

Published in final edited form as:

Int J Cancer. 2012 April 15; 130(8): 1879–1887. doi:10.1002/ijc.26194.

Switch from Cytology-based to HPV-based Cervical Screening: Implications for Colposcopy

Carolina Porras¹, Nicolas Wentzensen¹, Ana C. Rodríguez², Jorge Morales², Robert D. Burk³, Mario Alfaro², Martha Hutchinson⁴, Rolando Herrero², Allan Hildesheim¹, Mark E. Sherman¹, Sholom Wacholder¹, Diane Solomon¹, and Mark Schiffman¹

¹ Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA

² Proyecto Epidemiológico Guanacaste, Fundación INCIENSA, San José, Costa Rica

³ Albert Einstein Cancer Center, Albert Einstein College of Medicine, Bronx, New York

⁴ Women and Infants' Hospital, Providence, Rhode Island, USA

Abstract

Background—HPV testing is more sensitive than cytology; some cervical cancer prevention programs will switch from cytology to carcinogenic HPV test-based screening.

Objective—To evaluate the clinical implications of a switch to HPV test-based screening on performance and workload of colposcopy.

Methods—Women in the population-based, 7-year Guanacaste cohort study were screened at enrollment using cytology. We also took another specimen for HPV DNA testing and collected magnified cervical photographic images (cervigrams). A final case diagnosis (\geq CIN3, CIN2, or $<$ CIN2) was assigned at exit. Using the cervigram as a surrogate of colposcopy impression, we evaluated the impact of changing screening method from cytology to carcinogenic HPV testing on the distribution of enrollment colposcopic impression, and on the predictive values of positive and negative colposcopic impressions for the cumulative 7-year detection of \geq CIN2 and \geq CIN3.

Results—A program based on immediate colposcopic referral after positive HPV would immediately identify as high risk more of the cumulative \geq CIN2 cases than conventional cytology, because of an increased number of referrals. However, the proportion of women that would have visible lesions at referral to colposcopy, and the sensitivity-versus-specificity trade-off of the colposcopic impressions would be similar to programs using cytology (\geq ASCUS) for referral.

Conclusions—The major concern with switching from cytology to more sensitive HPV screening is management of the many HPV-positive women, including those with still non-visible \geq CIN2 lesions. Our data support the need for a non-visual diagnostic method to guide management and treatment of HPV-positive women.

Keywords

HPV testing; cytology; colposcopy; cervical cancer screening; cervicography

INTRODUCTION

Cervical cancer screening is designed to detect easily treatable cancer precursors, most commonly defined as cervical intraepithelial neoplasia grade 2 (CIN2) and greater, or more stringently as CIN3. The intention is to detect precursor lesions with prognostic accuracy, i.e., to find the lesions that will become cancer some years into the future to permit safe interval screening.

Multiple prospective observational studies¹⁻⁴ and randomized trials⁵⁻¹⁰ have shown that testing for carcinogenic types of human papillomavirus (HPV) is more sensitive than cytology. Prospectively, carcinogenic HPV testing (which we will simply call HPV testing) is much more sensitive, but less specific, than cytology for prediction of subsequent risk of CIN2, CIN3, and cancer (which together we will refer to as \geq CIN2). As a result, various countries are considering whether and how to switch from cytology-based to HPV test-based cervical screening. The specificity of HPV testing is inadequate particularly at young ages when transient sexually transmitted infections are very common. It is likely that a screening program based primarily on HPV, will be restricted to women past the peak of sexual transmission, e.g., 30 and older, as is currently approved for the use of the test in conjunction with cytology in the United States.¹¹ The ideal length of the screening interval for a test-based program is still under discussion; however, it is clear that given its high negative predictive value, it would be longer than the interval recommended for cytology based screening programs.

Currently, in cytology-based screening programs, colposcopy is used as the second test (diagnostic test) to distinguish among women with abnormal cytology those who need to be treated. If colposcopy remains as the diagnostic test after switching to HPV testing as the primary screening method, the central diagnostic role of colposcopy will remain but the clinical characteristics and size of the population seen at colposcopy will be altered. Referring only women who are carcinogenic HPV-positive would remove the substantial burden of examining women with only non-carcinogenic HPV type or HPV-negative, look-alike lesions that would never progress to cervical cancer. Thus, colposcopists would only examine patients with carcinogenic HPV infections, who in aggregate are at substantially elevated risk of subsequent \geq CIN2.

Despite its theoretical advantages due to increased sensitivity and causality the great majority of infections, especially new infections, would be benign, and ~90% would clear within 5 years.¹² Also, because HPV testing is so sensitive, it picks up infections very early in the carcinogenic process; early lesions may not be detectable in colposcopy. The need to evaluate so many infections, in the effort to choose those destined to lead to \geq CIN2, will challenge the capabilities of colposcopy which, when applied to small lesions is prone to sub-optimal reproducibility, and imperfect sensitivity and specificity.¹³

To evaluate how a switch to HPV-based screening might affect colposcopy practice, we used data from the population-based cohort study of HPV and cervical neoplasia in Guanacaste, Costa Rica. Specifically, we estimated the difference in colposcopic characteristics and colposcopic performance between the subsets of screened patients referred based on conventional cytology and the subsets of the screened patients that would have been referred in a program based on HPV testing. We evaluated the performance of colposcopy at the most sensitive threshold of abnormality (presence of acetowhite areas) comparing the predictive values of normal versus abnormal colposcopic impression for the detection of lesions (\geq CIN2 and \geq CIN3) through the 7 years of the study based on initial assessment. We included in our definition of disease outcome all prevalent and incident diagnoses of CIN2 and \geq CIN3 detected through the 7 years of the study because of recent

suggestions that HPV based screening would permit prolonged screening intervals approximately as long as the duration of our cohort study.¹⁴

