Screening Colposcopy: Evaluation of a Novel Method
For Cervical Cancer Screening in Low Resource Settings

Jessica J. Kenerson

A master’s paper submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of MSPH in the Department of Health Care and Prevention of the School of Public Health

Chapel Hill
2006

Approved by

First Reader: Dr. David Walmer

Second Reader: Dr. Vijaya Hogan
Abstract

Background: Strategies to decrease the burden of cervical cancer in developed countries are often difficult to implement in low resource settings, which has stimulated researchers and others to search for novel approaches for screening. Colposcopy offers some advantages that might make it a reasonable strategy as a standalone screen and treat protocol in communities where the incidence of disease is high. To date, results for the screening accuracy of colposcopy as a screening tool have varied widely from low to high sensitivity and specificity in different studies.

Methods: In this study, I analyzed cross sectional data from the medical records of 2,094 self referred women in a community cervical screening program in Haiti. The patients included in the final analyses received some combination of Pap smear, colposcopy and/or cervical biopsy, which allowed me to compare these techniques. The final study populations included 198 women who had both Pap smear and cervical biopsy, and 221 patients who had both colposcopy and cervical biopsy. The main outcome measure (cancer and precancer diagnoses) were operationalized as mild dysplasia or worse lesions on cervical biopsy. Diagnostic accuracy was determined by calculating the sensitivity, specificity, positive predictive values, and negative predictive values with 95% confidence intervals using cervical biopsy as the gold standard.

Results: The prevalence of dysplasia or cancer on biopsy for the study populations was 41.9%-48.0% while prevalence detected by Pap was 30.8%. Prevalence in the study populations detected by colposcopy was 66.7%-70.1%.
Pap smear had a sensitivity of 68.7% (CI 57.6-78.4) and specificity was 96.5% (CI 91.0-99.0) using biopsy as the gold standard. The sensitivity for colposcopy was 96.2% (CI 91.0-99.0) compared to biopsy, and a specificity of Sp 53.9% (CI 44.0-63.0).

**Conclusion:** Colposcopy identifies a population that includes most of the dysplasia at a price of including many false positives, placing the screened population at a significant risk of over treatment. Therefore, colposcopy is a reasonable screening strategy if the disease prevalence is high in the population of interest and the co-morbidity from over treatment is low.

**Introduction**

Cervical cancer disproportionately affects resource-poor nations, where 85% of deaths from cervical cancer and 80% of new cases arise annually \(^1\)\(^-\)\(^4\). Haiti shares a heavy amount of this burden. For example one study estimated an age-standardized mortality rate (ASR) of 53.5 per 100,000 person years (p-y) in 2005\(^5\). This contrasts starkly with rates in a mid-income country such as Argentina, in which the same study estimated the mortality rate at 7.3 ASR per 100,000 p-y\(^5\). The need for adequate prevention services is substantial and effective prevention efforts in low resource nations have been analyzed in many studies\(^1\)\(^,\)\(^3\)\(^,\)\(^6\)\(^-\)\(^9\). One strategy that has been developed from this research, the see and treat method, has demonstrated the ability to decrease prevalence of precursor cancer lesions using HPV DNA testing or Visual Inspection with Acetic Acid.
Colposcopy has characteristics that make it suitable for use in a screen and treat protocol; however diagnostic accuracy has varied widely between studies.

Barriers in access to health services in these nations, and especially Haiti, create special challenges in the development of screening programs. Factors including high rates of loss to follow-up, lack of transportation and finances for treatment, lack of existing health systems to address the issues, lack of knowledge, and cultural traits that hinder women from obtaining cervical screening have impeded efforts in many countries. Because of these difficulties, women that access care often do so because symptoms have impeded with their daily livelihood and/or they have an advanced and untreatable stage of cervical cancer with a tertiary care that is severely limited or non-existant. The interaction of these factors fuels the need for novel screening and treatment protocols. One-day screen and treat protocols have been instituted in the past decade in many low resource nations in order to curb high incidence and mortality from cervical cancer.

