A Comparison of Single and Combined Visual, Cytologic, and Virologic Tests as Screening Strategies in a Region at High Risk of Cervical Cancer[1]

Catterina Ferreccio, Maria C. Bratti, Mark E. Sherman, Rolando Herrero, Sholom Wacholder, Allan Hildesheim, Robert D. Burk, Martha Hutchinson, Mario Alfaro, Mitchell D. Greenberg, Jorge Morales, Ana C. Rodriguez, John Schussler, Claire Eklund, Guillermo Marshall, and Mark Schiffman[2]

Departamento de Salud Publica, Universidad Catolica de Chile, Santiago, Chile [C. F., G. M.]; Proyecto Epidemiologico Guanacaste, San Jose, Costa Rica [M. C. B., R. H., M. A., J. M., A. C. R.]; Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland 20852 [M. E. S., S. W., A. H., M. S.]; Albert Einstein College of Medicine, Bronx, New York 10461 [R. D. B.]; Womens and Infants' Hospital, Providence, Rhode Island 02905 [M. H., C. E.]; US Health Connect, Blue Bell, Pennsylvania 19034 [M. D.]; and Information Management Services, Silver Spring, Maryland 20904 [J. S.]

Abstract

Increased understanding of human papillomavirus (HPV) infection as the central cause of cervical cancer has permitted the development of improved screening techniques. To evaluate their usefulness, we evaluated the performance of multiple screening methods concurrently in a large population-based cohort of >8500 nonvirginal women without hysterectomies, whom we followed prospectively in a high-risk region of Latin America. Using Youden's index as a measure of the trade-off between sensitivity and specificity, we estimated the performances of a visual screening method (cervicography), conventional cytology, liquid-based cytology (ThinPrep), and DNA testing for 13 oncogenic HPV types. The reference standard of disease was neoplasia ≥ cervical intraepithelial neoplasia grade 3 (CIN 3), defined as histologically confirmed CIN 3 detected within 2 years of enrollment ($n = 90$) or invasive cancer detected within 7 years ($n = 20$). We analyzed each technique alone and in paired combinations ($n = 112$ possible strategies), and evaluated the significance of differences between strategies using a paired Z test that equally weighted sensitivity and specificity. As a single test, either liquid-based cytology or HPV DNA testing was significantly more accurate than conventional cytology or cervicography. Paired tests incorporating either liquid-based cytology or HPV DNA testing were not substantially more accurate than either of those two test strategies alone. However, a possibly useful synergy was observed between the conventional smear and cervicography. Consideration of age or behavioral risk profiles did not alter any of these conclusions. Overall, we conclude that highly accurate screening for cervical cancer and CIN 3 is now technically feasible. The remaining vital issue is to extend improved cervical cancer prevention programs to resource-poor regions.

Introduction

Increased understanding of the pathogenesis of cervical cancer should permit improved prevention methods. We now understand that ~15 oncogenic types of HPV[3] infection cause virtually all cases of cervical cancer and its immediate precursor, CIN 3 (1). Vaccination might ultimately prevent or treat oncogenic HPV infections and/or the lesions they produce but, in the shorter term, screening will be required for prevention. The challenge is to develop screening strategies, consistent with our understanding of the natural history of HPV infection and cervical cancer, which balance the need for sensitive detection of CIN 3 and cancer with acceptable specificity. Specificity is an issue, because infections with the oncogenic types of HPV are very common (2). Despite their oncogenic potential, most infections typically resolve within 1–2 years, and only a minority progress to CIN 3, which poses a high risk of invasion if left untreated (3, 4).

Screening tests can be categorized as visual (*e.g.*, colposcopy and its proxies), microscopic (cytology), or molecular. Visual screening methods rely on the identification of patterns of abnormal blood vessels or tissue whitening on application of acetic acid (5). The distinction between acute HPV infection and precancer (CIN 3) can be estimated roughly using grading of the same and additional visual criteria.

Cytology, specifically the conventional Pap smear, has been the mainstay of cervical cancer prevention for >50 years. The spectrum of cervical cytologic abnormalities ranges from equivocal changes to the pathognomonic nuclear and cytoplasmic effects of HPV infection ("koilocytosis") to severe cytologic neoplastic changes that suggest progression to CIN 3 (6, 7). New cytologic techniques might offer increased accuracy at increased cost, particularly liquid-based cytology produced in

Received 11/25/02; revised 6/2/03; accepted 6/5/03.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
[1] Supported by a series of National Cancer Institute contracts. In addition, this work was supported in part by Grant CA78527 (to R. D. B.) from the NIH.
[2] To whom requests for reprints should be addressed, at Division of Cancer Epidemiology, National Cancer Institute, NIH, Department of Health and Human Services, Room 7066, 6120 Executive Boulevard, Rockville, MD, 20852. Phone: (301) 496-1691; Fax: (301) 402-0916; E-mail: schiffmm@mail.nih.gov.

[3] The abbreviations used are: HPV, human papillomavirus; CIN, cervical intraepithelial neoplasia; Pap, Papanicolaou; ASCUS, atypical squamous cells of undetermined significance; LSIL, low-grade squamous intraepithelial lesion; HSIL, high-grade squamous intraepithelial lesion; HC2, Hybrid Capture 2; TaqGold, AmpliTaq Gold polymerase; CI, confidence interval.

standard, automated fashion relying on cells collected into preservative buffers (8, 9).

Molecularly, HPV DNA of known oncogenic types can now be directly measured, as can viral RNA, proteins, and related antibodies produced against viral antigens by some exposed individuals. At present, HPV DNA detection is the most accurate molecular technique for the detection of current HPV infection. The combination of HPV DNA testing and cytology for cervical cancer screening, particularly among women ≥30 years of age, is under evaluation by several groups including the American Cancer Society (10) and has been approved by the Food and Drug Administration. Screening using a combination of methods, although more expensive per screening, might be cost effective if the increased sensitivity permits lengthening of the screening interval.

In short, there is an increasingly large research literature on possible applications of new visual, microscopic, and virologic screening methods for prevention of cervical cancer. However, few if any investigators have examined the relative performance of new techniques, and combinations of techniques, in systematic surveys of whole populations where a broad and representative spectrum of cervical lesions is expected.

The present article summarizes the project experience related to performances of a visual screening method, conventional cytology, liquid-based cytology, and DNA testing for 13 oncogenic HPV types in a large cohort study in Guanacaste, Costa Rica.

Materials and Methods

Population. The enrollment (11, 12) and follow-up (13) phases of the Guanacaste Project have been described in detail elsewhere. Briefly, approximately one-sixth of the Guanacaste census tracts were selected randomly, and all of the women ≥18 years of age in the tracts were enumerated formally. A total of 10,049 women enrolled (>90% participation). This comparison of four screening technologies was nested in the natural history of HPV cohort, with a sample size and statistical power dictated by the occurrence of disease in the cohort. For this analysis, we excluded 583 virgins, 624 hysterectomized women, and 291 women without screening tests, leaving 8,551 women in the analytic cohort

Enrollment Screening Visit. Three highly trained study nurses followed a standardized protocol including three cytologic techniques and two visual techniques. Specifically, the nurses first visually examined the cervix without magnification or acetic acid. Conventional Pap smears were obtained by firmly rotating a Cervex brush (Unimar, Wilton, CT) five times (1800°) clockwise in the ecto- and endocervical area. The conventional smear was prepared and immediately spray-fixed with ethanol and carbowax Pap Perfect fixative (Medscand USA, Hollywood, FL). After the smear was made, the Cervex brush was swirled and pressed into 20 ml of methanol-based PreservCyt solution (Cytyc Corporation, Boxborough, MA) and then discarded. The solution was kept at ambient temperature for the preparation of liquid-based cervical slides (ThinPreps; Cytyc Corporation) in the United States (14).

For HPV DNA testing, additional cervical cells were collected with a Dacron swab, which was rotated in the endocervical canal and then swabbed over the ectocervix. The swabs were stored in Virapap DNA transport medium tubes (Digene Corporation, Gaithersburg, MD), which were kept in coolers at 4°C in the field until frozen at −30°C at the regional study offices.