METHODS

Study Procedures

The Guanacaste cohort study^{15, 16} was designed to investigate the role of HPV infection in the development of \geq CIN2, and to evaluate new screening technologies. In this population, a previous comparison of HPV testing to cytology confirmed the substantially increased sensitivity of HPV testing compared with conventional cytology.¹ The design and methods of the cohort study have been described in detail.^{15, 16} Briefly, approximately 20% of the census tracts in the area were selected randomly, and all resident women older than 18 years were enumerated (11,742 women). A total of 10,049 women (>90% participation among eligible women) were enrolled between June 1993 and December 1994. All participants signed informed consent and the protocol was approved by Costa Rican and the US National Cancer Institute Institutional Review Boards.

At enrollment, the cohort was intensively screened by a small group of experienced, highly-trained nurses. Women were screened by direct visual inspection, then a cytology specimen was taken using a Cervex Brush broom-like sampler (Unimar, Connecticut); two slides were produced, including a conventional Pap smear and liquid-based cytology (a split sample ThinPrep using the residual cells from the Pap smear placed into PreservCyt, Cytyc Corporation, now Hologic). The ThinPrep was interpreted by an expert screening/pathologist team in the United States. Of note, this technique was used during its developmental phase in which sensitivity of interpretation was emphasized over specificity leading to an extremely high rate of abnormal interpretations, for that reason the results of this technique were not included in this analysis. The conventional smear was screened in a clinically typical fashion in Costa Rica and read by an expert Costa Rican cytopathologist. The conventional smear was also interpreted using the semi-automated PapNet method (Neuromedical Systems, Suffern, NY) that selected abnormal cells and cell clusters that were screened and evaluated at Johns Hopkins University. PapNet is no longer widely used; moreover, its results resembled conventional cytology. For simplicity, it is not mentioned further in this manuscript.

After cytology and a second swab for HPV testing (see below), the nurses applied 5% acetic acid and took replicate magnified photographic images of the cervix (cervigrams) which were evaluated in the United States by a single expert evaluator (National Testing Laboratories, Fenton MO). The cervigrams were interpreted as: negative (no lesion seen), atypical (a lesion of doubtful significance was observed inside or outside of the transformation zone), P0 (remotely worrisome, although probably normal), P1 (compatible with an HPV induced low grade lesion), P2 (compatible with either CIN2 or CIN3) and P3 (compatible with cancer).

HPV detection and typing were done on a second specimen collected using a Dacron swab and stored in ViraPap DNA transport medium (Digene, now Qiagen) as described previously.¹⁷ In brief, extracted DNA was amplified using a MY09/M11 L1 degenerate primer PCR system with AmpliTaq gold polymerase. PCR products were detected by gel electrophoresis and hybridized with radiolabeled generic HPV DNA probes. Dot-blot hybridization was used for HPV typing using type-specific probes for HPV types: 6,11,16,18,26,31–33,35,39,40,45,51–56,58,59,61,62,64,66–74,81–85 and 89. Three experienced investigators who were masked to clinical outcomes interpreted each result, and discrepancies were resolved by consensus.

At enrollment, as shown in Figure 1, women were referred to colposcopy if they had any abnormal result (\geq ASCUS or atypical squamous cells of unknown significance) on any of the three cytologic methods, or positive cervigram (\geq P0, including all positive cervigram categories from P0 to P3). Also women with a possibly severe lesion suspected by visual inspection at the beginning of the pelvic exam were referred. Of note, the enrollment visit HPV testing results were not used to refer women to colposcopy. To guard against referral bias, a 2% random sample of all women regardless of screening results was referred to colposcopy but none had \geq CIN2. Because of the very liberal referral criteria meant to capture all women at possibly elevated risk of subsequent \geq CIN2, 24% of the 9175 women screened at enrollment were referred to colposcopy. The median time between the enrollment screening visit and the colposcopy visit was 90 days.

Based on the diagnosis given at the end of the enrollment phase women were followed up at different intervals. Women with cytologic LSIL or histologic CIN1 were re-screened every 6 months; women with ASCUS, cervigram P0 or P1, \geq 5 sexual partners and those positive for carcinogenic HPV-based on a now obsolete test called Hybrid Capture Tube were re-screened every year.¹⁶ Women with all negative screening at enrollment were seen again at 5–7 years, except for a random sample of 540 women followed yearly, which confirmed their extremely low risk for \geq CIN2.¹²

During follow up, only women with cytology or cervicography indicating possible incident \geq CIN2, or nurse concern over possible cancer were referred to colposcopy (Figure 1). In addition, at the exit phase of the cohort we referred women with cytologic results of ASCUS or LSIL, cervigrams of P0 and P1, or HPV 16 or HPV18 infection at either of the last two screening visits. We also referred women with ongoing persistent infection with other carcinogenic HPV types and a 6% random sample of the cohort.

The sensitivity of disease ascertainment in the Guanacaste cohort study was achieved by multiple screening visits over a long period, multiple screening methods, liberal colposcopic criteria. At enrollment and during follow-up lasting 7 years, women with suspected \geq CIN2 lesions were excluded from further follow-up and treated by excision if needed.

At colposcopy visits throughout the cohort study, a single colposcopist (JM) performed all treatments and biopsies under colposcopic guidance, if no lesions were seen no random biopsies were taken. Unless women with suspected \geq CIN2 lesions (e.g., HSIL cytology) were older and already had multiple children, a confirmatory biopsy was taken before treatment by excision. The histologic final diagnosis was assigned after review by an expert panel including two US experts. At the end of the study all women received a final case assignment as invasive cancer, CIN3, CIN2 or $<$ CIN2 according to an algorithm based on the independent masked review.