Colposcopy as a Screening Method

The best screening method for cervical cancer is one that can reduce the burden of disease while also proving to be inexpensive and reproducible, independent of pathology services, minimizing over treatment, side effects or complications, and socio-culturally acceptable. Because of the interplay of political and economic factors at the macro level in Haiti, this nation’s medical systems and public health
infrastructures have been severely weakened\textsuperscript{17,18}. As such, prior to this initiative, services for cervical cancer screening were nonexistent. For the past 13 years, Family Health Ministries, Inc. (FHM), a US based 501c3 nonprofit organization, has supported efforts to develop a cost-effective cervical cancer prevention program in Leogane, Haiti. This study is FHM’s first evaluation of the efficacy of this screening method, yet patient acceptability and cost effectiveness remain to be determined.

The screening and treatment protocol initially instituted thirteen years ago involved conventional methods used in the US and Europe, yet issues have arisen hindering effective screening. For example, physicians asked their patients to return for Pap smears, colposcopy, and biopsy results or evaluations, but follow up rates were very low. Furthermore, many Pap smear results could not be determined because inflammation on the slides obscured the sample. With assistance from FHM, they incorporated a new protocol for screening women into the typical screening strategy used in the United States. It involves making blinded preliminary diagnoses of the cervix using colposcopy, which has the potential for use in a one day screen and treat protocol. This method does not stray from US standard of care and allows for a comparison between colposcopy diagnoses and the gold standard diagnoses with biopsy. As a screening method, colposcopy involves the application of acetic acid to the cervix followed by visual inspection with the colposcope for dysplastic lesions. The physician’s diagnostic impression was recorded. If colposcopy demonstrates good accuracy,
cryotherapy could be performed to remove lesions on the same visit. This is called a “see and treat” method and is different from diagnostic colposcopy, the conventional use for colposcopy. Traditionally, a colposcopy procedure requires a biopsy when abnormalities are detected. The see and treat method does not require biopsy before treatment, and if accurate, pathology services would not be necessary.

Using colposcopy as a first line screening method is unique and currently published data addressing its efficacy in low resource settings is limited. Traditionally, colposcopy has only been used as a diagnostic tool in developing nations. It requires a skilled professional and the cost and maintenance of a colposcope may be prohibitive. Yet, once these resources have been acquired, as in the case of this Haitian program, it may prove to be more cost effective while also serving the main goal of decreasing cervical cancer mortality.

**Purpose of the Study**

The purpose of this study is to perform a diagnostic accuracy analysis of the see and treat colposcopy method and Pap smear analysis by determining the sensitivity and specificity of each test compared to biopsy results as recorded in the Leogane screening program. Few published studies describe the use of colposcopy as a screening method, or its use in a see and treat protocol. The screening colposcopy procedures used in Haiti might be performed effectively in situations where there is a high risk of loss to follow up in the patient population,
while still adequately screening for cervical cancer. Because of the unknown efficacy of this method, traditional colposcopy was also performed on each patient with biopsy when dysplastic-appearing lesions were present. If these lesions were present near the endocervical canal, endocervical curettage and/or biopsy of the ectocervix was performed. A Pap smear interpreted by adequate cytology services has been performed on nearly all colposcopy patients in order to ensure accuracy and precision of colposcopy.

In this study, I will compare the effectiveness of Pap smears and screening colposcopy as primary screening methods to identify cervical dysplasia in a population subject to poor access to resources. The colposcope may be ideal for use as a screening tool in Leogane because

1. It is more sensitive than Pap smears for identifying dysplasia\textsuperscript{20}
2. It is less dependent on pathology services at a site that has acquired equipment and a skilled physician
3. It is less expensive than Pap smear after the skills and equipment have been acquired
4. It decreases the number of visits that women have to make to complete their diagnosis and therapy\textsuperscript{19}

**Literature Review**

Many public health systems and/or advocacy organizations worldwide have evaluated the use of see and treat methods in low resource settings in order to
determine their efficacy compared to multiple visit screening, diagnostic and
treatment methods. A summary of the results for all methods is provided in Table
1. This section reviews currently used methods in order to identify factors that
may lead to the success or failure of a novel colposcopy screen and treat method.

