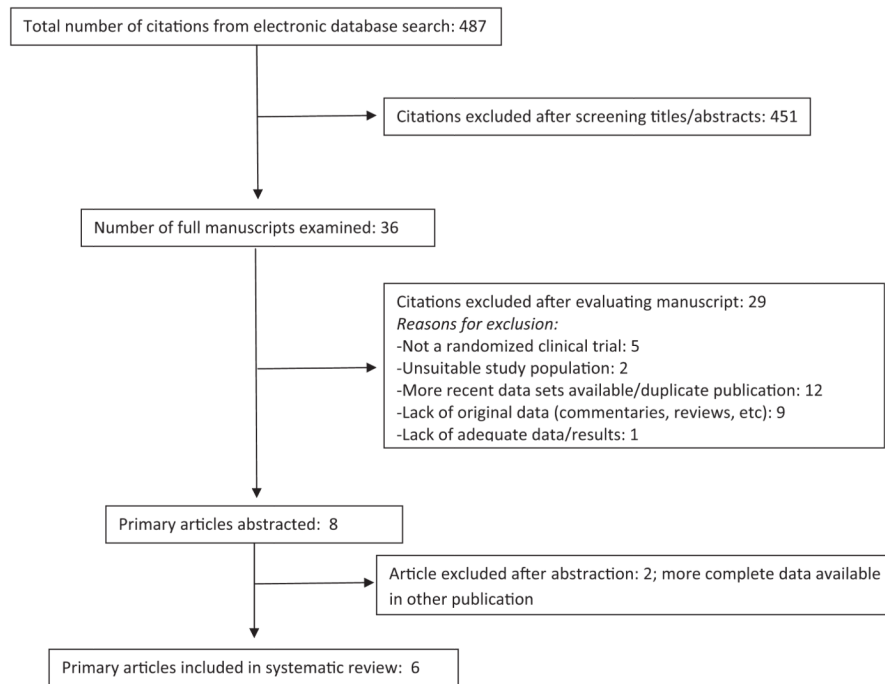


FIGURE 2. Systematic review study selection process

Flow diagram of articles identified and excluded for the systematic review is shown.
Patanwala. HPV testing in cervical cancer screening. Am J Obstet Gynecol 2013.