This analysis is based on the relationship of enrollment screening results, enrollment cervicography (to simulate enrollment colposcopic impressions), on histologic endpoints during the entire course (7 years) of the cohort study.

Statistical Analysis

Of the 10,049 women enrolled in the study, 8369 were included in this analysis after exclusions of women who at enrollment were virgins (n=583), refused the pelvic exam (n=291), had undergone hysterectomy (n=630), or had missing results for the cervigram (n=94), the HPV DNA testing (n=19), or the conventional cytology (n=63).

Cytology results were categorized as normal (normal, reactive changes) and \geq ASCUS (ASCUS, LSIL, HSIL, and rare glandular lesions). HPV DNA testing was categorized as

positive if there was at least 1 of the 13 carcinogenic HPV types as defined by the latest categorization of the International Agency for Research on Cancer (IARC) (HPV 16,18,31,33,35,39,45,51,52,56,58,59,68).¹⁸

To estimate how a change to HPV testing for primary screening would change colposcopic performance, we combined the categories of estimated colposcopic impressions into normal or abnormal and evaluated the actual disease yield for each category using final histologic disease outcomes. We did not attempt to evaluate finer categories of colposcopic abnormality because visual grading of severity of cervical lesions is inaccurate and not reproducible.¹⁹ The main reliable conclusion from previous work is that most women with \geq CIN3 show some abnormality (acetowhitening particularly), and we assessed colposcopic performance at this basic level.

Specifically, among women who screened positive for each method, we evaluated the positive predictive value (PPV) of abnormal colposcopic impression and negative predictive value (NPV) of normal colposcopic impression. It is important to re-emphasize that for the final diagnostic assignment, we included all prevalent and incident diagnoses of CIN2 and \geq CIN3 detected during the cohort study as the disease outcome. Thus, when we evaluated the performance of colposcopy we did not differentiate its diagnostic accuracy (performance in detecting the lesions that were present at enrollment) from prognostic accuracy (ability to predict future lesions), given that our aim was to evaluate its long-term performance in anticipation of a lengthened screening intervals with HPV testing.¹⁴

Finally, to look at the overall effect of changing from cytology screening followed by colposcopy to HPV screening followed by colposcopy, we calculated the overall programmatic sensitivity, specificity, PPV, and NPV in detecting \geq CIN3 for a HPV DNA test every 7 years leading to a single immediate colposcopy if positive versus conventional cytology at the same interval leading to a single immediate colposcopy if \geq ASCUS. We extended the consideration of programmatic accuracy to consider other screening strategies including the use of HPV testing for triage of ASCUS interpretations as is now commonly done, and HPV testing followed by conventional cytology triage. All the calculations were done stratifying by age group (<30 and \geq 30 years old), to permit a separate evaluation within the older group for which HPV testing is most commonly recommended.

Given that the two screening methods (HPV testing and conventional cytology) were performed on the same women, for comparing e.g. PPV, NPP we used the method described by Leisenring et al,²⁰ that incorporates GEE (Generalized Estimating Equations) methods in calculating p-values. The analysis was performed using STATA.

RESULTS

Validation of Cervicography as a Surrogate for Colposcopic Impression

Because we did not perform colposcopy at enrollment on all women in the cohort, we used the interpretation of the magnified photograph of the cervix (cervigram) taken at enrollment screening as a surrogate of colposcopic impression. The correlation of cervicographic impression and colposcopic impression is strong enough for this estimation, as previously reported.¹⁹

To ensure that the cervigram was a good surrogate of the colposcopic impression we evaluated their agreement among the 2,023 women referred to colposcopy during the enrollment period. Since cervigram interpretations of “Atypical” and “P0” could be considered as either normal or abnormal, we examined which categorization of these two cervigram interpretations yielded the highest agreement with the colposcopic impression.

We evaluated agreement by unweighted kappa (κ) statistics, which test whether observed agreement is beyond that expected by chance alone and the symmetry chi-square test. In general, kappa values are interpreted as follows: ≤ 0.20 , poor agreement; 0.21 to 0.40, fair agreement; 0.41 to 0.60, moderate agreement; 0.61 to 0.80, good agreement; > 0.80 , excellent agreement.

Table 1 shows the agreement between colposcopy impression and cervigram (considering Normal: Negative/P0; and Abnormal: Atypical/P1/P2/P3) that yielded the highest kappa value ($\kappa=0.73$). We chose this categorization. The agreement was even better ($\kappa=0.80$) when the women attended colposcopy within ≤ 90 days of enrollment, suggesting that cervicography as evaluated at that time by the NTL evaluator was a very good surrogate of enrollment colposcopic impression by our colposcopist. For completeness the other three possible categorizations of the cervigram had the following kappa values: 1. Normal (Negative, Atypical), Abnormal (P0, P1, P2, P3); $\kappa=0.54$; 2. Normal (Negative, Atypical, P0), Abnormal (P1, P2, P3); $\kappa=0.66$; 3. Normal (Negative), Abnormal (Atypical, P0, P1, P2, P3); $\kappa=0.62$. Thus, throughout the article, we use the term “estimated colposcopic impression” referring to the cervigram that was obtained in all women and which showed very good agreement with colposcopy in women for whom both tests were performed.

Effect of Switch to HPV Screening on Number of Colposcopic Referrals

Table 2 shows that a switch to HPV DNA testing as the primary screening method would increase the number of women referred to colposcopy (from 7.7% to 19.2%, among women < 30 years old and from 6.7% to 10.8% among women ≥ 30 years old). We compared the proportions by paired McNemar statistic, which tests the number of discordant results, the p values for both age groups were < 0.001 .