*Conventional Cervical Cytology*

Conventional cervical cytology with collection and examination of cervical
cytology has been the most widely implemented strategy worldwide, yet it
requires a health professional to collect the sample, cytotechnicians to process,
stain and read smears, and a cytopathologist to do final reporting. Reliable
testing depends on the presence of a quality controlled laboratory and personnel, a
method to relay results to patients, and usually a time lag between sample
collection, diagnosis, and reporting. Additionally, cytology does not identify the
site of the lesion. The sensitivity and specificity of cytology in developing
nations varies widely because of differences in quality of services. For instance,
in Zimbabwe, Pap smear sensitivity for mild dysplasia or worse lesions was
29.6% (95% CI 25.6%-33.8%), but was higher at 78.3% in S. Africa. Hutchinson et al’s study in Costa Rica demonstrated a sensitivity for High Grade
Squamous Intraepithelial Neoplasia of 77.8% (p<0.001). Specificity for Pap
smear in Zimbabwe was 92.3% (90.9-93.6) but was higher at 96.5% in South
Africa. While conventional cytology screening has been shown to reduce
cervical cancer mortality in developed nations, data from randomized controlled
trials in developing nations are lacking.
Liquid Based Cytology

Liquid based cytology (LBC) is another method of cervical cancer screening and uses fluid preservative to maintain cell features. Advantages to this method include improved transfer of cells to slides and readability because it is associated with fewer problems such as air-drying artifact, blood and inflammatory fragments, uneven layering of cells. One study done in Costa Rica found that LBC had a higher sensitivity than conventional cytology, yet data supporting there is no proof that this method is efficacious in reducing cervical cancer in a low income nation. Furthermore the cost of LBC is greater than conventional cytology, requires specific equipment and can thus be prohibitive to programs in developing nations. Testing for HPV DNA in developing nations encounters similar problems to those of LBC in terms of requiring health system infrastructure for sample analyses. The benefit to this method has been its fair to excellent level of sensitivity worldwide in detecting CIN 2-3 and invasive cancers at 61-100% and specificity of 61%-91% as highlighted in Franco et al’s systematic review.

Visual Inspection with Acetic Acid

One inexpensive method that has been implemented in many low resource nations because of the low level of training required is visual inspection with acetic acid (VIA). Performing the test involves application of 3-5% acetic acid to the cervix followed by examination of the cervix by the physician for acetowhite areas near
the transition zone. The transition zone is a circular area of the cervix where columnar cells of the endocervical canal meet with squamous cells of the external cervix. This site is important because cervical cancers are believed to originate at this site. While this method is not specific for cervical dysplasia, neoplastic areas appear more dense and whitened near the transition zone or close to the external os24. Most commonly, results are classified as negative or positive as uniform criteria for reporting have not yet been established internationally24. Because treatment such as cryotherapy may be performed in the same visit, this method bypasses complicated logistics required in diagnosis and treatment by other methods.

International studies have shown that the sensitivity of VIA measured in cross sectional studies ranged from 67-79% and specificity from 49-86% when performed in nations such as India, South Africa, Zimbabwe, and China24,26. Another large good quality clustered randomized controlled trial (n=142,701) done in India showed that VIA was significantly better than HPV testing and cytology in identifying condyloma and CIN1 (p<0.001) and there was not a significant difference in test ability to differentiate between CIN2 or 3 27. Because of the low specificity and fair sensitivity reported in most studies, this screening method lends itself to an increased risk for over-treatment.