After collection of the materials listed above, the cervix was rinsed twice with 5% acetic acid, and two photographic images of the cervix (cervigrams) were taken (15). The exposed, undeveloped film was sent to National Testing Laboratories, Worldwide (Fenton, MO) for developing and evaluation.

Interpretation of Enrollment Screening Tests and Colposcopic Referral. The conventional smears were stained and initially interpreted in Costa Rica by a team of cytotechnologists and one experienced pathologist. The staining process was reviewed and improved in mid-enrollment because of early variability. We had conventional cytologic interpretations for 8481 of the 8551 women. After the Costa Rican conventional interpretation, a cytotechnologist/pathologist team reinterpreted the smears using PapNet, a computer-assisted technology that added to the sensitivity of the overall enrollment screening but is no longer available (and thus not evaluated here; Ref. 16).

The PreservCyt vials were sent to the United States where ThinPreps were prepared using a prototype of the ThinPrep 2000 Processor, and a third cytology interpretation was made. During the early part of enrollment, the preparation of many ThinPreps was repeated to optimize the automated slide preparation technique, which was not yet standardized at that time. A few international shipments of PreservCyt vials were lost and, thus, we had results for 8082 of the 8551 women.

All three of the cytologic methods used the 1991 Bethesda System (17) for reporting, which included the following: normal including reactive changes, ASCUS, LSIL, HSIL, or carcinoma.

With regard to visual screening, during the enrollment examination, the nurses referred women to colposcopy if the direct visual appearance without acetic acid or magnification suggested malignancy or need for immediate gynecologic action (*e.g.*, mass or ulceration). The cervigram was interpreted in the United States as negative (no lesion seen); atypical (a lesion was seen but the site and/or morphology of the lesion was such that colposcopy was not recommended); or positive. The positive category had the following classifications: P0, probably normal variant but colposcopy preferable to rule out serious neoplasia; P1, compatible with CIN 1; P2, compatible with CIN 2-CIN3; and P3, compatible with cancer (15). We obtained cervigram results for 8457 of the 8551 women.

All of the patients with abnormal results in any of the three Pap smears (ASCUS or worse), positive (P) cervigrams, or worrisome direct visual appearance were referred for colposcopy and biopsy of visible lesions. A group of controls with negative screening examinations was referred as well. Of note, cervigram photographs marked as to possibly significant areas were available to the colposcopist, who also was aware of referring high-grade cytologic (but not virologic) screening data.

HPV DNA Testing. Many of the swab-based specimens collected for virologic studies were tested repeatedly for HPV DNA before the final testing that forms the basis of this report (18, 19). Initially, all of the specimens were tested by Hybrid Capture Tube test, which was used to define which women should be followed most actively in the prospective cohort (see the following section) but is now an obsolete test because of inadequate sensitivity. Specimens from women with prevalent CIN (and controls) were also tested by MY09-MY11 consensus primer PCR using AmpliTaq polymerase and by HC2. However, the basis of this report is the masked retesting of remaining aliquots of all of the specimens in the full cohort, again using MY09-MY11 consensus primer PCR using TaqGold, a more sensitive assay for HPV DNA (19). For 33 women,

completely depleted specimens led us to use the Amplitaq PCR results instead of missing TaqGold results. We exclude the women with no PCR results ($n = 32$) from HPV DNA test-related analyses.

Specimen preparation for PCR, amplification by MY09/11, detection of PCR products by gel electrophoresis, Southern blot transfer, and hybridization with radiolabeled probes for HPV types 11, 16, 18, and 51 (generic probe) have been described elsewhere (19). Two observers evaluated the signal strength of the PCR products. All of the PCR products were hybridized with type-specific probes for >40 HPV types. Three experienced investigators interpreted each dot-blot result, and discrepancies were resolved by consensus. For this analysis, to maximize the clinical relevance of the results, a specimen was considered HPV DNA positive if it was found to contain at least 1 of the 13 oncogenic types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, or 68) contained in HC2, the only HPV DNA test approved by the United States Food and Drug Administration. In total, we had HPV results for 8519 of 8551 women.

Follow-Up. All of the women with CIN 2 or worse at enrollment screening and resultant colposcopic evaluation were treated by loop electrosurgical excision procedure, cold knife cone, or hysterectomy/radiation as appropriate, and censored from additional follow-up. We rescreened at 6–12 month intervals all of the women with any equivocal or definite cytologic abnormality by any technique; P cervigrams; HPV DNA positivity by Hybrid Capture Tube test; or a lifetime history of >4 sexual partners. We also actively followed via 12 monthly rescreening a random sample of 416 additional women as controls. The total size of the actively followed subcohort of sexually active women was 2627 (30.7%) of 8551 women in the total cohort. We rescreened the remainder of the cohort not followed actively at least once between years 5 and 7 after enrollment. Participation rates in all of the components of the Guanacaste Project were very high (consistently >80% of scheduled visits were kept). For all of the groups, the repeat screening examination was the same as the enrollment examination except for the deletion after year 1 of follow-up of PapNet reinterpretation of the conventional cytology smears.

During follow-up, women were again referred to the same expert colposcopist if cytologic HSIL/cancer (conventional or ThinPrep) or P2/P3 cervigrams were observed. However, because this cohort study was designed to examine the etiologic role of HPV infection, HPV test results were not used (at enrollment or during follow-up) as a basis of referral to colposcopy, colposcopic biopsy, or treatment until after the exit phase of the project.

During the exit phase of the cohort study, we used more liberal colposcopic referral algorithms in a final effort to maximize participant safety. Specifically, in addition to those with any screening results suggestive of CIN 2 or worse, women with cytologic findings of ASCUS or LSIL, P0/P1 cervigrams or persistent HPV infection (*i.e.*, the same DNA type at enrollment and a second test at years 5–7) were referred to exit colposcopy as an additional safety check against CIN 3 or cervical cancer.

Reference Standard of Disease. Colposcopic examinations were performed by a single expert gynecologic oncologist, who biopsied the worst appearing part of the cervix. We included, as the reference standard of disease, all of the cases of CIN 3 found at enrollment ($n = 73$) and during the first 2 years of follow-up ($n = 17$), and all of the cases of cancer detected during enrollment ($n = 12$), and during follow-up and exit ($n = 8$). Therefore, a total of 90 CIN 3 (73 + 17) and 20 cancer cases (12 + 8) are included in the analysis discussed below.

Statistical Methods. We evaluated the test performance for each screening technique and combination of two techniques. When we evaluated a screening method that combined two tests, women with either missing were dropped from the analysis. For each of the two cytologic methods, performance was considered at colposcopic referral thresholds of ≥ASCUS, ≥LSIL, or ≥HSIL. For cervicography, referral thresholds were ≥A, ≥P0, ≥P1, ≥P2, and ≥P3. HPV tests were categorized simply as positive or negative for at least 1 of the 13 oncogenic types in HC2, the only widely available HPV DNA test. Inclusion of additional, possibly oncogenic types including HPV 26, 66, and 73 did not meaningfully affect the conclusions (HPV testing was made slightly more sensitive but less specific as expected). We investigated combinations that required both tests to be positive, and those requiring either of a pair to be positive for referral. When all of the single and double combinations, using all of the thresholds, were combined, there were 112 possible testing strategies.

For all 112 of the strategies, we calculated the sensitivity and specificity of detection for CIN 3 and cancer together, and then repeated the analysis for cancer alone. To take into account tradeoffs between increasing sensitivity and decreasing specificity (20), we compared the overall accuracy of the strategies using Youden's index (21). We chose the index because of its simplicity, clear meaning, and easily calculable CIs. The score is calculated as the sensitivity plus specificity (expressed as proportions) minus 1.0. More formally, if true positives are a, false negatives b, false positives c, and true negatives d, the formula for the index is $(ad - bc)/[(a + b)(c + d)]$ with a well-defined variance of $[ab/(a + b)^3] + [cd/(c + d)^3]$ that can be used to calculate CIs. The values range theoretically from 1.0 (perfect) to 0.0 (randomly useless) to −1.0 (always wrong).