Effect of Switch to HPV Screening on Colposcopic Performance

As shown in Table 3, we evaluated the 7-year cumulative disease outcome for women with abnormal and normal estimated colposcopic impressions at baseline by age, depending on which screening test would have referred them to colposcopy.

As expected from a previous analysis in this Guanacaste cohort (1) and numerous other studies, HPV screening was more sensitive and less specific than conventional cytology. Among women < 30 years old, HPV captured 24 more $\geq \text{CIN}2$ cases than conventional cytology (Table 3a, 37 $\geq \text{CIN}2$ cases among 176 referred for cytology versus Table 3b, 61 of 440 referred for HPV). Among women aged 30 years or older, HPV captured 36 more $\geq \text{CIN}2$ cases (Table 3c, 84 $\geq \text{CIN}2$ cases among 409 referred for cytology versus 120 cases among 653 referred for HPV).

Although absolute numbers of colposcopic referrals and the number of $\geq \text{CIN}2$ cases would increase with use of HPV testing, the *percentage* of referred women with abnormal colposcopic impressions (as estimated by cervicography) was very similar regardless of which method was used for referral to colposcopy (among women < 30 years: 37.5% for conventional, 34.8% for HPV testing, $p=0.44$; among women ≥ 30 years old: 22.0% for conventional, 18.8% for HPV testing, $p=0.09$).

The two most revealing metrics of colposcopic performance are also highlighted in table 3: positive predictive value (PPV) or absolute risk of $\geq \text{CIN}2$ among women with abnormal colposcopy and negative predictive value (NPV) or reassurance against $\geq \text{CIN}2$ among those with normal colposcopy.

We observed that a switch from conventional cytology to HPV screening would not improve these critical measures of colposcopic performance. Although our data suggested that the

switch to HPV screening would reduce the PPV (risk of \geq CIN2) following abnormal colposcopic impressions, the difference was not significant. Among women <30 years old, the PPV dropped from 31.8% for abnormal colposcopy after abnormal cytology (Table 3a) to 24.2% for abnormal colposcopy after positive HPV (Table 3b); $p=0.10$. A switch to HPV screening would increase the NPV of normal colposcopic impressions (from 85.5% after abnormal conventional cytology to 91.6% after HPV positivity, $p=0.02$). Among women ≥ 30 years old, the PPV would drop from 47.8% for abnormal colposcopy after abnormal conventional cytology (Table 3c) to 41.5% for abnormal colposcopy after positive HPV (Table 3d); $p=0.15$; and would barely change the NPV of normal colposcopic impressions (~87% for both, $p=0.92$).

It is important that many HPV positive women ≥ 30 years old (Table 3d) who were ultimately found to have \geq CIN2 did not have abnormal estimated colposcopic impressions at enrollment. Raising the colposcopy cutoff to women with more obvious acetowhite lesions (i.e., ignoring Atypical, and including only P1, P2 or P3), only 9 of 38 (23.7%) of the CIN2 and 29 of 82 (35.4%) of the CIN3/cancer cases detected or predicted by HPV testing would have been classified as colposcopically abnormal if referred immediately.

Evaluating Accuracy of Screening and Colposcopy Combined

To examine accuracy of the total screening program (screening test plus colposcopy) we calculated sensitivity, specificity, PPV and NPV for the \geq CIN3 cases detected throughout the 7 years of the study, taking into account the possible compounded errors of the screening phase and the colposcopy phase.

Independently of the women age, Table 4 first shows that, assuming perfect colposcopic performance (i.e. that colposcopy would detect all the \geq CIN2 cases positive for the primary screening test), HPV screening is much more sensitive than conventional cytology for the detection or prediction of \geq CIN3 cases although it is less specific. This is concordant with the finding that the major effect of HPV on the screening program is to improve its negative predictive value, which is a function of sensitivity. The use of HPV testing for triage of ASCUS prior to colposcopic referral increased the specificity of conventional cytology, with slight loss in sensitivity. In contrast, the use of conventional cytology for triage of HPV positive women (considering as screening positive women who were both HPV positive and \geq ASCUS) increased specificity but at expense of a great decrease of the sensitivity originally gained with HPV testing. In general, we observed that the screening strategies were more accurate among women older than 30 years; for HPV screening the improvement in this age group was mainly increasing specificity (assuming perfect colposcopic performance specificity increased from 82.0% among women <30 years to 90.4% among women ≥ 30 years).

However, to evaluate properly the effect of the shift in the screening program including a downstream impact on colposcopy performance (Table 4, Section B), we estimated the same measures but considered as positive only women who screened positive at enrollment and had concomitant estimated colposcopic abnormalities, also at enrollment. We observed that regardless of a woman's age at baseline or the screening strategy (cytology, HPV, or combinations), the sensitivity and NPV of the screening program for prediction of the subsequent 7 years decreased because of enrollment colposcopic inaccuracy. This reduction in sensitivity was most pronounced for stand-alone HPV testing among older women. In short, as a result of colposcopic inaccuracy, the added sensitivity of HPV screening compared to conventional cytology did not translate into an absolute increase in sensitivity because of the inability to find many eventual lesions cases of \geq CIN3 among the HPV positive women who did not (yet) have acetowhite lesions in a one-time colposcopic evaluation.