Recent research has supported the use of VIA as a tool to reduce the incidence of cervical dysplasia in a screen and treat protocol in a South African randomized
controlled trial comparing efficacy of VIA and HPV DNA testing in relation to efficacy of screen and treat methods. Denny et al (2005) published a good quality study measuring the efficacy of the visual inspection with acetic acid (VIA) method versus HPV DNA testing in the South African township of Khayelitsha. This randomized control trial screened 6,555 non-pregnant women age 35-65 with VIA and HPV DNA testing after which they were randomized to one of three groups: In one group, those patients who were HPV positive were treated with cryotherapy; in a second group, all patients with positive VIA results were treated; in the third group all patients were subject to further evaluation and treatment until six months later. At the six month endpoint, colposcopy was performed on all women by a blinded physician. This physician also performed an endocervical curettage on all women and biopsied all acetowhite lesions. After twelve months all women who had positive HPV or VIA results and a subset of negative tests received repeat colposcopy at twelve months.

The results at six months demonstrated that 0.80% of the HPV DNA group had CIN 2 (95% CI 0.4%-1.2%) and 2.23% of the VIA group had CIN 2 (95% CI 1.57%-2.89%). For the delayed evaluation group 3.55% had CIN 2 (95% CI 2.71%-4.39%). These results were significant as p<0.001 for the HPV DNA group and p=0.02 for the VIA groups. Similar results were found at twelve months where the HPV group had the largest reduction followed by the VIA method and the delayed group had a 5.41% detection prevalence. Overall this study demonstrated that the HPV DNA method was slightly better than the VIA
method in preventing the development of increased dysplasia, yet both methods were efficacious in reducing cervical dysplasia and inhibiting the progression to cervical cancer.

In the Denny et al study, the patients were randomized by a computer generated schedule and patients, health care providers (except for nurses performing cryotherapy treatment) and patients were blinded to treatment strategies. In addition, allocation concealment of the groups hid diagnostic and treatment from the researchers. It was not clear how recruitment of this population took place other than the fact that they came from the township of Khayelitsha. While systematic measurement bias in colposcopic evaluation may have occurred, this issue is inherent in this diagnostic procedure and inter-observer discrepancy may have led to variable results. Furthermore, the size of the study population lends itself to decreased random error. The study is generalizable to women age 35-65 in other developing nations as the study population was large and gave statistical power to the analysis. Of note, the township of Khayelitsha is comprised of 90% black African women and 10% “colored” women thus this study is applicable to similar populations.

Cost-effectiveness for Other Methods

Cost effective analyses comparing VIA method to cytology and HPV testing have been done to evaluate the role this method may play in developing nations. Goldie et al’s estimated cost effectiveness by using a computer generated model
based on a 30 year old black South African woman's risk, incidence, natural history and mortality of cervical cancer. They compared three screening methods in low resource settings and found that HPV testing was usually more effective with a 27% cancer incidence reduction, but was more costly ($39/year of life saved)\textsuperscript{29}. In addition, HPV testing required two visits for screening and treatment. Visual inspection methods were slightly less effective (26% cancer incidence reduction) yet only required one visit and were not only cost effective but cost saving\textsuperscript{29}. Finally cytology was less effective than VIA or HPV testing in low resource settings with a 19% incidence reduction and highest cost a $81/YLS\textsuperscript{29}. Overall, VIA and HPV testing were the best methods for screening according to that analysis. Mandelblatt et al’s cost effective analysis was based on a simulation model of a Thai population and evaluated seven screening methods including VIA, HPV testing, and combinations of same day, multiple day screening and treatment approaches. Their study revealed that VIA in a screen and treat program performed at 5 year intervals in ages 35-55 was the least expensive and saved the most lives and had while HPV screening cost slightly more\textsuperscript{30}. Cost effectiveness studies and efficacy for LBC in reducing the burden of suffering in developing nations remains to be established.