We present the single test strategies and the six optimal paired combinations (the optimal screening threshold for each paired combination of the four techniques as determined by Youden's index) in the tables. We evaluated the statistical significance of differences in accuracy (*i.e.*, test positivity for cases and test negativity for noncases) between these 10 strategies using a newly developed Z test[5] that accounts for the fact that the screening accuracy is evaluated in the same individuals and gives equal weight to sensitivity and specificity, just as does Youden's index. The Z statistic is computed according the following formula:

$$Z = \frac{\frac{n_{+-} - n_{-+}}{n} - \frac{m_{+-} - m_{-+}}{m}}{\sqrt{\frac{n_{-+} + n_{+-} - [n_{+-} - n_{-+}]^2/n}{n^2} + \frac{m_{-+} + m_{+-} - [m_{+-} - m_{-+}]^2/m}{m^2}}}$$

where n equals the number of cases, m the number of noncases, n_{+-} the number of cases classified properly as positive only by test strategy 1, n_{-+} the number of cases classified as positive by test strategy 2 only, m_{-+} the number of noncases classified properly as negative by test strategy 1 only, and m_{+-} the number of noncases classified properly as negative by test strategy 2 only. Only discordant test results figure into this

[4] G. Marshall, manuscript in preparation.

Table 1 Characteristics of the 8551 women in the Guanacaste study population, by case status

	Percentage of noncases	Percentage of women with ≥CIN 3
Age	(n = 8441)	(n = 110)
18–19	1.9	0.9
20–29	25.2	28.2
30–39	29.2	38.2
40–49	18.8	13.6
50–59	11.1	7.3
60–69	8.0	7.3
≥70	5.8	4.6
Years of education	(n = 8433)	(n = 110)
0	8.3	10.0
1–3	18.7	13.6
4–6	39.6	47.3
7–9	11.7	10.9
≥ 10	21.6	18.2
Number of live births[a]	(n = 8441)	(n = 110)
0	5.8	0.9
1–2	33.1	26.4
3–4	28.6	31.8
5–6	12.7	24.6
7–10	13.3	10.9
≥11	6.5	5.4
Number of sex partners[a]	(n = 8440)	(n = 110)
1	54.3	36.4
2	21.5	22.7
3	12.7	17.3
≥4	11.6	23.6
Age at first intercourse[a]	(n = 8441)	(n = 110)
≤14	8.8	12.7
15–16	23.6	37.3
17–18	28.6	23.6
19–20	17.5	12.7
≥21	21.5	13.6
Smoking[a]	(n = 8436)	(n = 110)
Never	89.1	80.9
Past	5.6	8.2
Current	5.2	10.9
Oral contraceptive use[a]	(n = 8435)	(n = 110)
Never	36.7	27.3
Past	43.0	38.2
Current	20.4	34.6

[a] Difference between cases and noncases statistically significant, $P \leq 0.01$.

statistic, and women with missing values for either test are excluded.

Estimates of Youden's index for all of the possible screening thresholds are shown in the Appendix, but without pairwise comparisons for statistical significance.

We also calculated and provide in the Appendix the predictive values, which will vary markedly according to the disease prevalence in the screened population. We stratified on age (<30, 30+) and also on risk. We defined as high-risk women those who ever smoked cigarettes, who had ≥4 live births, or who had ≥5 lifetime sexual partners.

Results

In Table 1, the most relevant characteristics of the study population are tabulated, stratified by case status. The great majority of adult women in Guanacaste are literate, although most complete few years of schooling. The population is highly parous (mean = 4.1), and women typically marry young, having few sexual partners (mean = 2.1). According to previous research, the male sexual partners are likely to have additional sexual contacts (22). Smoking is uncommon. Oral contraceptive use is common, but not long-term in accordance with high parity. Cases tended to have more sexual partners, earlier ages at first intercourse, more smoking, higher parity, and more oral contraceptive use than noncases (23). The 90 cases of CIN 3 had a mean age at enrollment of 36.9 (median, 33.5; range, 19–85). The 20 women with cancer (n.b., detected by screening not symptoms) had a mean age at enrollment of 43.4 (median, 41.5; range, 20–72).

The single-technique strategies aimed at detecting ≥CIN 3 and the most promising ways to combine each pair of techniques are listed in Table 2. As single techniques, liquid-based cytology (≥ASCUS) and HPV DNA testing performed almost identically. Either was significantly more accurate than conventional cytology (≥ASCUS) or cervicography (≥A), because of greatly increased sensitivity with less extreme decreases in specificity ($P < 0.001$). Conventional cytology was nonsignificantly more accurate than cervicography ($P = 0.11$) because of greater specificity given similar sensitivity.

Table 2 also shows the most promising test thresholds for the six possible pairings of the four screening techniques. When combining two tests, the two alternatives are to use both tests independently with a positive on either counting as a positive screening result; or to use them sequentially with the second test applied only to those women for whom the first test is positive. Strategies with sequential testing were less accurate than single tests, because somewhat increased specificity did not compensate for decreased sensitivity. Therefore, except for the Appendix, we only present here the results for tests combined in strategies in which either test being positive equaled overall positivity.

Adding any of the other three techniques to cervicography or to the conventional Pap smear resulted in combinations that were much more accurate than the single techniques alone ($P < 0.001$). Adding any of the other three techniques to HPV DNA testing increased accuracy but to a lesser extent ($P = 0.03–0.04$ for three paired comparisons), with tradeoffs of slightly improved sensitivity and decreased specificity. Adding another technique to liquid-based cytology did not significantly improve accuracy. In fact, combining the conventional Pap and liquid-based cytology was less accurate than liquid-based cytology alone ($P < 0.001$) because of lower specificity.

When we compared the most accurate combinations of the techniques with each other, the first four combinations listed in Table 2 were equivalent. The combination of liquid-based cytology and conventional Pap smear was nonsignificantly less accurate than any of the first four combinations listed in Table 2 ($P = 0.13–0.15$) but, in turn, was nonsignificantly more accurate than a combination of cervicography and the conventional Pap smear (which was significantly less accurate than the first four; $P = 0.005–0.006$).

A few other combinations should be mentioned because of extremely high sensitivity. Referring women with either a cervigram ≥ P0 or HPV positivity yielded a sensitivity of 93.5%, but was less specific than the most accurate combination of cervicography (≥P2) and HPV testing, because decreased specificity (84.2%) reduced the Youden's index (0.78; 95% CI, 0.73–0.82). The combination of liquid-based cytology (≥ASCUS) and HPV testing was 94.3% sensitive but only 80.2% specific, with a Youden's index of 0.74 (95% CI, 0.70–0.79).

When we repeated the analyses in Table 2, restricting the disease end point to the 20 invasive cancers, the results were similar although less reliable because of small numbers. As single techniques, liquid-based cytology or HPV testing were

Table 2 Single and selected[a] two-test strategies for detection of ≥CIN 3

Strategy[b]	Sensitivity[c]	Specificity[d]	Youden's index[e]	95% CI
HPV (+)	85.3%	88.2%	0.74	0.67–0.80
Liquid-based (≥ASCUS)	85.7%	87.8%	0.74	0.67–0.80
Smear (≥ASCUS)	63.0%	93.7%	0.57	0.48–0.66
Cervigram (≥A)	61.7%	84.8%	0.46	0.37–0.56
Smear (≥HSIL) or HPV (+)	90.7%	87.8%	0.79	0.73–0.84
Liquid-based (≥HSIL) or HPV(+)	90.5%	88.0%	0.78	0.73–0.84
Cervigram (≥P2) or HPV (+)	89.7%	88.1%	0.78	0.72–0.84
Liquid-based (≥ASCUS) or cervigram (≥P0)	93.2%	83.9%	0.77	0.72–0.82
Smear (≥HSIL) or liquid based (≥ASCUS)	86.5%	87.6%	0.74	0.68–0.81
Smear (≥LSIL) or cervigram (≥P0)	74.5%	90.9%	0.65	0.57–0.74

