ABSTRACT

Objective: To design a cervical cancer screening algorithm for the developing world highly sensitive for CIN II, III, and cancer, and highly specific for CIN II and III, making it possible to ablate the transformation zone without histologic confirmation.

Methods: In rural Shanxi Province, China, we examined 1997 women ages 35-45. Each subject underwent a self-test for HPV (by HC-II assay), fluorescence spectroscopy, a liquid based pap (read manually and by computer and used as a direct test for HPV), a visual inspection (VIA) diagnosis, and colposcopy with multiple cervical biopsies.

Results: Mean age was 39.1 ± 3.16 years, mean number of births of 2.6 ± 0.93. Based on tests administered 43% subjects had ≥ CIN II. All subjects with ≥ CIN II had either a ThinPrep Pap (≥ ASCUS), or a positive HPV direct test. The sensitivity and specificity for the detection of ≥ CIN II was respectively 83% and 86% for the HPV self-test, 95% and 85% for the HPV direct test, 94% and 78% for the ThinPrep Pap (≥ ASCUS), 77% and 98% for the ThinPrep Pap (≥ HGSIL), 94% and 99% for fluorescence spectroscopy, 71% and 74% for VIA, and 84% and 77% for colposcopy.

Conclusion: Based on these data and the existing healthcare infrastructure in China, we believe further refinement of primary HPV screening using centralized labs is indicated. Self-testing in the local villages may be effective with improvements in the devices and techniques.
INTRODUCTION

The incidence of cervical cancer varies widely between and within regions throughout the world \[1,2\]. This is attributable in large part to the variable access to cytologic screening programs to detect and treat pre-invasive disease of the cervix \[3\]. For example, although age-adjusted mortality rate from cervical cancer in The People’s Republic of China (4.29/100,000) is slightly greater than that of the United States \[4\], rates in rural provinces such as Shanxi Province where screening is rarely performed are much higher (52/100,000) \[5,6\]. Methods of screening for cervical neoplasia which are less expensive than cytologic screening and do not require as great a healthcare infrastructure such as visual inspection aided by acetic acid (VIA) \[7,8\] and self-testing for high-risk types of human papillomavirus \[9\] have been proposed.

The Shanxi Province Cervical Cancer Screening Study (SPOCCS) was designed to determine the sensitivity and specificity (critical in low resource settings) of six screening technologies in order to develop low-cost screening for rural China. We searched for a screening algorithm that was sensitive for cervical intraepithelial neoplasia (CIN) II, III, and cancer, and highly specific for CIN II and III. With such an algorithm, it would be possible to ablative the transformation zone in test positive subjects without histological confirmation of the diagnosis. A pilot study of 136 women between the ages of 30 and 45 years was conducted in Xiangyuan and Yangcheng counties in southern Shanxi Province. The prevalence of >CIN II was 8.8%. In addition to confirming the high prevalence of cervical neoplasia, we were able to establish the feasibility of the SPOCCS trial in this rural environment \[10\].

METHODS

The human subject review boards of both the Cleveland Clinic Foundation and The Cancer Institute/Hospital of the Chinese Academy of Medical Sciences (CICAMS) approved the study. A field supervisor from the Department of Epidemiology of CICAMS in Beijing and the local commune nurses recruited two thousand forty-seven women between 35 and 45 years of age. The patients came from villages in four separate communes in Xiangyuan County. Only non-pregnant women with no history of cervical screening, pelvic radiation, or hysterectomy were eligible. All eligible women were invited to participate, and virtually everyone agreed and was accepted until our target number was reached. The women were brought to the clinic where the trial was explained and informed consent was obtained. Each subject was administered a demographic, epidemiological, sexual, and nutritional questionnaire. Blood was drawn (15cc), separated into multiple vials of serum, plasma, white cells, and red cells, and then stored at -84°C. Two American and five Chinese gynecologic oncologists made clinical observations and collected specimens June-July, 1999 (steps 2 thru 6), (figure 1).

The 6 screening steps took place in three examination rooms of the clinic. Steps 1-5 occurred in room #1, and step 6 took place in rooms #2 and #3. All women were subjected to all tests and procedures and all tests were performed by gynecologic oncologists or their fellows, with the few exceptions noted below.