*Overall Efficacy of Screen and Treat Methods*

Screen and treat methods have thus been demonstrated to be efficacious and cost-effective. Furthermore, patients found at least one method acceptable in a study performed in Thailand in which 98.5% of the study population with abnormal
VIA exam accepted immediate treatment within the same visit (n=798). This study reported no complications and a 4.4% return rate of treated women for a perceived problem secondary to treatment. Many of these studies have used colposcopy and colposcopy-directed biopsy as a diagnostic adjunct to verify the efficacy of these methods, yet few studies have evaluated colposcopy as a screening tool in a see and treat approach. This method is currently used in clinical practice in addition to a conventional screening strategy with Pap smear and biopsy in Leogane, Haiti.

*Colposcopy as a primary screening method*

Colposcopy was initiated into US practice in the 1970’s as a diagnostic procedure for women who had abnormal Pap smears, cervical, vaginal or vulvar lesions and colposcopic-guided biopsy remains the gold standard for cervical diagnosis in high income nations. It is not often used in low-resource settings because it requires a skilled technician, and the cost of equipment can be prohibitive to prevention programs in resource poor nations. Colposcopic evaluation is also similar to other visual screening or diagnostic methods in that it can be subjective based on the observer and their skill level. Yet its use as a screening method may be adequate for use in Cervical Cancer Prevention Program for several reasons. Local Haitian physicians have complained that Pap smears have not been sufficient to screen the local population as many Pap results have obscuring inflammation, and may mask accurate pathological diagnosis. In a community whose patients have high risk for loss to follow-up and that has also acquired a
facility, equipment and trained staff to carry out this method, it may provide a satisfactory method for screening and prevention as it allows for same day care.

Mitchell et al performed a meta-analysis in order to characterize colposcopy in the diagnosis of squamous intraepithelial lesions in patients that had abnormal Pap smears. This was done in light of the development of new technologies currently researched in developed nations, such as speculoscopy and fluoroscopy, but have not been implemented in underdeveloped nations at this time. They pooled data from nine studies and noted that in the US and Europe, the study with the lowest sensitivity was 87.0% (+/-5%) while the study with the highest sensitivity had results of 99% (+/-6%). The lowest specificity was 23% (+/-6%) and the highest was 87% (+/-1%). Mitchell et al found that the specificity increased slightly when only distinguishing between high grade dysplasia and cancer from lower grade lesions. Because of high sensitivity and low specificity of colposcopy in a screen and treat protocol for all cervical lesions, this method lends itself to a risk of over treating patients. Because this meta-analysis demonstrated that the weighted average for sensitivity was higher than that for VIA, and specificity was higher than most results for VIA, colposcopy is generally a more accurate method for screening. This suggests that this tool is better suited for a population with higher prevalence of more advanced lesions.

Schneider et al performed a study with 5,455 women in Germany comparing screening methods for high grade dysplasia (CIN2/3) and cancer including Pap
smear, HPV DNA testing, and screening colposcopy using biopsy as the gold standard diagnostic tool. They found that the corrected sensitivity for colposcopy with 95% confidence intervals was 13.3 (7.0-20.5) and specificity was 99.3 (99.0-99.6). The positive predictive value (PPV) for colposcopy was 38.1 (21.3-58.3) and the negative predictive value (NPV) was 97.3 (96.7-97.8). The sensitivity for cytology was slightly higher at 20.0 (10.8-28.6) and specificity was 99.2% (98.8-99.5). Cytology had a PPV of 70.6 (41.1-95.3) and an NPV of 97.5 (96.9-98.0). Overall, traditional cytology was able to identify patients with disease better than screening colposcopy, although the ability to identify patients that were negative for CIN II or III was very close. Screening colposcopy had a higher PPV than cytology, however, and the NPV for each was similar. This study was of fair to good quality as patients were not randomized and therefore only those with positive cytology and/or colposcopy screening results received biopsy. Because not all patients received biopsy, the authors adjusted prevalence estimates using the prevalence of cancer in the biopsy group and applied an incident rate to negative women based on an eight month follow up period. While internal validity is good, it is not known whether these results would apply to populations in low resource settings as this study took place in private gynecology practices in Germany.