[a] For each of the six possible two-technique combinations, the table shows the performance for the cut-points with the highest accuracy as measured by Youden's index.
[b] There were three possible thresholds for conventional and liquid-based cytology (≥ASCUS, ≥LSIL, ≥HSIL), five possible thresholds for cervicography [≥Atypical, ≥Positive(0), ≥Positive(1), ≥Positive(2), ≥Positive(3)], and a single threshold for HPV DNA testing (positive *versus* negative). Techniques were considered singly and in pairs at all thresholds. Two kinds of combinations were evaluated, either requiring both techniques to be positive or at least one. Overall, there were 112 strategies considered, which were ranked in order of decreasing Youden's index.
[c] Sensitivity calculated as the percentage of cases of ≥CIN 3 detected by the screening strategy.
[d] Specificity calculated as the percentage of women without CIN 3 or cancer who tested negative by the screening strategy.
[e] Youden's index calculated as sensitivity plus specificity (expressed as proportions) minus 1.00, with 95% CI.

again significantly more accurate than conventional cytology or cervicography. The performance of liquid-based cytology (sensitivity of 83%, specificity of 87%, Youden's index of 0.70) was not substantially improved by the addition of any of the other techniques, although a combined screening with both liquid-based cytology and HPV testing did detect 89% of 18 women with both test results who were diagnosed with cancer in the 7-year study. The results indicate that repeated testing of specimens from the few cancer cases might have compromised the sensitivity of HPV testing. Seventeen of 20 cancer cases (85%) were HPV positive on at least one assay (Hybrid Capture tube test, HC2, AmpliTaq PCR, or TaqGold PCR). But only 14 of 19 (73.7%) were positive on the final round of TaqGold retesting of barely remaining aliquots used for these analyses, despite the relatively high analytic sensitivity of the TaqGold technique shown in direct comparisons to the other assays (19).

The positive predictive values of the most accurate single or two-technique strategies (Table 3) demonstrate the difference between clinical and public health needs. To detect ≥CIN 3, the most accurate tests as judged by high Youden's indexes had a positive predictive value <10% (and <2% for invasive cancers). On the other hand, the negative predictive values of these same strategies (reassurance regarding the absence of prevalent or incipient ≥CIN 3) were very high, uniformly around 99.9%. In other words, women testing negative by these strategies would have a probability of only 1/1000 of presenting with CIN 3 within 2 years, and much less for cancer within 7.

We attempted to discern subpopulations that might benefit from specific screening strategies, by dividing the population by age (<30, ≥30) and risk profile. To begin screening at older ages (even ≥25) limited the sensitivity of all screening strategies, because so many cases of CIN 3 in particular were young. Although some differences by age/risk group were observed, they were generally not large or consistent enough to affect the main conclusions. As an exception worth mentioning, cervicography performance (≥A) decreased substantially with age (Youden's index of 0.56, 0.44, and 0.35 for ages <30, 30–44, and ≥45, respectively), with limited accuracy especially in older, multiparous women because of low sensitivity presumably because the transformation zone was not adequately visible.

Table 3 Positive and negative predictive values for single and selected[a] two-test strategies for the detection of ≥CIN 3

Strategy[b]	Positive predictive value[c]	Negative predictive value[d]
Liquid-based (≥ASCUS)	8.5%	99.8%
HPV (+)	8.6%	99.8%
Smear (≥ASCUS)	11.5%	99.5%
Cervigram (≥A)	4.9%	99.4%
Smear (≥HSIL) or HPV (+)	8.8%	99.9%
Liquid-based (≥HSIL) or HPV (+)	9.0%	99.9%
Cervigram (≥P2) or HPV (+)	8.8%	99.9%
Liquid-based (≥ASCUS) or cervigram (≥P0)	7.0%	99.9%
Smear (≥HSIL) or liquid-based (≥ASCUS)	8.4%	99.8%
Smear (≥LSIL) or cervigram (≥P0)	9.5%	99.6%

[a] For each of the six possible two-technique combinations, the table shows the predictive values for the cut-points with the highest accuracy as measured by Youden's index.
[b] There were three possible thresholds for conventional and liquid-based cytology (≥ASCUS, ≥LSIL, ≥HSIL), five possible thresholds for cervicography [≥Atypical, ≥Positive(0), ≥Positive(1), ≥Positive(2), ≥Positive(3)], and a single threshold for HPV DNA testing (positive *versus* negative). Techniques were considered singly and in pairs at all thresholds. Two kinds of combinations were evaluated, either requiring both techniques to be positive or at least one. Overall, there were 112 strategies considered, which were ranked in order of decreasing Youden's index.
[c] Positive predictive value was calculated as the percentage of women with a positive screening result that had CIN 3 or cancer diagnosed.
[d] Negative predictive value was calculated as the percentage of women with a negative screening result that did not have CIN 3 or cancer diagnosed.

Discussion

Primary Results. There are now several good techniques to screen for CIN 3 and cervical cancer. All four of the techniques that we evaluated demonstrated worth; nonetheless, there were clearly two levels of performance: liquid-based cytology or HPV DNA testing were more accurate than conventional cytology or cervicography. Two-technique combinations including liquid-based cytology or HPV DNA testing were not substantially more accurate than either of the two new techniques alone. In particular, adding another technique to liquid-based cytology as interpreted in this project did not additionally improve screening performance. Sequential testing strategies

were relatively insensitive, although their specificity could be important in low-resource settings (see below).

Both liquid-based cytology and HPV DNA testing were very sensitive in detecting ≥CIN 3, but were less specific than conventional Pap smear or cervicography. Combining them, there was a marginally significant gain in accuracy compared with use of HPV testing alone, and a nonsignificant improvement compared with liquid-based cytology alone. The performance of HPV DNA testing was fairly typical compared with other published studies, although possibly compromised in a few cases because of specimen depletion because of prior rounds of testing. However, the performance of liquid-based cytologic interpretations in this study was unusual, with a very high percentage of ASCUS interpretations that raised sensitivity and lowered specificity to a level similar to the HPV DNA test. If liquid-based cytology had performed differently, the combination of liquid-based cytology and HPV testing might have been different.

The conventional Pap smear performed well in our study, with sensitivity in the upper range of its published historical performance (8, 24). This may reflect efforts to improve smear quality initiated during the project, and the expertise of the single pathologist who interpreted all of the smears. Nonetheless, adding any of the other three tests increased the accuracy of Pap smear alone. The highest accuracy was obtained when HPV testing was added, a performance equivalent to that of HPV combined with liquid-based cytology.

Cervicography was the least accurate single test. Some authors have described higher sensitivity for cervicography than conventional Pap smear (25, 26), but we did not, partly because of notably decreased sensitivity in older women (27). Cervicography performed better in combinations with other tests than alone, in line with the stated expectations of National Testing Laboratories Worldwide, which performed cervicography evaluations for this study. The improvement in the performance of Pap smears when cervigrams were added suggests a possibly exploitable synergy between cytological and visual screening (28). Cervicography or another similar, visual technique could be relatively inexpensive and easy to introduce. The images could be quickly interpreted in large numbers if the interpretations were centralized in a few expert regional centers. Theoretically the results obtained with cervigrams might be reproduced by other visual techniques (colposcopy, direct visual inspection) if performed in similar conditions.

Strengths of the Study. Our conclusions were robust, in that they were unaltered by changing the analytic strategy. Liquid-based cytology and HPV DNA testing were the most accurate single tests, and the same combinations tended to be noteworthy as judged by Youden's index or a variety of other measures that we explored but did not report [including comparing CIs for sensitivity, specificity, or predictive values (29); adjusting the observed two-technique results for random effects (20); computing effectiveness scores (30, 31); and receiver operating characteristic (ROC) curve-based approaches (32)]. As additional strengths of the study, it was truly population-based and participation rates were high enough to rule out meaningful biases. Finally, we were able to optimize each of the screening techniques used in the study by extensive training, quality monitoring, and retraining as needed.

Limitations. All of the studies of screening efficacy must be concerned with verification bias. On the basis of multiple screening tests we referred a large percentage, ~25%, of women in the cohort to colposcopic evaluation at enrollment. Also, we counted all of the cases of CIN 3 diagnosed within 2 years and all of the cases of invasive cancer found during 7 years of follow-up, because routine colposcopy misses a meaningful fraction of ≥CIN 3 (33). Virtually all of the women in the cohort underwent a multiple-technique rescreening at least once. We believe that very few cases were missed by this approach. Nonetheless, it is of great importance that a few cases escaped detection both at enrollment and during follow-up by all four of the methods: visual, two cytological, and molecular. One postmenopausal woman was discovered at year 7 to have advanced cervical cancer that apparently had progressed unseen and unsampled in the endocervical canal. Even with the newest techniques and full participation, there will be a very few cases that will escape detection, keeping us from completely eradicating cervical cancer.