DISCUSSION

Using data from the population based cohort study in Guanacaste, Costa Rica, we found that immediate colposcopic referral based on HPV positivity without further triage, even among women thirty and older, would be sensitive but non-specific. We evaluated the implications for colposcopy of a switch from conventional cytology-based screening program to primary HPV-based cervical screening in terms of number of referrals, percentage of women with acetowhite lesions and number of CIN2 and CIN3/cancer cases detected by abnormal colposcopic impression.

Our data indicate that, as expected, moving to HPV-based screening without further triage would substantially increase the number of colposcopy referrals, as compared to conventional cytology. At colposcopy, for each woman, the probability of having visible lesions would remain about the same. Comparing the PPV of abnormal colposcopic impressions and NPV of normal colposcopic impressions, the performance of colposcopy in a referral population derived from HPV screening would not be obviously better than its performance in a population referred by cytology with an ASCUS threshold.

The most important consideration is the impact of switching from cytology to HPV testing on the performance of the entire program (whose objective, in this analysis, was to detect and predict risk of CIN3 7 years out) encompassing the screening test and one single immediate colposcopy, and taking into account colposcopic performance. Our data indicated that an HPV-based screening program is more sensitive than conventional cytology mainly because of additional referral of women at risk of developing \geq CIN3. While HPV screening would refer more women at risk compared to cytology, it would also lead to referral of many other women with benign transient infections. Also our data indicate that insensitivity of immediate colposcopy blunted the superior sensitivity of HPV screening, since colposcopy especially failed at predicting which HPV infections would progress in the next 7 years.

Particularly among women 30 years or older, most of the \geq CIN2 cases detected or predicted by HPV testing that were diagnosed through the 7 years of the study did not have visible lesions at baseline. This result was strengthened when we used a stricter definition for colposcopic abnormality (women with clear visible lesions: cervigrams P1, P2, P3). This confirms that additional follow up or other forms of triage are needed for HPV positive women.

One limitation of our approach is that we only used the presence of acetowhite lesions as indicator of successful colposcopy. Recent data have demonstrated that even in abnormal appearing cervixes, it can be challenging to target the worst lesion with a biopsy. There is a clear need for a thoughtful protocol for colposcopic referral, and for a more thorough colposcopically directed biopsy protocol for screen positive women who are referred to colposcopy especially if only due to positive HPV testing. These results parallel the comparison of HPV testing vs. repeat cytology for triage of ASCUS in the ALTS trial.²¹ In that large trial, many of the CIN3 lesions diagnosed among HPV positive women within 2 years were not detectable by immediate colposcopy. As the first step to improve colposcopic sensitivity, taking more biopsies has been shown to increase detection of CIN3.²²

Because a visual technique cannot be expected to diagnose a molecular/microscopic process with sensitivity, it would be preferable in an ideal world to avoid colposcopy altogether if better risk stratifiers justifying immediate treatment were available. In the Guanacaste Natural History Study, we did not find a clinically appropriate combination of HPV testing and cytology. However, Naucler et al. reported that primary HPV testing screening followed with cytology triage and repeat HPV testing among HPV positive women with normal cytology may be a good strategy combined with cervical cancer screening.²³ Ideally, HPV

positive women should wait at least for 6–12 months to confirm at least short-term persistence to improve performance, because new infections at any age are typically benign and persistent infections are risky.¹² Perhaps HPV typing (for HPV 16 or 18), restriction of the test positivity threshold to exclude infections with ultra-low viral load (shown to pose a low PPV or absolute risk of subsequent \geq CIN3),^{3, 24} or triage with biomarkers (e.g. p16) might help to identify which women are at higher risk of progression. The best follow-up to an HPV-positive test will depend on large part on the resources and health infrastructure of the screening setting.

When interpreting our results, the fact that we calculated sensitivity using as outcome of disease all prevalent and incident cases detected through the 7 years of the study should be kept in mind, thus the lower sensitivity of HPV testing may be partly explained by infections acquired after baseline screening. Even shorter intervals would miss some interval CIN2 or CIN3, which is probably not harmful as long as cancer does not develop. We acknowledge that some might reasonably disagree with the view that colposcopy, as part of cervical screening, should be evaluated on its longer-term prognostic values.

As pertinent to all studies of subjective visual tests such as colposcopy, other colposcopists might have interpreted findings differently in individual cases. Still, we expect that the overall results and the conclusions would not change substantially. Although great variability exists in colposcopic impressions, this particular pair of cervicographer and colposcopist had worked together and had good agreement. We are confident that we could predict colposcopic impression from cervicography. Our colposcopist has been extensively compared, and found to be mainstream, in large collaborative studies of ASCCP experts.^{25, 26}

Also, unlike HPV test performance cytologic interpretations vary considerably between laboratories. While this is a clear advantage of HPV testing, it does not allow to generalize the quantitative differences between cytology and HPV screening found in our study, since the magnitude of change in moving from conventional cytology to HPV tests will depend on the accuracy of the existing screening program.

Despite the limitations, we believe the conclusions to be clear. Switching from cytology-based to HPV-based screening would not improve colposcopic performance. Colposcopy, in turn, can not cure the low specificity of HPV screening; visual methods like colposcopy, cervicography, and possibly even visual inspection with acetic acid, are not good triage tests for deciding which HPV-infected women should be followed up more intensively. Specific to each setting considering the introduction of HPV testing, clear guidelines for the clinical management of HPV positive women are needed to limit colposcopic referral and prevent (over) treatment; these guidelines must be in place before the widespread implementation of standalone, primary HPV-based screening.

Acknowledgments

The authors express special thanks to the women of Guanacaste who take part in this cohort study and the staff in Costa Rica for their dedicated work.