A study done by Wu et al (2005) in China compared the efficacy of colposcopy, HPV DNA testing, liquid based cytology, and visual inspection with acetic acid in identifying any squamous intraepithelial lesions and worse lesions. The study
population included 450 women age 20-74 years of age without history of cervical biopsy, all of whom received VIA and HPV testing. Only 273 underwent all four screening procedures “because of the good results of visual inspection”19. VIA had the lowest sensitivity (38.9%) and specificity (68.5%) while colposcopy was the second lowest with a sensitivity of 55.6% and specificity of 79.4%. These values are low for a screening test. Liquid based cytology had a higher sensitivity than both methods with 77.2% sensitivity and 98.63% specificity. This study was of fair quality because it was not randomized (patients chose which methods they did not want to receive), blinding was not described in the article and no reference to the colposcopist’s experience was given. Furthermore they did not discuss whether the results were significant or not, and the meanings behind this. They concluded that because of the high levels of false negative tests from VIA and colposcopy for patients with low grade dysplasia, neither should be used in the early detection of cervical dysplasia19.

Massad and Collins (2003) evaluated the correlation of colposcopic impression with biopsy results in a study of 2112 women with both colposcopy and biopsy results35. They found that the association between colposcopy results and biopsy pathology was significant (p<0.001), although the kappa statistic suggested a weak correlation of 0.2 between results. They found that there was exact agreement in the grade of cytological lesion in only 37% of cases, although this increased to 75% when the colposcopy result lay within one histological grade of the biopsy result. The sensitivity of colposcopy in detecting lesions worse than
CIN II was 56% and specificity for the detection of atypia or worse was 52%. They concluded that while colposcopy may assist in estimating the grade of cervical lesions, it is not precise and increases risk of over diagnosis and overtreatment of patients when used alone in screening. The Massad and Collins study was of good quality overall. This study was not randomized as performing biopsy without indication would be unethical. The patients were referred after having abnormal Pap smears, and thus were at higher risk than an average individual and the residents performing the examination were not blinded to this information. Measurement bias may have been introduced as the physicians already had a high index of suspicion. Inter-observer variability may be at issue as different residents performed colposcopy. This study is generalizable to urban populations with a high proportion of women of African American and Hispanic descent and an average age of 33 years.

Finally, Cantor et al (1998) did a cost effectiveness study comparing new screening methods (speculoscopy) to colposcopy in traditional screening methods and see and treat methods. There has not been a cost effective analysis comparing other see and treat methods to see and treat colposcopy. Cantor et al used data from nine studies and their clinic to generate probabilities of true positives and negatives and false positives and negatives using weighted average of prevalence from each referral population (all based in high resource nations). Colposcopy had a high sensitivity at 94% and low specificity at 48% in distinguishing between presence versus absence of cervical dysplasia.
distinguishing between high grade versus low grade dysplasia, the sensitivity dropped to 79% and specificity to 76% however. In order to approximate cost, Cantor et al used charge data from one US based hospital system and compared it to cost per high grade lesion detected. They found that colposcopic directed biopsy was most expensive (this requires 2-3 visits from the patient) at $311,808 to detect 45.78 cases. When a see and treat protocol was used 46.05 cases were found per $285,133. Colposcopy dominated spectroscopy in cases found per cost in both screening protocols. This analysis was based on US health system costs and did not incorporate a societal perspective of benefits and costs (such as life years gained). This study further supports the ability of colposcopy to identify presence versus absence of cervical dysplasia, but inability to distinguish between high and low grade dysplasia in developed nations. Cost effectiveness must also be determined in low resource countries prior to discarding this method.