Screening performance is likely to be affected by many study-specific factors that must be viewed as caveats to the generalization of our findings. First, Guanacaste is a high-risk province (annual age-adjusted cervical cancer incidence rates averaged >30/100,000, 1983–1993) and the prevalent cases of ≥CIN 3 might be more easily detected than the smaller lesions seen in better screened populations. Moreover, the specificity of any cervical cancer screening technique is influenced by the underlying population prevalence of HPV infection, and associated mild cytology or visual abnormalities. As a result, all four of the fundamental screening statistics (sensitivity, specificity, and positive and negative predictive values) will vary for cervical cancer screening of greatly different populations.

Three other potentially important limitations of study design deserve specific mention. First, the intervals of follow-up were 6 months, 1 year, or 5–7 years as determined by the perceived risk of incident ≥CIN 3. The only HPV test results available at enrollment to guide this choice were from Hybrid Capture Tube test, which is insensitive. As a result, many women that were actually HPV DNA positive by PCR were placed into a group not seen again until years 5–7. As a related second point, HPV testing was not used as a basis for colposcopic referral unless the HPV type persisted for at least 5 years. However, we only included cases of CIN 3 diagnosed within 2 years. Therefore, it is possible that we undercounted HPV-positive CIN 3 due to delayed diagnosis, leading to decreased estimates of the sensitivity, specificity, and positive predictive value of HPV testing relative to the other three techniques. Third, conventional and liquid-based cytology were interpreted by different pathologists. We attempted to optimize each technique with support from a pathologist that was expert with that technique. As we stated earlier, the interpretation of liquid-based cytology as performed was extremely sensitive and less specific than is typically reported. For pathologists to change preferred methods is not a simple matter; a learning curve definitely exists. We do not mean to suggest that our comparisons of liquid-based cytology and conventional cytology were pathologist-independent.

Conclusions

The applicability of our study data will vary greatly by region. In resource-poor settings with high prevalence of cervical cancer, the main gains in controlling cervical cancer come from the establishment of integrated programs of high coverage among the groups at risk (*i.e.*, ones that include communications, screening, diagnoses, treatment, and follow-up (34)). When no such program is practical, test accuracy is not the limiting factor. Sequential screening strategies might have some role in regions where specificity is essential to limit cost and permit viable screening programs. In medium-resource regions, the transfer of the newest, most accurate screening strategies such as liquid-based cytology performed by an expert and HPV

DNA testing might be economically and technically feasible, but only if the cost of new technologies is drastically tiered.

In areas with already-high cytology screening coverage and reasonable compliance with follow-up recommendations, new technology might soon permit lengthened screening intervals (10). To reduce the incidence of invasive cancer additionally it might be important to increase screening sensitivity (35); however, to the degree that compliance to screening recommendations and follow-up of detected abnormalities are imperfect, greater gains against incidence and mortality might come from improving coverage and follow-up.[6]

It is important that we did not see great performance variability by age and risk factor profile. We believe that clinicians can probably not identify a high-risk group based on interview to be screened differently than a lower-risk majority. And we could not identify a safe, later age such as 30 to begin screening, because many cases (even a few women with cancer) occurred among younger women.

From the clinician and patient perspective, the most important parameters are the predictive values of positive and negative test results. In fact, the positive predictive values of the best single or two test combinations were very low, because HPV and its associated microscopic or visual changes are very common. As a result, the reassurance of negative screening can be very high, but the challenge now will be to find an affordable and efficient strategy for the management (triage) of women who screen positive.

We hope that the details given in the Appendix will help somewhat in making choices regarding screening strategies. Given the abundant options for detection of its precursors, the hopeful fundamental message should be that cervical cancer is controllable, if not eradicable, using a choice of several techniques. It is now a matter of policy and financing (36, 37), not a conceptual or technical challenge, which prevents us from eliminating this malignancy.

[6] C. Ferreccio, unpublished observations.

Acknowledgments

We thank Deidra Kelly (Johns Hopkins University) and Dr. Laurie Mango (Neuromedical Systems, Suffern, NY) for their collaboration in the interpretation of cytological specimens. Reagents and services were supplied or discounted by Cytyc Inc. (Boxborough, MA), National Testing Laboratories (Fenton, MO), Utah Medical (Midvale, UT), and Neuromedical Systems (Suffern, NY). We offer special recognition for the excellent work of the study staff in Costa Rica, in particular Fernando Cardenas, Manuel Barrantes, Elmer Perez, Lidia Ana Morera, and Iris Ugarte. We also acknowledge the collaboration of health authorities in Costa Rica for their enthusiastic support of this project.

Appendix

One-hundred and twelve screening strategies for detection of ≥CIN 3, sorted by decreasing Youden's Index (Prevalence of ≥CIN 3 = 1.286%).

Appendix. 112 Screening strategies for detection of ≥CIN 3, sorted by decreasing Youden's Index (Prevalence of ≥CIN 3 = 1.286%)

Screening combination	% Sensitivity	Specificity	Positive predict. value	Negative predict. value	Youden's index (with 95% CI)	Total number screened by combination
Smear = HSIL+ or HPV+	90.74%	87.77%	8.76%	99.86%	0.79 (0.73–0.84)	8456
ThinPrep = HSIL+ or HPV+	90.48%	87.97%	9.03%	99.86%	0.78 (0.73–0.84)	8058
Cervigram = P2+ or HPV+	89.72%	88.07%	8.82%	99.85%	0.78 (0.72–0.84)	8428
Cervigram = P0+ or HPV+	93.46%	84.15%	7.05%	99.90%	0.78 (0.73–0.82)	8428
Smear = LSIL+ or HPV+	91.67%	85.79%	7.70%	99.87%	0.77 (0.72–0.83)	8456
Cervigram = P1+ or HPV+	91.59%	85.55%	7.54%	99.87%	0.77 (0.72–0.82)	8428
ThinPrep = ASCUS+ or Cervigram = P0+	93.20%	83.90%	7.02%	99.89%	0.77 (0.72–0.82)	7996
ThinPrep = LSIL+ or HPV+	90.48%	86.61%	8.19%	99.86%	0.77 (0.71–0.83)	8058
ThinPrep = ASCUS+ or Cervigram = P1+	91.26%	85.19%	7.44%	99.87%	0.76 (0.71–0.82)	7996
Smear = ASCUS+ or HPV+	91.67%	84.21%	6.99%	99.87%	0.76 (0.71–0.81)	8456
ThinPrep = LSIL+ or Cervigram = P0+	84.47%	91.26%	11.20%	99.78%	0.76 (0.69–0.83)	7996
Cervigram = P3 or HPV+	86.92%	88.21%	8.66%	99.81%	0.75 (0.69–0.82)	8428
ThinPrep = ASCUS+ or Cervigram = P2+	87.38%	87.53%	8.38%	99.81%	0.75 (0.68–0.81)	7996
ThinPrep = ASCUS+ or HPV+	94.29%	80.21%	5.92%	99.91%	0.74 (0.70–0.79)	8058
ThinPrep = LSIL+ or Cervigram = P1+	81.55%	92.69%	12.71%	99.74%	0.74 (0.67–0.82)	7996
Smear = HSIL+ or ThinPrep = ASCUS+	86.54%	87.58%	8.37%	99.80%	0.74 (0.68–0.81)	8036
ThinPrep = ASCUS+ or Cervigram = P3	86.41%	87.69%	8.39%	99.80%	0.74 (0.67–0.81)	7996
HPV+	85.32%	88.22%	8.58%	99.78%	0.74 (0.67–0.80)	8519
ThinPrep = ASCUS+	85.71%	87.80%	8.47%	99.79%	0.74 (0.67–0.80)	8082
ThinPrep = ASCUS+ and HPV+	77.14%	95.84%	19.66%	99.69%	0.73 (0.65–0.81)	8058
Smear = LSIL+ or ThinPrep = ASCUS+	86.54%	85.92%	7.46%	99.79%	0.72 (0.66–0.79)	8036
Smear = HSIL+ or ThinPrep = LSIL+	75.96%	95.13%	16.99%	99.67%	0.71 (0.63–0.79)	8036
ThinPrep = LSIL+ or Cervigram = P2+	75.73%	95.31%	17.41%	99.67%	0.71 (0.63–0.79)	7996
Smear = ASCUS+ or ThinPrep = ASCUS+	86.54%	84.39%	6.78%	99.79%	0.71 (0.64–0.78)	8036
Cervigram = A+ or HPV+	94.39%	75.89%	4.79%	99.91%	0.70 (0.66–0.75)	8428
ThinPrep = HSIL+ or Cervigram = P0+	75.73%	94.24%	14.63%	99.67%	0.70 (0.62–0.78)	7996
ThinPrep = ASCUS+ or Cervigram = A+	94.17%	75.46%	4.77%	99.90%	0.70 (0.65–0.74)	7996
Smear = ASCUS+ or ThinPrep = LSIL+	77.88%	91.17%	10.37%	99.68%	0.69 (0.61–0.77)	8036
Smear = LSIL+ or ThinPrep = LSIL+	75.96%	92.93%	12.34%	99.66%	0.69 (0.61–0.77)	8036
ThinPrep = HSIL+ or Cervigram = P1+	72.82%	95.69%	18.07%	99.63%	0.69 (0.60–0.77)	7996
ThinPrep = LSIL+ or Cervigram = A+	86.41%	81.96%	5.88%	99.78%	0.68 (0.62–0.75)	7996
ThinPrep = LSIL+ or Cervigram = P3	72.82%	95.48%	17.36%	99.63%	0.68 (0.60–0.77)	7996
ThinPrep = LSIL+	71.43%	95.55%	17.44%	99.61%	0.67 (0.58–0.76)	8082
Smear = ASCUS+ or ThinPrep = HSIL+	73.08%	93.24%	12.42%	99.62%	0.66 (0.58–0.75)	8036
Smear = LSIL+ or ThinPrep = HSIL+	71.15%	95.13%	16.09%	99.60%	0.66 (0.58–0.75)	8036
Smear = LSIL+ or Cervigram = P0+	74.53%	90.91%	9.50%	99.64%	0.65 (0.57–0.74)	8389