Step 1 -- Self-Test for HPV. Patient placed a Dacron swab 5-7cm into the vagina for 30 seconds, and then placed the swab in a collection tube with STM transport media (Digene Corp., Gaithersburg, MD USA).
Step 2--Fluorescence Spectroscopy. The Optical Biopsy device (BiOptics CX-100, Optical Biopsy Technologies, LLC [Knoxville, Tenn.]) was used to take eight readings, one in each of the quadrants of the exocervix at the squamo-columnar junction (2, 4, 8, and 10 o’clock), and one at each of the same clock positions of the endocervical canal. All diagnoses were determined by the device’s algorithm and the investigators were blinded to the results.

Step 3/4—Liquid based cytology/ Direct HPV test. Using a plastic spatula and an endocervical brush, cervical samples were collected into a liquid medium and a ThinPrep Pap (Cytyc Corp. Boxborough, Mass) was prepared. The fluid remaining was used to obtain cells for direct HPV testing.

Step 5--Visual inspection (VIA). Acetic acid (5%) was applied to the cervix, and after waiting one-minute VIA was performed. The observations were recorded by quadrant. Quadrants defined as low-grade (HPV and CIN I) had pale white lesions that might or might not abut the squamo-columnar junction. Quadrants defined as high-grade (CIN II, III) had dense white lesions with sharp borders; one border of these high-grade lesions lesions always abutted the squamo-columnar junction. Cancer was diagnosed when a friable mass with an irregular surface was seen.

Step 6--Colposcopy. Acetic acid 5% was reapplied to the cervix, colposcopy performed and observations were recorded by quadrant. If a quadrant was abnormal, the data sheet reflected the most severe abnormality within that quadrant. If the examination showed no abnormalities in a quadrant, a biopsy was taken from that quadrant at 2, 4, 8 or 10 o’clock on the exocervix at the squamo-columnar junction depending on the quadrant. It was acceptable to take more than one biopsy per quadrant. All biopsies were performed with a bronchoscopy biopsy instrument that has 2 mm jaws and is virtually painless for most patients. An endocervical curetage (ECC) was also done on every patient.

The sealed collection tubes containing the Dacron HPV swabs were sent to Digene Corp. (Gaithersburg, MD) for analysis. They were tested with the second-generation Hybrid Capture micro plate–based human papillomavirus test (HC II test, Digene Corp. Gaithersburg, MD). The test uses an RNA probe mixture of 13 intermediate and high-risk anogenital HPV viral types [11]. A cut-off value of 10 pg HPV DNA was used for positive. HPV testing was performed blinded to the results of other tests. PreservCyt fluid from the ThinPrep samples was also sent to Digene for HPV testing, coded and labeled separately from the self-test swabs.

The ThinPrep Pap tests were read locally by a cytopathologist from Beijing so appropriate follow-up could be initiated. The Cleveland Clinic cytopathologists reviewed all abnormal and 5% of normal slides. Pap smear diagnoses followed the Bethesda system. The biopsies were also processed and interpreted in Beijing, and then sent to the Cleveland Clinic for a similar review. The pathologists were not aware of the results from any of the other tests. No significant differences in diagnoses were seen between Beijing and Cleveland Clinic. The final diagnosis therefore represents the Beijing diagnosis and in the case of the histology was the worst histological diagnosis from any quadrant or the ECC.

Once the cytological and biopsy results were known, the patients with pathology were identified. Patients needing follow-up and/or therapy returned to the Xiangyuan County Women and Children’s Clinic where with the assistance of the Beijing staff, appropriate treatment was rendered. Patients needing radical surgery and/or radiotherapy were referred to a larger facility.
A sample size of 2000 women was chosen to provide 95% power to detect a 10% difference (2 sided \( p=0.05 \) test) in diagnostic accuracy between any pair of tests. For each test, the statistics were calculated based on the number of test procedures performed rather than the number of results obtained. Subjects for whom a valid test result was not available were classified as having less than CIN II or negative. The various tests were compared, with respect to performance, by comparing the partial area under the receiver operating characteristic curve (ROC) in a pair-wise manner [12]. The ROC curves presented are based on values obtained after fitting a binormal distribution to the subjects’ results obtained from each test [13].