Two methods have been shown to decrease prevalence of dysplasia in a see and treat protocol in low resource settings, yet the large ranges in colposcopic sensitivity and specificity found in other studies introduces doubt as to the accuracy of this method. Prior studies examining accuracy have varied widely for many reasons including differences in the characteristics of study populations, variability of colposcopy skills, presence in high resource settings and more. This study will characterize the accuracy of colposcopy in Haiti, and these results will determine whether this method is adequate for a see and treat protocol using colposcopy as the primary method.
Methods

Study Questions and Hypotheses

Reports on screening and diagnostic accuracy of screening colposcopy have varied significantly in the literature. Medical records data from the Leogane, Haiti Cervix Clinic were used to characterize the accuracy of these methods at that specific clinical site through the examination of two questions:

Analysis 1: How accurately does the US gold standard screening method, Pap smear, identify mild dysplasia (CINI) or worse lesions as compared to gold standard diagnostic testing with biopsy?

Analysis 2: How accurately does screening colposcopy identify mild dysplasia or worse lesions as compared to diagnostic biopsy?

Ho: Colposcopy will perform with less sensitivity and specificity than Pap smears when comparing each to the gold standard, biopsy.

Alternative hypothesis: Colposcopy will perform the same as Pap smear in sensitivity and specificity analysis when compared to the gold standard, biopsy.

Study Populations

Recruitment to the Leogane clinic occurred via word of mouth from patients who visited the clinic, and often included women with various vaginal symptoms seeking acute care. This study includes three different accuracy analyses and therefore three final study groups were created based on which clinical
evaluations were received by the source population. Inclusion into at least one of the two study populations required that a patient have a minimum of two clinical procedures, including Pap smear, colposcopy and/or cervical biopsy. See Figure 1 for a visual depiction of the study populations.

Clinical Protocol and Data Collection

The clinic nurse gathered demographic data via interview with each new patient. The physician visually examined the vulva, vagina and cervix of each patient. The physician performed Pap smear and colposcopy on each patient and biopsy if indicated. Two different protocols were instituted over time, the first of which took place from 01/2000 through 08/2005. During this phase, the physician would not perform cervical screening tests until after the administration of treatment to all women that had vaginal or cervical infections. Loss to follow up was high and it is believed that patients did not return to the clinic if their symptoms had resolved, or other barriers intervened such as finances. From 08/05 until 12/05 the predominant strategy was to perform screening or diagnostic evaluation with Pap smear at the initial visit regardless of whether or not there was cervical or vaginal infection. All of these patients were required to return to the clinic in order to receive and pay for the Pap results and also to receive colposcopy.

After the initial interview and physical examination, the women underwent different combinations of the three procedures: Pap smear, colposcopy, and
cervical biopsy, depending on the predominant protocol and patient behaviors. Pap smear was performed by rotating a wooden stick at the opening of the cervical os and smeared the sample on a glass slide. Thereafter the physician rotated a cytobrush in the internal cervical os smeared this sample on a clean section of the glass slide. The slide was then sprayed with a fixer. The physician was blinded from Pap results prior to colposcopy as the cytology lab withheld Pap results in a sealed envelope.

The physician performed colposcopy after applying 3% acetic acid to the cervix and made a colposcopic diagnosis and recorded the results. He opened the envelope with Pap smear results for women that returned to the clinic from a previous visit to ensure that a lesion was not missed on a normal appearing cervix. Cervical biopsy and/or endocervical curettage was performed if the physician identified dysplastic-appearing lesions and cervical samples were placed in 10% neutral buffered formalyn.