Appendix. Continued

Screening combination	% Sensitivity	Specificity	Positive predict. value	Negative predict. value	Youden's index (with 95% CI)	Total number screened by combination
Smear = HSIL+ or ThinPrep = HSIL+	67.31%	98.08%	31.53%	99.56%	0.65 (0.56–0.74)	8036
Smear = LSIL+ or Cervigram = P1+	72.64%	92.54%	11.08%	99.62%	0.65 (0.57–0.74)	8389
ThinPrep = HSIL+ or Cervigram = A+	80.58%	84.29%	6.27%	99.70%	0.65 (0.57–0.73)	7996
Smear = ASCUS+ or Cervigram = P0+	75.47%	89.12%	8.15%	99.65%	0.65 (0.56–0.73)	8389
Smear = ASCUS+ or Cervigram = P1+	73.58%	90.69%	9.19%	99.63%	0.64 (0.56–0.73)	8389
ThinPrep = LSIL+ and HPV+	66.67%	97.18%	23.81%	99.55%	0.64 (0.55–0.73)	8058
Smear = LSIL+ or Cervigram = A+	81.13%	81.54%	5.33%	99.70%	0.63 (0.55–0.70)	8389
Smear = HSIL+ or Cervigram = P0+	68.87%	93.79%	12.44%	99.58%	0.63 (0.54–0.71)	8389
Smear = ASCUS+ and ThinPrep = ASCUS+	65.38%	97.09%	22.74%	99.53%	0.62 (0.53–0.72)	8036
Smear = HSIL+ or Cervigram = P1+	66.98%	95.44%	15.81%	99.56%	0.62 (0.53–0.71)	8389
Smear = ASCUS+ or Cervigram = A+	82.08%	79.98%	4.99%	99.71%	0.62 (0.55–0.69)	8389
Smear = LSIL+ or Cervigram = P2+	66.04%	95.36%	15.42%	99.55%	0.61 (0.52–0.70)	8389
Smear = ASCUS+ or Cervigram = P2+	67.92%	93.44%	11.71%	99.56%	0.61 (0.52–0.70)	8389
ThinPrep = HSIL+ or Cervigram = P2+	62.14%	98.66%	37.65%	99.50%	0.61 (0.51–0.70)	7996
Smear = HSIL+ or Cervigram = A+	75.47%	83.94%	5.67%	99.63%	0.59 (0.51–0.68)	8389
Smear = LSIL+ and ThinPrep = ASCUS+	61.54%	97.49%	24.33%	99.49%	0.59 (0.50–0.68)	8036
Smear = ASCUS+ and ThinPrep = LSIL+	59.62%	98.03%	28.44%	99.46%	0.58 (0.48–0.67)	8036
ThinPrep = HSIL+ or Cervigram = P3	58.25%	98.86%	40.00%	99.45%	0.57 (0.48–0.67)	7996
Smear = LSIL+ or Cervigram = P3	61.32%	95.56%	15.01%	99.48%	0.57 (0.48–0.66)	8389
Smear = ASCUS+ or Cervigram = P3	63.21%	93.64%	11.28%	99.50%	0.57 (0.48–0.66)	8389
Smear = ASCUS+	62.96%	93.73%	11.47%	99.49%	0.57 (0.48–0.66)	8481
ThinPrep = HSIL+	57.14%	98.92%	41.10%	99.43%	0.56 (0.47–0.66)	8082
Smear = HSIL+ or Cervigram = P2+	57.55%	98.44%	32.11%	99.45%	0.56 (0.47–0.65)	8389
Smear = LSIL+ and ThinPrep = LSIL+	57.69%	98.21%	29.70%	99.44%	0.56 (0.46–0.65)	8036
Smear = ASCUS+ and HPV+	57.41%	97.74%	24.70%	99.44%	0.55 (0.46–0.64)	8456
Smear = LSIL+	59.26%	95.63%	14.88%	99.45%	0.55 (0.46–0.64)	8481
Smear = HSIL+ and ThinPrep = ASCUS+	52.88%	98.94%	39.57%	99.38%	0.52 (0.42–0.61)	8036
Smear = LSIL+ and HPV+	53.70%	98.06%	26.36%	99.39%	0.52 (0.42–0.61)	8456
ThinPrep = HSIL+ and HPV+	52.38%	99.21%	46.61%	99.37%	0.52 (0.42–0.61)	8058
Smear = HSIL+ or Cervigram = P3	52.83%	98.66%	33.53%	99.39%	0.51 (0.42–0.61)	8389
Smear = HSIL+	50.93%	98.75%	34.38%	99.36%	0.50 (0.40–0.59)	8481
ThinPrep = ASCUS+ and Cervigram = A+	52.43%	97.20%	19.64%	99.37%	0.50 (0.40–0.59)	7996
Cervigram = A+ and HPV+	52.34%	97.18%	19.24%	99.37%	0.50 (0.40–0.59)	8428
Smear = ASCUS+ and ThinPrep = HSIL+	50.00%	99.36%	50.49%	99.34%	0.49 (0.40–0.59)	8036
Smear = HSIL+ and ThinPrep = LSIL+	49.04%	99.12%	42.15%	99.33%	0.48 (0.39–0.58)	8036
Smear = LSIL+ and ThinPrep = HSIL+	48.08%	99.39%	51.02%	99.32%	0.47 (0.38–0.57)	8036
Cervigram = A+	61.68%	84.78%	4.94%	99.42%	0.46 (0.37–0.56)	8457
Smear = HSIL+ and HPV+	46.30%	99.19%	42.37%	99.30%	0.45 (0.36–0.55)	8456
Cervigram = P0+	50.47%	94.85%	11.16%	99.34%	0.45 (0.36–0.55)	8457
ThinPrep = LSIL+ and Cervigram = A+	45.63%	98.51%	28.48%	99.28%	0.44 (0.35–0.54)	7996
Smear = HSIL+ and ThinPrep = HSIL+	43.27%	99.56%	56.25%	99.26%	0.43 (0.33–0.52)	8036
Cervigram = P1+	45.79%	96.51%	14.41%	99.29%	0.42 (0.33–0.52)	8457
ThinPrep = ASCUS+ and Cervigram = P0+	42.72%	98.91%	33.85%	99.25%	0.42 (0.32–0.51)	7996
Cervigram = P0+ and HPV+	42.06%	98.98%	34.62%	99.25%	0.41 (0.32–0.50)	8428
Smear = ASCUS+ and Cervigram = A+	41.51%	98.50%	26.19%	99.25%	0.40 (0.31–0.49)	8389
ThinPrep = ASCUS+ and Cervigram = P1+	39.81%	99.13%	37.27%	99.21%	0.39 (0.29–0.48)	7996
Cervigram = P1+ and HPV+	39.25%	99.24%	40.00%	99.22%	0.38 (0.29–0.48)	8428
Smear = LSIL+ and Cervigram = A+	38.68%	98.87%	30.37%	99.21%	0.38 (0.28–0.47)	8389
ThinPrep = HSIL+ and Cervigram = A+	36.89%	99.56%	52.05%	99.18%	0.36 (0.27–0.46)	7996
ThinPrep = LSIL+ and Cervigram = P0+	36.89%	99.35%	42.70%	99.18%	0.36 (0.27–0.46)	7996
Smear = ASCUS+ and Cervigram = P0+	36.79%	99.42%	44.83%	99.19%	0.36 (0.27–0.45)	8389
Smear = HSIL+ and Cervigram = A+	35.85%	99.57%	51.35%	99.18%	0.35 (0.26–0.45)	8389
ThinPrep = LSIL+ and Cervigram = P1+	34.95%	99.43%	44.44%	99.15%	0.34 (0.25–0.44)	7996
Smear = LSIL+ and Cervigram = P0+	33.96%	99.55%	49.32%	99.16%	0.34 (0.24–0.43)	8389
Smear = ASCUS+ and Cervigram = P1+	33.96%	99.52%	47.37%	99.16%	0.33 (0.24–0.42)	8389
Smear = HSIL+ and Cervigram = P0+	31.13%	99.77%	63.46%	99.12%	0.31 (0.22–0.40)	8389
ThinPrep = HSIL+ and Cervigram = P0+	31.07%	99.76%	62.75%	99.11%	0.31 (0.22–0.40)	7996
Smear = LSIL+ and Cervigram = P1+	31.13%	99.59%	49.25%	99.12%	0.31 (0.22–0.40)	8389
ThinPrep = HSIL+ and Cervigram = P1+	29.13%	99.81%	66.67%	99.08%	0.29 (0.20–0.38)	7996
Smear = HSIL+ and Cervigram = P1+	28.30%	99.79%	63.83%	99.09%	0.28 (0.20–0.37)	8389
Cervigram = P2+	21.50%	99.68%	46.00%	99.00%	0.21 (0.13–0.29)	8457
ThinPrep = ASCUS+ and Cervigram = P2+	19.42%	99.91%	74.07%	98.96%	0.19 (0.12–0.27)	7996
Cervigram = P2+ and HPV+	16.82%	99.90%	69.23%	98.94%	0.17 (0.10–0.24)	8428
ThinPrep = LSIL+ and Cervigram = P2+	16.50%	99.94%	77.27%	98.92%	0.16 (0.09–0.24)	7996
Smear = ASCUS+ and Cervigram = P2+	16.04%	99.95%	80.95%	98.94%	0.16 (0.09–0.23)	8389
ThinPrep = HSIL+ and Cervigram = P2+	15.53%	99.97%	88.89%	98.91%	0.16 (0.09–0.23)	7996
Smear = HSIL+ and Cervigram = P2+	14.15%	99.98%	88.24%	98.91%	0.14 (0.07–0.21)	8389
Smear = LSIL+ and Cervigram = P2+	14.15%	99.95%	78.95%	98.91%	0.14 (0.07–0.21)	8389
Cervigram = P3	8.41%	99.89%	50.00%	98.84%	0.08 (0.03–0.14)	8457
ThinPrep = ASCUS+ and Cervigram = P3	7.77%	99.96%	72.73%	98.81%	0.08 (0.03–0.13)	7996
Smear = ASCUS+ and Cervigram = P3	7.55%	99.98%	80.00%	98.83%	0.08 (0.02–0.13)	8389
ThinPrep = HSIL+ and Cervigram = P3	6.80%	99.97%	77.78%	98.80%	0.07 (0.02–0.12)	7996
ThinPrep = LSIL+ and Cervigram = P3	6.80%	99.97%	77.78%	98.80%	0.07 (0.02–0.12)	7996
Cervigram = P3 and HPV+	6.54%	99.98%	77.78%	98.81%	0.07 (0.02–0.11)	8428
Smear = HSIL+ and Cervigram = P3	5.66%	99.98%	75.00%	98.81%	0.06 (0.01–0.10)	8389
Smear = LSIL+ and Cervigram = P3	5.66%	99.98%	75.00%	98.81%	0.06 (0.01–0.10)	8389