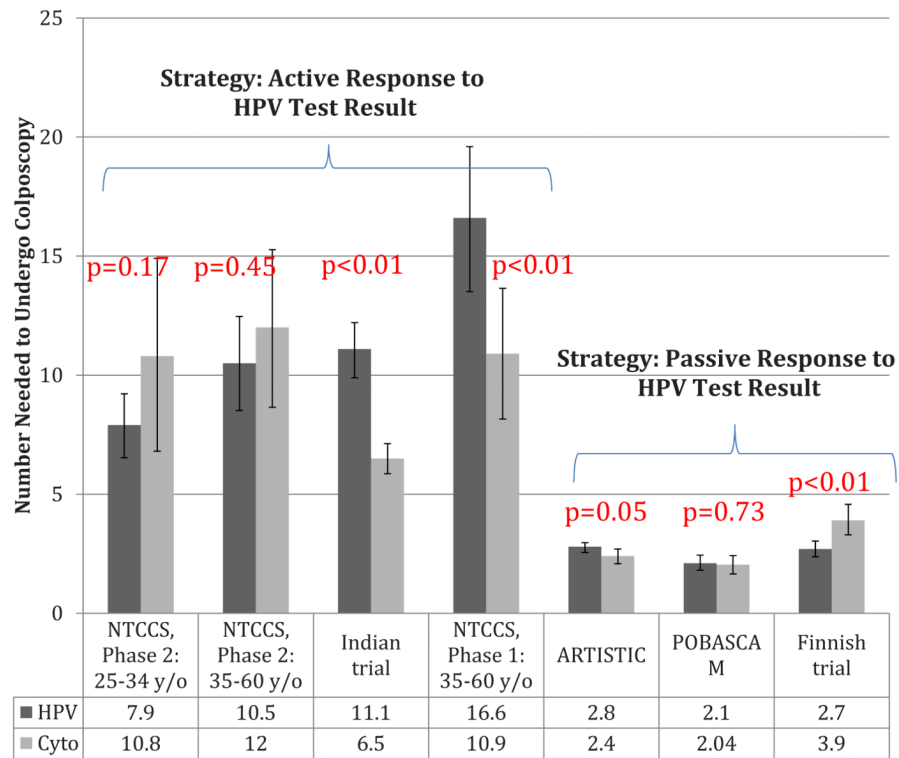


FIGURE 3. Number of colposcopies needed to detect 1 case of CIN2 or greater in round 1
 Values are calculated as described in *Materials and Methods*.
 CIN, cervical intraepithelial neoplasia.
 Patanwala. HPV testing in cervical cancer screening. *Am J Obstet Gynecol* 2013.

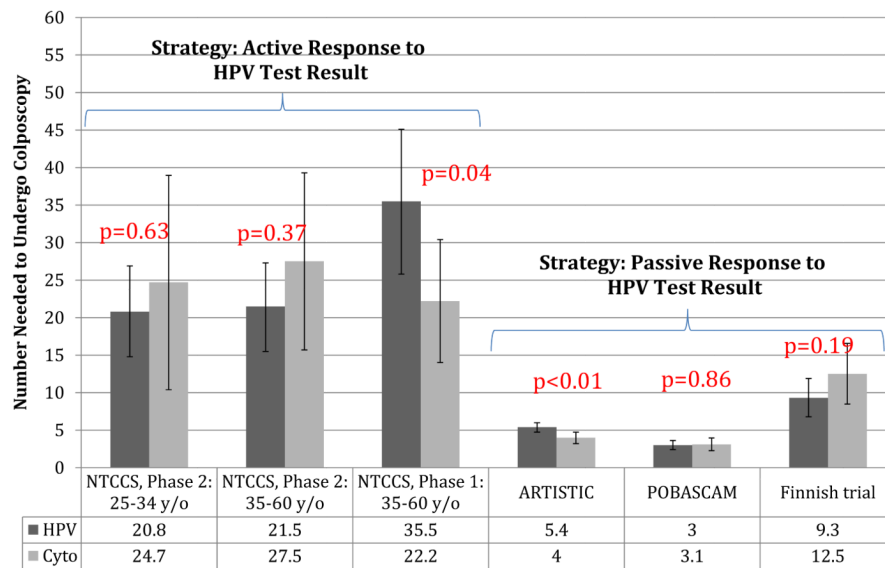


FIGURE 4. Number of colposcopies needed to detect 1 case of CIN3 or greater in round 1
 Values are calculated as described in *Materials and Methods*.
 CIN, cervical intraepithelial neoplasia.
 Patanwala. HPV testing in cervical cancer screening. *Am J Obstet Gynecol* 2013.