FUNDING: National Institutes of Health (N01-CP-21081, N01-CP-33061, N01-CP-40542, N01-CP-50535, N01-CP-81023, and intramural program CA78527 to R.D.B.). The Guanacaste cohort (design and conduct of the study, sample collection, management, analysis and interpretation of the data) for the enrollment and follow-up phases were supported by the National Cancer Institute, National Institutes of Health, Department of Health and Human Services. C.P. was supported by an appointment to the Senior Fellowship Program at the National Institutes of Health for analysis and manuscript preparation. The program is administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the US Department of Energy and the National Institutes of Health.

References

1. Ferreccio C, Bratti MC, Sherman ME, Herrero R, Wacholder S, Hildesheim A, Burk RD, Hutchinson M, Alfaro M, Greenberg MD, Morales J, Rodriguez AC, 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 Epidemiol Biomarkers Prev.* 2003; 12:815–23. [PubMed: 14504189]
2. Kjaer S, Hogdall E, Frederiksen K, Munk C, van den Brule A, Svare E, Meijer C, Lorincz A, Iftner T. The absolute risk of cervical abnormalities in high-risk human papillomavirus-positive, cytologically normal women over a 10-year period. *Cancer Res.* 2006; 66:10630–6. [PubMed: 17062559]
3. Mesher D, Szarewski A, Cadman L, Cubie H, Kitchener H, Luesley D, Menon U, Hulman G, Desai M, Ho L, Terry G, Williams A, et al. Long-term follow-up of cervical disease in women screened by cytology and HPV testing: results from the HART study. *Br J Cancer.* 2010; 102:1405–10. [PubMed: 20354519]
4. Sherman ME, Lorincz AT, Scott DR, Wacholder S, Castle PE, Glass AG, Mielzynska-Lohnas I, Rush BB, Schiffman M. Baseline cytology, human papillomavirus testing, and risk for cervical neoplasia: a 10-year cohort analysis. *J Natl Cancer Inst.* 2003; 95:46–52. [PubMed: 12509400]
5. Bulkmans NW, Berkhof J, Rozendaal L, van Kemenade FJ, Boeke AJ, Bulk S, Voorhorst FJ, Verheijen RH, van Groningen K, Boon ME, Ruitinga W, van Ballegooijen M, et al. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. *Lancet.* 2007; 370:1764–72. [PubMed: 17919718]
6. Kitchener HC, Almonte M, Thomson C, Wheeler P, Sargent A, Stoykova B, Gilham C, Baysson H, Roberts C, Dowie R, Desai M, Mather J, et al. HPV testing in combination with liquid-based cytology in primary cervical screening (ARTISTIC): a randomised controlled trial. *Lancet Oncol.* 2009; 10:672–82. [PubMed: 19540162]
7. Mayrand MH, Duarte-Franco E, Rodrigues I, Walter SD, Hanley J, Ferenczy A, Ratnam S, Coutlee F, Franco EL. Human papillomavirus DNA versus Papanicolaou screening tests for cervical cancer. *N Engl J Med.* 2007; 357:1579–88. [PubMed: 17942871]
8. Naucler P, Ryd W, Tornberg S, Strand A, Wadell G, Elfgrén K, Radberg T, Strander B, Johansson B, Forslund O, Hansson BG, Rylander E, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *N Engl J Med.* 2007; 357:1589–97. [PubMed: 17942872]
9. Ronco G, Giorgi-Rossi P, Carozzi F, Confortini M, Dalla Palma P, Del Mistro A, Gillio-Tos A, Minucci D, Naldoni C, Rizzolo R, Schincaglia P, Volante R, et al. Results at recruitment from a randomized controlled trial comparing human papillomavirus testing alone with conventional cytology as the primary cervical cancer screening test. *J Natl Cancer Inst.* 2008; 100:492–501. [PubMed: 18364502]
10. Sankaranarayanan R, Nene BM, Shastri SS, Jayant K, Muwonge R, Budukh AM, Hingmire S, Malvi SG, Thorat R, Kothari A, Chinoy R, Kelkar R, et al. HPV screening for cervical cancer in rural India. *N Engl J Med.* 2009; 360:1385–94. [PubMed: 19339719]
11. Wright TC Jr, Schiffman M, Solomon D, Cox JT, Garcia F, Goldie S, Hatch K, Noller KL, Roach N, Runowicz C, Saslow D. Interim guidance for the use of human papillomavirus DNA testing as an adjunct to cervical cytology for screening. *Obstet Gynecol.* 2004; 103:304–9. [PubMed: 14754700]
12. Rodriguez AC, Schiffman M, Herrero R, Hildesheim A, Bratti C, Sherman ME, Solomon D, Guillen D, Alfaro M, Morales J, Hutchinson M, Katki H, et al. Longitudinal study of human papillomavirus persistence and cervical intraepithelial neoplasia grade 2/3: critical role of duration of infection. *J Natl Cancer Inst.* 2010; 102:315–24. [PubMed: 20157096]
13. Jeronimo J, Schiffman M. Colposcopy at a crossroads. *Am J Obstet Gynecol.* 2006; 195:349–53. [PubMed: 16677597]
14. Dillner J, Rebolj M, Birembaut P, Petry KU, Szarewski A, Munk C, de Sanjose S, Naucler P, Lloveras B, Kjaer S, Cuzick J, van Ballegooijen M, et al. Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. *BMJ.* 2008; 337:a1754. [PubMed: 18852164]