RESULTS

We recruited 2047 women for the study. Fifty patients were excluded from analysis: 44 had a pap and colposcopy only due to a changeover of personnel during the study; two were <35 years of age; two were having their menses; one was pregnant; and one had been screened in the pilot study. Therefore 1997 patients are the subject population.

The mean age of the women screened was 39.1 ± 3.16 years. The mean number of pregnancies was 3.1 ± 1.27, range: 0-13, and a mean number of births of 2.6 ± 0.93, range: 0-7. 98.5% of the women were currently married, 93.3% never smoked, and 0.3% had a history of condyloma. 20% had evidence of trichomoniasis on their cervical cytology.

Biopsy specimens were unsatisfactory in 21/7988 (0.3%) of the quadrants. No patient had more than a single unsatisfactory quadrant. The endocervical biopsies were unsatisfactory in 131/1997 (6.6%). Overall of the enrolled women, 1784 (89.3%) of the subjects had a final diagnosis of negative, 127(6.4%) had CIN I, 43(2.2%) had CIN II, 31(1.6%) had CIN III, and 12(0.6%) had cancer. Therefore 4.3% subjects had \(^\geq\) CIN II. All subjects with \(^\geq\) CIN II had either an abnormal ThinPrep Pap (\(^\geq\) ASCUS), or a positive HPV direct test. The sensitivity and specificity and the positive and negative predictive values of the tests studied are presented in Table 1.

Sixteen patients had false negative colposcopy. The ECC was the only positive biopsy in 2 of 12 with CIN II, and 3 of 4 with CIN III. Twelve cancers were diagnosed. All were squamous cell type. Only two of 12 cancers occurred in women under the age of 40.

Comparison of the partial areas under these ROC curves suggests that the test procedures can be combined into three groups. The area selected for comparison corresponds to false positive rates <10% and reflects the range of false positive rates we were most interested in for screening. Fluorescence spectroscopy had a very low specificity in this study, 9%; therefore it was not included in these comparisons (figure 2).

The ThinPrep Pap test performed significantly better than the other procedures in the range of false positive rates considered, \( p<0.0001 \) compared to each of the other procedures. The HPV direct test and colposcopy performed similarly (\( p=0.61 \)) but were each significantly better than either the HPV self-test (\( p<0.0001 \) compared to HPV direct test and .01 compared to colposcopy) or visual inspection (\( p<0.0001 \) compared to HPV direct test and \( p=0.002 \) compared to colposcopy). There was no significant difference between the HPV self test and visual inspection (\( p=.94 \)).

DISCUSSION

The major reason for performing this complex study in rural China is that the prevalence of pre-invasive and invasive cervical cancer in this Province is so high. In addition if we hope to develop a screening algorithm for rural China, then this rural environment is where the studies...
need to be done. The SPOCCS study is unique since all patients received identical screening tests and more importantly, all patients were biopsied. In most other screening studies, it is the abnormal Pap smear, HPV test, and/or abnormal colposcopy that lead to biopsy [9,14]. Women with negative screening tests and/or negative colposcopic evaluations are presumed to be true negatives. In such trials, estimates of sensitivity and specificity may be affected by work-up bias resulting in improved performance of a screening test when the prevalence of disease is low [15]. In this trial, all 86 women with CIN II or worse had either an abnormal ThinPrep Pap test or a positive direct test for high-risk HPV. The presumption that > CIN II would not occur among the 1,352 women who had tested negative on ThinPrep Pap and direct HPV test proved to be valid. The presumption that > CIN II would not occur among the 1,478 women who had negative colposcopic evaluations, however, was not proven valid as 16 of these women had CIN II or worse.

In applying our results to other populations it is important to note the current study was performed in women age 35 to 45 years old. Limiting the study to women age 35 to 45 years was done purposefully because we anticipate screening the population only once and ablating the transformation zone without histologic confirmation. In this population, women age 35 have a high risk of CIN II and CIN III and a low risk of invasive cancer. Therefore they are most likely to benefit from this strategy. When compared to screening trials that include young women (age 18 to 35 years), the specificities of some of the screening tests in the current trial are likely to be higher. This is particularly apparent for the HPV tests in which the prevalence of HPV is high in young women [16].