TABLE 1

Characteristics of trials included in systematic review

Author, year, country	Comparison	Exclusion criteria	HPV test used ^d	Cytology method	Outcome measured	Threshold for colposcopy	Size, N _C /N _{HPV} ^b
Ronco et al. ¹⁶ 2010, Italy (NTCCS trial) Total n = 94,370	Phase 1, 25-34 y old: cytology alone vs HPV plus cytology (passive response)	Pregnant; history of hysterectomy; treatment for CIN in past 5 y; age 25 y or younger or 60 y or older	hc2	Conventional and LBC	CIN2 or greater; CIN3 or greater; cancer ^c	Cyto: ASCUS or greater (7/9 centers) or LSIL or greater (2/9 centers) HPV: cytology positive as above or HPV test remains positive 12 months after initial testing	12,410 5808/6602
	Phase 1, 35-60 y old: cytology alone vs HPV plus cytology (active response)					Cyto: ASCUS or greater (7/9 centers) or LSIL (2/9 centers) HPV: positive HPV test or positive cytology, as above	33,364 16,658/16,706
	Phase 2, all ages: cytology alone vs HPV alone					Cyto: ASCUS or greater (7/9 centers) or LSIL or greater (2/9 centers) HPV: positive HPV test	49,196 24,535/24,661
Leinonen et al. ²⁰ 2009, Finland (Finnish trial)	Cytology alone vs HPV plus cytology triage	Age 25 y or younger or 65 y or older	hc2	LBC	CIN1 or greater; CIN2 or greater; CIN3 or greater	Cyto: LSIL or greater HPV: positive HPV test followed by cytology LSIL or greater	71,337 35,500/35,837
Kitchener et al. ¹⁸ 2009, United Kingdom (ARTISTIC trial)	Cytology alone vs HPV plus cytology (passive response)	Age 20 y or younger or 60 y or older	hc2	Conventional	CIN2 or greater; CIN3 or greater	Cyto: HSIL or greater, 2 consecutive LSIL results, or 3 consecutive ASCUS results HPV: HPV test remains positive 12 months after initial testing or HSIL or greater	24,510 6124/18,386
Sankaranarayanan et al. ¹⁷ 2009, India (India trial)	Cytology alone vs HPV alone	Pregnant; history of cervical cancer, hysterectomy, or uterine prolapse; no history of marriage; age 30 y or younger or 59 y or older	hc2	Conventional	Cancer, cancer death	Cyto: ASCUS or greater HPV: positive HPV test	131,746 32,058/34,126
Rikjaart et al. ¹³ 2012, Netherlands (POBASCAM trial)	Cytology alone vs HPV plus cytology (passive response)	History of hysterectomy; CIN2 or greater or worse, or abnormal cytology in previous 2 y; age 29 y or younger or 56 y or older	GP5/GP6	Conventional	CIN2 or greater; CIN3 or greater	Cyto: HSIL or greater initially or repeated ASCUS results at 6 months HPV: HPV test remains positive at 6 or 18 months after initial testing or positive cytology, as above	40105 19999/20106

Author, year, country	Comparison	Exclusion criteria	HPV test used ^a	Cytology method	Outcome measured	Threshold for colposcopy	Size, N _C /N _H ^b
Naucler et al, ¹⁹ 2007, Sweden (SwedeScreen trial)	Cytology alone vs HPV plus cytology (passive response)	Age 32 y or younger or 38 y or older	GP5/GP6	Conventional	CIN2 or greater, CIN3 or greater	Cyto: ASCUS or greater (Stockholm) or HSIL or greater (all other cities) HPV: HPV test remains positive 12 months after initial testing or positive cytology, as above	12527 6270/6257

ARTISTIC, A Randomised Trial of HPV Testing in Primary Cervical Screening; *ASCUS*, atypical squamous cells of undetermined significance; *CIN*, cervical intraepithelial neoplasia; *Cyto*, cytology; *G*, G serotype; *HPV*, human papillomavirus; *HSIL*, high-grade squamous intraepithelial lesion; *LBC*, liquid based-cytology; *LSIL*, low-grade squamous intraepithelial lesion; *NTCCS*, New Technologies for Cervical Cancer Screening; *P*, P genotype; *POBASCAM*, Population Based Screening Study Amsterdam.

Patanwala. HPV testing in cervical cancer screening. *Am J Obstet Gynecol* 2013.

^aThe hc2 (hybrid capture 2) was considered positive if 1 relative luciferase unit or greater;

^bN_C/N_H is the number of study subjects in the cytology group/number of study subjects in the HPV group;

^cIn most recent publication, CIN2+ signifies CIN2, CIN3, and AIS combined. CIN3+ signifies CIN3 and AIS combined.

TABLE 2

Quality of trials included in systematic review

Author, year	Rounds, n	Time frame	Randomization	Blinding ^a	Intention to screen	Loss to follow-up (HPV %/Cyto %)	QUADAS score (x/11)
Ronco et al. ¹⁶ 2010 (NTCCS trial)	2	February 2002 through November 2008	Computerized blocks and series of random numbers	1/5	Yes	Varies by phase	9
Leinonen et al. ²⁰ 2009 (Finnish trial)	1 ^b	January 2003 to present	Based on national ID number	None	Yes	NR	6
Kitchener et al. ¹⁸ 2009 (ARTISTIC trial)	2	July 2001 through July 2008	Series of random numbers	2/5	Yes	34.2/34.9	9
Sankaranarayanan et al. ¹⁷ 2009 (India trial)	1	January 2000 through December 2007	Method not stated	None	Yes	20.3/20.3	9
Rikjaart et al. ¹³ 2012 (POBASCAM trial)	2	January 1999 through February 2007	Series of random numbers	2/5	Yes	14.1/14.9 ^c	9
Naucler et al. ¹⁹ 2007 (SwedeScreen trial)	2	May 1997 through August 2005	Series of random numbers	3/5	Yes	12.9/12.2	10