15. Herrero R, Schiffman MH, Bratti C, Hildesheim A, Balmaceda I, Sherman ME, Greenberg M, Cardenas F, Gomez V, Helgesen K, Morales J, Hutchinson M, et al. Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: the Guanacaste Project. *Rev Panam Salud Publica*. 1997; 1:362–75. [PubMed: 9180057]
16. Bratti MC, Rodriguez AC, Schiffman M, Hildesheim A, Morales J, Alfaro M, Guillen D, Hutchinson M, Sherman ME, Eklund C, Schussler J, Buckland J, et al. Description of a seven-year prospective study of human papillomavirus infection and cervical neoplasia among 10000 women in Guanacaste, Costa Rica. *Rev Panam Salud Publica*. 2004; 15:75–89. [PubMed: 15030652]
17. Castle PE, Schiffman M, Gravitt PE, Kendall H, Fishman S, Dong H, Hildesheim A, Herrero R, Bratti MC, Sherman ME, Lorincz A, Schussler JE, et al. Comparisons of HPV DNA detection by MY09/11 PCR methods. *J Med Virol*. 2002; 68:417–23. [PubMed: 12226831]
18. Bouvard V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, Benbrahim-Tallaa L, Guha N, Freeman C, Galichet L, Cogliano V. A review of human carcinogens--Part B: biological agents. *Lancet Oncol*. 2009; 10:321–2. [PubMed: 19350698]
19. Massad LS, Jeronimo J, Katki HA, Schiffman M. The accuracy of colposcopic grading for detection of high-grade cervical intraepithelial neoplasia. *J Low Genit Tract Dis*. 2009; 13:137–44. [PubMed: 19550210]
20. Leisenring W, Alonzo T, Pepe MS. Comparisons of predictive values of binary medical diagnostic tests for paired designs. *Biometrics*. 2000; 56:345–51. [PubMed: 10877288]
21. Guido R, Schiffman M, Solomon D, Burke L. Postcolposcopy management strategies for women referred with low-grade squamous intraepithelial lesions or human papillomavirus DNA-positive atypical squamous cells of undetermined significance: a two-year prospective study. *Am J Obstet Gynecol*. 2003; 188:1401–5. [PubMed: 12824969]
22. Gage JC, Hanson VW, Abbey K, Dippery S, Gardner S, Kubota J, Schiffman M, Solomon D, Jeronimo J. Number of cervical biopsies and sensitivity of colposcopy. *Obstet Gynecol*. 2006; 108:264–72. [PubMed: 16880294]
23. Naucler P, Ryd W, Tornberg S, Strand A, Wadell G, Elfgrén K, Radberg T, Strander B, Forslund O, Hansson BG, Hagmar B, Johansson B, et al. Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. *J Natl Cancer Inst*. 2009; 101:88–99. [PubMed: 19141778]
24. Lorincz AT, Castle PE, Sherman ME, Scott DR, Glass AG, Wacholder S, Rush BB, Gravitt PE, Schussler JE, Schiffman M. Viral load of human papillomavirus and risk of CIN3 or cervical cancer. *Lancet*. 2002; 360:228–9. [PubMed: 12133661]
25. Jeronimo J, Massad LS, Castle PE, Wacholder S, Schiffman M. Interobserver agreement in the evaluation of digitized cervical images. *Obstet Gynecol*. 2007; 110:833–40. [PubMed: 17906017]
26. Massad LS, Jeronimo J, Schiffman M. Interobserver agreement in the assessment of components of colposcopic grading. *Obstet Gynecol*. 2008; 111:1279–84. [PubMed: 18515509]

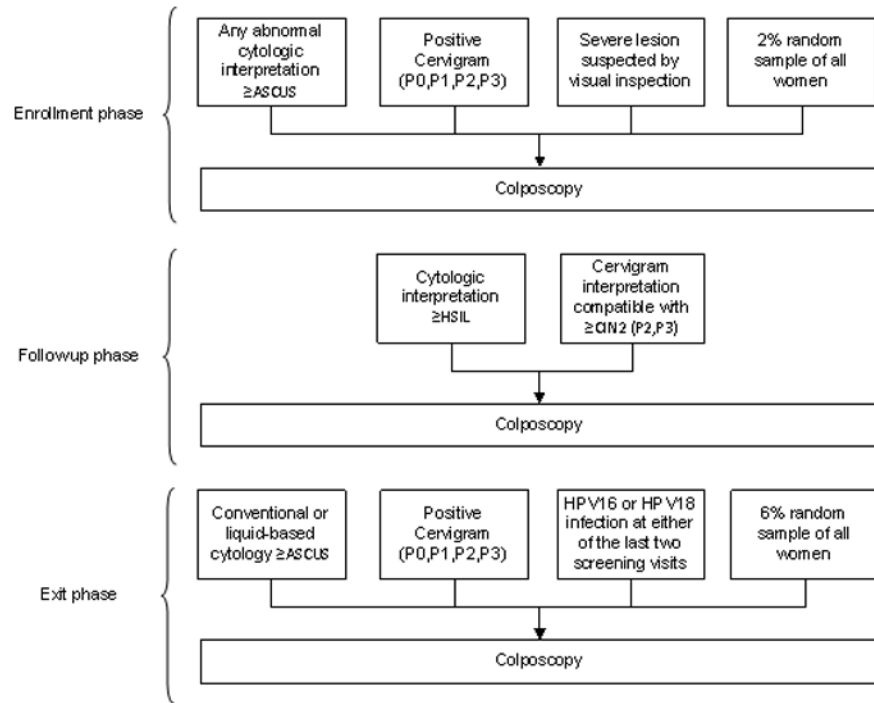


Figure 1. Algorithm for referral to colposcopy during the entire course of the cohort study.

Table 1

Baseline cervigram interpretations versus colposcopic impressions, among women referred to colposcopy at enrollment.