The 95% sensitivity of the direct HPV test in detection of > CIN II found in this series is somewhat higher than the 84% reported by Wright et al. and the 88% reported by Schiffman et al. [9,14] and less than the 100% reported by Clavel et al. [17]. The specificity of the HPV direct test in the detection of CIN II or worse is similar to the 83% reported by Wright et al. [9] and somewhat lower than the 89% reported by Schiffman et al. [14]. If 35 to 45 year-olds were screened with direct HPV and ablation of the transformation zone were restricted to women who had positive direct tests and did not have visual inspections consistent with invasive cancer, 18% of women would undergo ablation. The sensitivity of the ThinPrep Pap in the detection of CIN II or worse in this trial is identical to the 94% reported by Bolick [18]. The sensitivity of the ThinPrep Pap in this study is much higher than the 51% reported by the Evaluation of Cervical Cytology, AHCPR Report for conventional cytology and the 80% reported for thin layer cytology [15]. By changing the definition of an abnormal ThinPrep Pap from >ASCUS to >HGSIL, it is possible to increase the specificity of the later from 78% to 98% at the cost of decreasing the sensitivity from 94% to 77%. If only 35 to 45 year-olds were screened with the ThinPrep Pap tests and positives limited to > HGSIL (5% of the screened population), 91% of > CIN III would be identified.

The disadvantage of screening cytology, whether liquid based or conventional, is the infrastructure required and the need for skilled caregivers. Also, at least one additional visit is required for ablation of the transformation zone in women who test Pap positive. Although HPV testing also requires waiting for results, it is easily centralized and does not require the highly skilled personnel necessary to interpret cytology smears. In addition centralization diminishes the overall cost of the laboratory equipment.
Fluorescence spectroscopy performed poorly in this trial. In the United States trials examining the cervix the device had sensitivity 86.7% and specificity 70.1% [19]. A possible explanation for the poor performance of the device in the current trial is that it was used before the application of acetic acid solution while in the previous trials it was used after the acetic acid solution [20]. In addition many women in this trial were observed to have severe chronic cervicitis.

Visual inspection after the application of acetic acid (VIA) is receiving considerable attention as a potential screening technique in the developing world [7,8]. In 1996, Megevand reported VIA to have a sensitivity for ≥ CIN II of >60% [21]. In Zimbabwe the sensitivity and specificity for ≥ CIN II was 77% and 64% respectively [7]. In our study VIA had a sensitivity for ≥ CIN II of 71%, a specificity of 74%, and it failed to identify 1/5 of the cancers. Although disappointing, we recognize the amount of acute cervicitis in these patients made visual inspection quite difficult. If 35 to 45 year-olds were screened with VIA and ablation of the transformation zone was limited to women who had a lesion that was not consistent with invasive cancer then 27% of women would be treated.

Screening systems based on VIA have two significant advantages over those based on HPV testing or ThinPrep Pap tests. First, VIA is inexpensive, and second it allows for immediate treatment. In rural situations any delay in the screening, diagnosis, and treatment sequence may lead to high rates of loss to follow-up. Visual inspection, like cytological interpretation, requires training, some expertise, and quality control.

Visual inspection in this trial, was performed by a Chinese Gynecologic Oncologist and the cytological interpretation of ThinPrep smears was done by the Chief of Cytopathology of CICAMS. Reproducibility could be a problem in routine screening if those less experienced perform visual inspection and cytologic interpretation. Reproducibility of tests is less of a problem with a HPV Hybrid Capture II [14].

Some decisions are easy. Fluorescence spectroscopy, with the specific machine and algorithm used in this study should not be adopted. Colposcopy as a screening test should not be adopted because it was no better than direct HPV testing and inferior to ThinPrep in the range of false positive rates we considered (< 10%); in addition, it is expensive and would be difficult to implement in rural China. HPV self-test could be ideal in certain environments and especially if newer collection devices show sensitivity approaching that of direct HPV testing. Visual inspection might be adopted despite its shortcomings (decreased sensitivity), since it would likely be less expensive and would not require a second visit.

We plan to perform a randomized trial comparing screening with HPV to no screening. New technology may allow this to be by self-test. To limit the number of screened women who already have invasive cancer, the one-time screening will be performed on women age 30 to 35 years rather than on women age 30 to 45 years. Women with positive HPV tests will have VIA. The few women with VIA consistent with invasive cancer will be referred for biopsy, all others will undergo ablation of the transformation zone.