ARTISTIC, A Randomised Trial of HPV Testing in Primary Cervical Screening; *Cyto*, cytology; *HPV*, human papillomavirus; *NR*, not relative; *NTCCS*, New Technologies for Cervical Cancer Screening; *POBASCAM*, Population Based Screening Study Amsterdam; *QUADAS*, Quality Assessment of Diagnostic Accuracy.

Patanwala. HPV testing in cervical cancer screening. *Am J Obstet Gynecol* 2013.

^a A point was assigned for each entity blinded: patient, clinician, pathologist, cytologist, statistician;

^b Data from 1 round is available, whereas the second round is still in progress;

^c Percentages calculated from raw numbers regarding eligibility and attendance published in paper.

TABLE 3
Round 1: positivity rate, disease rate per woman screened, disease rate per positive result

Author	Comparison	Age range, y	Positivity rate, (%) ^a		Rate per woman screened (CIN2 or greater), %		Rate per woman screened (CIN3 or greater), %		Rate per positive result (CIN3 or greater), %			
			HPV	Cyto	HPV	Cyto	HPV	Cyto	HPV	Cyto		
Ronco et al ¹⁶ (2010) (NTCCS)	HPV plus cytology (passive response) vs cytology alone	25-34	11.7	4.6	1.2	0.6	10.1	17.7	0.4	0.4	3.0	11.0
Ronco et al ¹⁶ (2010) (NTCCS)	HPV plus cytology (active response) vs cytology alone	35-60	10.6	3.6	0.6	0.3	6.0	9.2	0.3	0.2	2.8	4.5
Ronco et al ¹⁶ (2010) (NTCCS)	HPV alone vs cytology alone	25-34	13.1	4.0	1.7	0.4	12.7	9.2	0.6	0.2	4.8	4.1
Ronco et al ¹⁶ (2010) (NTCCS)	HPV alone vs cytology alone	35-60	5.8	3.1	0.5	0.3	9.5	8.4	0.3	0.1	4.7	3.6
Leinonen et al ²⁰ (2009) (Finnish trial)	HPV plus cytology triage vs cytology alone	25-65	1.2	1.2	0.4	0.3	37.0	25.4	0.1	0.1	10.7	8.0
Kitchener et al ¹⁸ (2009) (ARTISTIC trial)	HPV plus cytology (passive response) vs cytology alone	20-64	6.8	5.2	2.5	2.2	36.2	41.8	1.3	1.3	18.6	25.1
Sankaranarayanan et al ¹⁷ (2009) (India trial) ^b	HPV alone vs cytology alone	30-59	10.3	7.0	0.9	1.1	9.1	15.4	0.4	0.5	3.6	6.8
Rikjaart et al ¹³ (2012) (POBASCAM trial) ^c	HPV plus cytology (passive response) vs cytology alone	29-56	2.3	1.3	1.1	0.7	47	49	0.8	0.5	33	32
Naucler et al ¹⁹ (2007) (SwedeScreen trial)	HPV plus cytology (passive response) vs cytology alone	32-38	NR	NR	1.8	1.2	NA	NA	1.2	0.9	NA	NA

ARTISTIC, A Randomised Trial of HPV Testing in Primary Cervical Screening; CIN, cervical intraepithelial neoplasia; Cyto, cytology; HPV, human papillomavirus; NA, not applicable; NR, not relative; NTCCS, New Technologies for Cervical Cancer Screening; POBASCAM, Population Based Screening Study Amsterdam; QUADAS, Quality Assessment of Diagnostic Accuracy.