References

1. Bosch, F. X., Lorincz, A., Munoz, N., Meijer, C. J., and Shah, K. V. The causal relation between human papillomavirus and cervical cancer. J. Clin. Pathol., *55:* 244–265, 2002.

2. Koutsky, L. Epidemiology of genital human papillomavirus infection. Am. J. Med., *102:* 3–8, 1997.

3. Nobbenhuis, M. A., Walboomers, J. M., Helmerhorst, T. J., Rozendaal, L., Remmink, A. J., Risse, E. K., van der Linden, H. C., Voorhorst, F. J., Kenemans, P., and Meijer, C. J. Relation of human papillomavirus status to cervical lesions and consequences for cervical-cancer screening: a prospective study. Lancet, *354:* 20–25, 1999.

4. Ho, G. Y., Burk, R. D., Klein, S., Kadish, A. S., Chang, C. J., Palan, P., Basu, J., Tachezy, R., Lewis, R., and Romney, S. Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia. J. Natl. Cancer Inst., *87:* 1365–1371, 1995.

5. Reid, R., Stanhope, C. R., Herschman, B. R., Crum, C. P., and Agronow, S. J. Genital warts and cervical cancer. IV. A colposcopic index for differentiating subclinical papillomaviral infection from cervical intraepithelial neoplasia. Am. J. Obstet. Gynecol., *149:* 815–823, 1984.

6. Koss, L. The Papanicolaou test for cervical cancer detection. A triumph and a tragedy. JAMA, *261:* 737–743, 1989.

7. Kurman, R. J., Malkasian, G. D., Jr., Sedlis, A., and Solomon, D. From Papanicolaou to Bethesda: the rationale for a new cervical cytologic classification. Obstet. Gynecol., *77:* 779–782, 1991.

8. Nanda, K., McCrory, D. C., Myers, E. R., Bastian, L. A., Hasselblad, V., Hickey, J. D., and Matchar, D. B. Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review. Ann. Int. Med., *132:* 810–819, 2000.

9. Sulik, S. M., Kroeger, K., Schultz, J. K., Brown, J. L., Becker, L. A., and Grant, W. D. Are fluid-based cytologies superior to the conventional Papanicolaou test? A systematic review. J. Fam. Pract., *50:* 1040–1046, 2001.

10. Saslow, D., Runowicz, C., Solomon, D., Moscicki, A-B., Smith, R. A., Eyre, H. J., and Cohen, C. American Cancer Society Cervical Cancer Screening Guidelines 2002. CA: Cancer J. Clin., *52:* 375–376, 2002.

11. Herrero, R., Schiffman, M. H., Bratti, C., Hildesheim, A., Balmaceda, I., Sherman, M. E., Greenberg, M., Cardenas, F., Gomez, V., Helgesen, K., Morales, J., Hutchinson, M., Mango, L., Alfaro, M., Potischman, N. W., Wacholder, S., Swanson, C., and Brinton, L. A. Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: the Guanacaste Project. J. Pan. Am. Health Org., *1:* 362–375, 1997.