Colposcopic impression ²	Cervigram interpretation ¹		Total
	Normal	Abnormal	
Normal	1373	118	1491
Abnormal	94	438	532
Total	1467	556	2023

kappa=0.73 symmetry χ^2 p value=0.10

¹ Cervigram interpretations: Normal (Negative, P0), Abnormal (Atypical, P1, P2, P3).

² Colposcopic impressions: Normal and Abnormal (low grade, high grade and cancer)

Table 2

Conventional cytology and HPV testing results at baseline, by age.

		Women <30 years old			Women ≥ 30 years old		
		Any carcinogenic HPV type			Any carcinogenic HPV type		
Conventional cytology	Negative	Positive	TOTAL	Conventional cytology	Negative	Positive	TOTAL
Normal	1771	350	2121 92.3%	Normal	5167	496	5663 93.3%
≥ ASCUS	86	90	176 7.7%	≥ ASCUS	252	157	409 6.7%
	1857 80.8%	440 19.2%	2297		5419 89.2%	653 10.8%	6072

Table 3

Worst diagnosis through the 7 years of the study by estimated colposcopic impression at baseline, among women who tested positive for each of the screening methods at baseline.

		Women younger than 30 years				Women aged 30 years or older			
		Table 3a. Conventional Cytology ≥ ASCUS				Table 3c. Conventional Cytology ≥ ASCUS			
		Worst Diagnosis				Worst Diagnosis			
Estimated colposcopic impression		≥CIN3	CIN2	< CIN2	Total	≥CIN3	CIN2	< CIN2	Total
Abnormal [†]		14 21.2	7 10.6	45 68.2	66 100.0%	28 31.1	15 16.7	47 52.2	90 100.0%
Normal		6 5.5	10 9.1	94 85.5	110 100.0%	23 7.2	18 5.6	278 87.2	319 100.0%
Total		20 11.4	17 9.7	139 79.0	176 100.0%	51 12.5	33 8.1	325 79.5	409 100.0%
Table 3b. HPV testing positive									
Estimated colposcopic impression		Worst Diagnosis				Worst Diagnosis			
Estimated colposcopic impression		≥CIN3	CIN2	< CIN2	Total	≥CIN3	CIN2	< CIN2	Total
Abnormal [†]		24 15.7	13 8.5	116 75.8	153 100.0%	36 29.3	15 12.2	72 58.5	123 100.0%
Normal		12 4.2	12 4.2	263 91.6	287 100.0%	46 8.7	23 4.3	461 87.0	530 100.0%
Total		36 8.2	25 5.7	379 86.1	440 100.0%	82 12.6	38 5.8	533 81.6	653 100.0%

[†] Abnormal: equivalent to estimated colposcopic impressions of low grade, high grade or cancer.

Programatic accuracy of HPV testing and conventional cytology screening for the detection of \geq CIN3 cases through the 7 years of the study, stratified by age.

Table 4

Screening method	Section A. Assuming perfect colposcopic performance							
	Women < 30 years (n=2297)				Women \geq 30 years (n=6072)			
	Sens.	Spec.	PPV	NPV	Sens.	Spec.	PPV	NPV
HPV DNA positive ¹	70.6	82.0	8.2	99.19	80.4	90.4	12.6	99.63
Conventional cytology \geq ASCUS ²	39.2	93.1	11.4	98.54	50.0	94.0	12.5	99.10
Conventional \geq ASCUS (ASCUS with HPV triage) ³	37.3	94.7	13.8	98.52	50.0	95.6	16.1	99.11
HPV positive (followed by conventional cytology triage) ⁴	33.3	96.8	18.9	98.46	44.1	98.1	28.7	99.04

Screening method	Section B. Taking into account colposcopic performance when visible lesions are present							
	Women < 30 years (n=2297)				Women \geq 30 years (n=6072)			
	Sens.	Spec.	PPV	NPV	Sens.	Spec.	PPV	NPV
HPV DNA positive and abnormal colposcopy ⁵	47.1	94.3	15.7	98.74	35.3	98.5	29.3	98.89
Conventional cytology \geq ASCUS and abnormal colposcopy ⁶	27.5	97.7	21.2	98.34	27.5	99.0	31.1	98.76
Conventional \geq ASCUS (ASCUS followed with HPV testing) and abnormal colposcopy ⁷	27.5	98.1	25.0	98.35	27.5	99.1	34.2	98.76
HPV positive (followed by conventional cytology triage) and abnormal colposcopy ⁸	23.5	98.6	27.3	98.27	25.5	99.5	44.8	98.74

Sens: Sensitivity; Spec: Specificity; PPV: Positive predictive value; NPV: Negative predictive value.

- ¹ <30 years, screened positive n=440, negative n=1857; \geq 30 years screened positive n=653, negative n=5419
- ² <30 years, screened positive n=176, negative n=2121; \geq 30 years screened positive n=409, negative n=5663
- ³ <30 years, screened positive n=138, negative n=2159; \geq 30 years screened positive n=316, negative n=5756
- ⁴ <30 years, screened positive n=90, negative n=2207; \geq 30 years screened positive n=157, negative n=5915. Conventional cytology for triage of HPV positive women means considering as screening positive just those women who were both HPV positive and cytologically \geq ASCUS.
- ⁵ <30 years, screened positive n=153, negative n=2144; \geq 30 years screened positive n=123, negative n=5949
- ⁶ <30 years, screened positive n=66, negative n=2231; \geq 30 years screened positive n=90, negative n=5982
- ⁷ <30 years, screened positive n=56, negative n=2241; \geq 30 years screened positive n=82, negative n=5990
- ⁸ <30 years, screened positive n=44, negative n=2253; \geq 30 years screened positive n=58, negative n=6014