Patanwala. HPV testing in cervical cancer screening. Am J Obstet Gynecol 2013.

^aWhen positivity rate was not recorded, the colposcopy referral rate was used instead because both values should theoretically be the same

^bCIN3 or greater was not an outcome measured in this study; instead the detection rates for CIN3 or greater actually refer to the detection rates for cancer

^cValues recorded in table abstracted and calculated based on numbers printed in earlier publication because the most recent publication does not give enough information to calculate a new positivity rate (Bulkman, 2007).¹⁴

TABLE 4
Relative sensitivities and relative positive predictive values of HPV based strategies compared with cytology-based strategies

Author	n	Round 1			Round 2			Overall				
		Relative sensitivity CIN2 or greater	Relative sensitivity CIN3 or greater	Relative PPV CIN2 or greater	Relative sensitivity CIN2 or greater	Relative sensitivity CIN3 or greater	Relative PPV CIN2 or greater	Relative sensitivity CIN2 or greater	Relative sensitivity CIN3 or greater	Relative PPV CIN2 or greater		
Ronco et al ¹⁶ (2010), HPV with cytology (passive response), 25-34 y old, NTCCS ^a	12,410 HPV: 6602 Cyto: 5808	2.0 ^b	0.9	0.6 ^b	0.9	1.3	NR	NR	1.7 ^b	1.0	0.8	0.3 ^b
Ronco et al ¹⁶ (2010), HPV with cytology (active response), 35-60 y old, NTCCS ^a	33,364 HPV: 16,706 Cyto: 16,658	1.9 ^b	1.9 ^b	0.4 ^b	0.7	0.7	NR	NR	1.7 ^b	1.6 ^b	0.6 ^b	0.5 ^b
Ronco et al ¹⁶ (2010), HPV alone, 25-34 y old, NTCCS ^a	13,725 HPV: 6937 Cyto: 6788	4.5 ^b	3.9 ^b	0.9	0.4 ^b	0.2 ^b	NR	NR	2.8 ^b	2.1 ^b	NR	NR
Ronco et al ¹⁶ (2010), HPV alone, 35-60 y old, NTCCS ^a	35,471 HPV: 17,724 Cyto: 17,747	2.1 ^b	2.4 ^b	0.8	0.3 ^b	0.3	NR	NR	1.6 ^b	1.7 ^b	NR	NR
Leinonen et al ²⁰ (2009), Finnish trial	71,337 HPV: 35,837 Cyto: 35,500	1.4 ^b	1.2	1.3 ^b	NA	NA	NR	NR	NA	NA	NR	NR
Kitchener et al ¹⁸ (2009), ARTISTIC trial	24,510 HPV: 18,386 Cyto: 6124	1.1	1.0	1.1	0.6 ^b	0.5 ^b	0.7	0.4	1.0	0.9	NR	NR
Sankaranarayanan et al ¹⁷ (2009), India trial	66,184 HPV: 34,126 Cyto: 32,058	0.9	1.0	0.6	NA	NA	NR	NR	0.9	1.0	0.6	0.6
Rikjaart et al ¹³ (2012), (POBASCAM trial) ^a	40,105 HPV: 19,999 Cyto: 20,106	1.3 ^b	1.2	NR	0.9	0.7 ^b	NR	NR	1.0	1.1	NR	NR

ARTISTIC, A Randomised Trial of HPV Testing in Primary Cervical Screening; CIN, cervical intraepithelial neoplasia; Cyto, cytology; HPV, human papillomavirus; NA, not applicable; NR, not relative; NTCCS, New Technologies for Cervical Cancer Screening; POBASCAM, Population Based Screening Study Amsterdam; PPV, positive predictive value.

Patanwala. HPV testing in cervical cancer screening. Am J Obstet Gynecol 2013.

^aRound 2 screening in Ronco et al¹⁶ was the same for both groups: cytology only. Round 2 screening in Rikjaart et al¹³ was also the same for both groups: HPV plus cytology;

^b Significant value. The significance of values for Kirchener¹⁸ (PPV's only), Sankaranaiyanan¹⁷, and Nauclet¹⁹ could not be verified because these values were calculated and not published in original papers.