It is critical to carry out cost–benefit analyses to assess the appropriate application of screening technologies. In 1998 the annual health care expenditure per capita in rural China was $16.00 [22]. In the developing world, the competition for health care dollars is great and it is necessary to prioritize cancer screening alongside childhood immunization, nutrition, sanitation, refrigeration, and the management of infectious diseases.

**ACKNOWLEDGEMENTS**

Funding: Taussig Cancer Center, Cleveland Clinic Foundation; Cancer Institute/Hospital, Chinese Academy of Medical Sciences; Terry Fox Foundation; Transamerica Corporation; DiGene Corporation; Cytyc Corporation; Optical Biopsy Tech., LLC.; Carl Zeiss, Inc.
We are indebted to Mr. Anthony Yen for his counsel and assistance throughout this project. We wish to thank Dr. Alex Ferenczy for reviewing our manuscript, and Monique Washington RN, MS for her assistance in preparation of our manuscript.

REFERENCES


ARTICLE PRÉCIS

This study determines the sensitivity and specificity of six screening technologies for cervical cancer. All patients had a minimum of five cervical biopsies.

Figure 1- The Study Protocol

Table 1 – Sensitivity and Specificity of the Screening Tests

<table>
<thead>
<tr>
<th></th>
<th># Positive</th>
<th>Sensitivity for CIN II</th>
<th>Specificity for CIN II</th>
<th>Sensitivity for CIN III</th>
<th>Specificity for Cancer</th>
<th>Positive Predictive Value</th>
<th>Negative Predictive Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPV Self-Test</td>
<td>17%</td>
<td>83%</td>
<td>86%</td>
<td>81%</td>
<td>75%</td>
<td>21%</td>
<td>99.1%</td>
</tr>
<tr>
<td>(&gt;10 pg/ml)</td>
<td>340/1,997</td>
<td>71/86</td>
<td>35/43</td>
<td>9/12</td>
<td>71/340</td>
<td>1642/16</td>
<td>57</td>
</tr>
<tr>
<td>Test</td>
<td>Sensitivity</td>
<td>Specificity</td>
<td>Accuracy</td>
<td>Positive Predictive Value</td>
<td>Negative Predictive Value</td>
<td>Overall Accuracy</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>----------</td>
<td>---------------------------</td>
<td>---------------------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>Fluorescence Spectroscopy</td>
<td>91%</td>
<td>94%</td>
<td>93%</td>
<td>100%</td>
<td>5%</td>
<td>97%</td>
<td></td>
</tr>
<tr>
<td>ThinPrep® Pap (³ ASCUS)</td>
<td>25%</td>
<td>94%</td>
<td>78%</td>
<td>98%</td>
<td>100%</td>
<td>99.7%</td>
<td></td>
</tr>
<tr>
<td>ThinPrep® Pap (³ LGSIL)</td>
<td>10%</td>
<td>87%</td>
<td>94%</td>
<td>93%</td>
<td>100%</td>
<td>99.4%</td>
<td></td>
</tr>
<tr>
<td>ThinPrep® Pap (³ HGSIL)</td>
<td>5%</td>
<td>77%</td>
<td>98%</td>
<td>91%</td>
<td>100%</td>
<td>99%</td>
<td></td>
</tr>
<tr>
<td>HPV Direct Test (³ 1.0 pg/ml)</td>
<td>18%</td>
<td>95%</td>
<td>85%</td>
<td>98%</td>
<td>100%</td>
<td>99.8%</td>
<td></td>
</tr>
<tr>
<td>Visual Inspection (Any abnormal)</td>
<td>28%</td>
<td>71%</td>
<td>74%</td>
<td>79%</td>
<td>67%</td>
<td>98%</td>
<td></td>
</tr>
<tr>
<td>Colposcopy (Any abnormal)</td>
<td>26%</td>
<td>81%</td>
<td>77%</td>
<td>91%</td>
<td>100%</td>
<td>99%</td>
<td></td>
</tr>
</tbody>
</table>

17 insufficient material, 2 not reported 17 insufficient material, 2 not reported
Sixty-three excluded due to no probes

3 unsatisfactory, 1 results not reported

40 insufficient material, 17 not received/lost (using results obtained instead of tests administered, sensitivity of 

\(^5\text{CIN II}=98\%, \text{and } ^5\text{CIN III}=100\%)\)