12. Herrero, R., Hildesheim, A., Bratti, C., Sherman, M. E., Hutchinson, M., Morales, J., Balmaceda, I., Greenberg, M. D., Alfaro, M., Burk, R. D., Wacholder, S., Plummer, M., and Schiffman, M. Population-based study of human papillomavirus infection and cervical neoplasia in rural Costa Rica. J. Natl. Cancer Inst., *92:* 464–474, 2000.

13. Bratti, M. C., Rodriguez, A. C., Schiffman, M., Hildesheim, A., Morales, J., Alfaro, M., Guillen, D., Hutchinson, M., Sherman, M. E., Eklund, C., Schussler, J., Buckland, J., Cardenas, F., Barrantes, M., Perez, E., Cox, J. T., Burk, R. D., and Herrero, R. Description of a seven-year prospective study of HPV infection and cervical neoplasia among 10, 049 women in Guanacaste, Costa Rica, Rev Panam Salvd Publica, in press.

14. Hutchinson, M., Zahniser, D., Sherman, M. E., Herrero, R., Alfaro, M., Bratti, C., Hildesheim, A., Greenberg, M., Lorincz, A., Morales, J., and Schiffman, M. Utility of liquid-based cytology for cervical carcinoma screening: results of a population-based study conducted in a region of Costa Rica with a high incidence of cervical carcinoma. Cancer (Phila.), *87:* 48–55, 1999.

15. Schneider, D. L., Herrero, R., Bratti, C., Greenberg, M. D., Hildesheim, A., Sherman, M. E., Morales, J., Hutchinson, M. L., Sedlacek, T. V., Lorincz, A., Mango, L., Wacholder, S., Alfaro, M., and Schiffman, M. Cervicography screening for cervical cancer among 8460 women in a high-risk population. Am. J. Obstet. Gynecol., *180:* 290–298, 1999.

16. Sherman, M. E., Schiffman, M., Herrero, R., Kelly, D., Bratti, C., Mango, L. J., Alfaro, M., Hutchinson, M. L., Mena, F., Hildesheim, A., Morales, J., Greenberg, M. D., Balmaceda, I., and Lorincz, A. T. Performance of a semiautomated Papanicolaou smear screening system: results of a population-based study conducted in Guanacaste, Costa Rica. Cancer (Phila.), *84:* 273–280, 1998.

17. Kurman, R. J., and Solomon, D. The Bethesda System. Springer-Verlag, New York, 1994.

18. Schiffman, M., Herrero, R., Hildesheim, A., Sherman, M. E., Bratti, M., Wacholder, S., Alfaro, M., Hutchinson, M., Morales, J., Greenberg, M. D., and Lorincz, A. T. HPV DNA testing in cervical cancer screening: results from women in a high-risk province of Costa Rica. JAMA, *283:* 87–93, 2000.

19. Castle, P. E., Schiffman, M., Gravitt, P. E., Kendall, H., Fishman, S., Dong, H., Hildesheim, A., Herrero, R., Bratti, M. C., Sherman, M. E., Lorincz, A., Schussler, J. E., and Burk, R. D. Comparisons of HPV DNA detection by MY09/11 PCR methods. J. Med. Virol., *68:* 417–423, 2002.

20. Franco, E. L., and Ferenczy, A. Assessing gains in diagnostic utility when human papillomavirus testing is used as an adjunct to papanicolaou smear in the triage of women with cervical cytologic abnormalities. Am. J. Obstet. Gynecol., *181:* 382–386, 1999.

21. Youden, J. W. Index for rating diagnosis tests. Cancer (Phila.), *3:* 32–35, 1950.

22. Brinton, L. A., Reeves, W. C., Brenes, M. M., Herrero, R., Gaitan, E., Tenorio, F., de Britton, R. C., Garcia, M., and Rawls, W. E. The male factor in the etiology of cervical cancer among sexually monogamous women. Int. J. Cancer, *44:* 199–203, 1989.

23. Hildesheim, A., Herrero, R., Castle, P. E., Wacholder, S., Bratti, M. C., Sherman, M. E., Lorincz, A. T., Burk, R. D., Morales, J., Rodriguez, A. C., Helgesen, K., Alfaro, M., Hutchinson, M., Balmaceda, I., Greenberg, M., and Schiffman, M. HPV co-factors related to the development of cervical cancer: results from a population-based study in Costa Rica. Br. J. Cancer, *84:* 1219–1226, 2001.

24. Fahey, M. T., Irwig, L., and Macaskill, P. Meta-analysis of Pap test accuracy. Am. J. Epidemiol., *141:* 680–689, 1995.

25. Tawa, K., Forsythe, A., Cove, J. K., Saltz, A., Peters, H. W., and Watring, W. G. A comparison of the Papanicolaou smear and the cervigram: sensitivity, specificity, and cost analysis. Obstet. Gynecol., *71:* 229–235, 1988.

26. Spitzer, M. Cervical screening adjuncts: recent advances. Am. J. Obstet. Gynecol., *179:* 544–556, 1998.

27. Schneider, D. L., Burke, L., Wright, T. C., Spitzer, M., Chatterjee, N., Wacholder, S., Herrero, R., Bratti, M. C., Greenberg, M. D., Hildesheim, A., Sherman, M. E., Morales, J., Hutchinson, M. L., Alfaro, M., Lorincz, A., and Schiffman, M. Can cervicography be improved? An evaluation with arbitrated cervicography interpretations. Am. J. Obstet. Gynecol., *187:* 15–23, 2002.

28. Autier, P., Coibion, M., De Sutter, P., and Wayemberg, M. Cytology alone versus cytology and cervicography for cervical and cervicography for cervical cancer screening: a randomized study. Obstet. Gynecol., *93:* 353–358, 1999.

29. Fletcher, R. H., Fletcher, S. W., and Wagner, E. H. Clinical Epidemiology. The Essentials. 3rd ed. Baltimore: Williams and Wilkins, 1996.

30. Gardner, I. A., and Greiner, M. Advanced methods for test validation and interpretation in veterinary medicine. Freie Universitat Berlin, 1999.

31. Hasselblad, V., and Hedges, L. V. Meta-analysis of screening and diagnostic tests. Psychol. Bull., *117:* 167–178, 1995.

32. Sackett, D. L., Haynes, R. B., Guyatt, G. H., and Tugwell, P. Clinical Epidemiology. A Basic Science for Clinical Medicine. 2nd Ed. Boston: Little, Brown and Company, 1991.

33. Belinson, J., Qiao, Y. L., Pretorius, R., Zhang, W. H., Elson, P., Li, L., Pan, Q. J., Fischer, C., Lorincz, A., and Zahniser, D. Shanxi Province Cervical Cancer Screening Study: a cross-sectional comparative trial of multiple techniques to detect cervical neoplasia. Gynecol. Oncol., *83:* 439–444, 2001.

34. Suba, E. J., Nguyen, C. H., Nguyen, B. D., and Raab, S. S. The Viet/American Cervical Cancer Prevention Project. *De novo* establishment and cost-effectiveness of Papanicolaou cytology screening services in the Socialist Republic of Vietnam. Cancer (Phila.), *91:* 928–939, 2002.

35. Hutchinson, M., and Farber, F. L. Cost-effectiveness of new cytology technologies. Am. J. Clin. Pathol., *113:* 149–151, 2000.

36. Goldie, S. J., Kuhn, L., Denny, L., Pollack, A., and Wright, T. C. Policy analysis of cervical cancer screening strategies in low-resource settings: clinical benefits and cost-effectiveness. JAMA, *285:* 3107–3115, 2001.

37. Kulasingam, S. L., Hughes, J. P., Kiviat, N. B., Mao, C., Weiss, N. S., Kuypers, J. M., and Koutsky, L. A. Evaluation of human papillomavirus testing in primary screening for cervical abnormalities: comparison of sensitivity, specificity, and frequency of referral. JAMA, *288:* 1749–1757, 2002.

A Comparison of Single and Combined Visual, Cytologic, and Virologic Tests as Screening Strategies in a Region at High Risk of Cervical Cancer

Catterina Ferreccio, Maria C. Bratti, Mark E. Sherman, et al.

Cancer Epidemiol Biomarkers Prev 2003;12:815-823.