In order to receive cryotherapy as treatment for cervical dysplasia the physician required patients to return to the clinic for the biopsy results. The nurse recorded all data obtained into a computer database in Filemaker Pro. The clinical practice in which this medical records data was gathered rarely practices a screen and treat approach. Patients who were certain that they could not return were offered cryotherapy prior to the return of biopsy results. The data is being analyzed to evaluate the potential effectiveness of a single day screen and treat approach.
A total of 2106 women attended the clinic between January of 2001 and December of 2005 and were ages 19-80 from Leogane and surrounding regions. Of the 2106 women, 82 women had no record of cervical screening and they were excluded leaving a total of 2024 women with at least one documented evaluation (See Group A, Figure 1). From these 2024 women, a total of 23 women did not have a Pap smear as part of their evaluation. This meant that 23 women had results available only from colposcopy and biopsy (See Group B, Figure 1). Two reasons for missing results include failure of patients to take slides or biopsy results to the laboratory, failure of patients to return to the clinic for lab results, and finally the clinic ran out of supplies for certain tests at different times.

A total of 2001 women underwent a Pap smear as part of their evaluation. Of these, 158 women had a Pap smear as part of the one day screening protocol and 20 were excluded because they had only a Pap smear and no further evaluation. A total of 127 had both Pap smear and colposcopy while 11 (Group C) women had Pap smear, colposcopy and biopsy during the second clinical screening protocol.

From the 2001 women that had Pap smears, 1843 did not undergo the one day screening protocol and were asked to return. A total of 689 women were excluded from this analysis because they did not return after the Pap smear results
and colposcopy. This means that 34% of the women that were asked to return to receive Pap smear results did not return for their second visit.

After initial evaluation with Pap smear, 1,154 returned for further evaluation. The remaining women were included in at least one of the two analyses performed in this study. A total of 187 women had Pap smear, colposcopy and cervical biopsy, allowing for a comparison of results from each of those tests (Group D, Figure 1), while 967 women had Pap smear and colposcopy only.

The final study group for the first analysis included 198 women and will be used to calculate diagnostic sensitivity and specificity of Pap smear against biopsy, (diagnostic gold standard for cervical cancer). The second analysis had a total of 221 women with colposcopy and biopsy results, for which and sensitivity and specificity of colposcopy will be assessed by comparing it to biopsy as the gold standard. A total of 1,315 women had at least two evaluations performed. See Table 2 to view the final study populations for each analysis.

Variables:

Pap Smear Results

Cytological results of Pap smears obtained from the clinic’s laboratory were graded independently for inflammation and cellular atypia. Cellular atypia was graded as normal, mild moderate or severe inflammation, mild dysplasia (CIN I), moderate dysplasia (CIN II), severe dysplasia (CIN III) or carcinoma. All patients
with any level of dysplasia or carcinoma were considered to have a positive Pap
test. In 2000, Pap specimens were sent to a Pittsburgh laboratory for validation.
The lab performed independent cytological evaluation and found that results
obtained by the Leogane lab were comparable (data not available)*.

Colposcopy Results
The diagnostic result used for see and treat colposcopy is a “colposcopic
impression”, or preliminary diagnosis that a Haitian physician records after
viewing the cervix through a colposcope. This includes normal, infection,
inflammation, mild dysplasia, moderate dysplasia, indeterminate grade dysplasia,
and cancer. For this study, dysplasia or cancer constitutes a positive result while
infection, inflammation and normal constitute a negative result.

Biopsy Results
The specimens for biopsy results were taken from the endocervical canal, the
outer cervix or both. The highest grade lesion seen on biopsy was recorded as the
final diagnosis. Results included normal cervix, chronic cervicitis, HPV change,
mild dysplasia, moderate dysplasia, severe dysplasia, microinvasive cancer and
carcinoma. A positive result includes any level of dysplasia, or any stage of
carcinoma. Pathology services were provided by the Hopital St. Croix of

* Validity of Pap smear verified through personal communication with Dr. David Walmer,
Chairman of FHM on 05/15/06 and 06/20/06. He reports that this evaluation was performed at
Mercy Hospital of Pittsburgh, PA by Dr. Rosemary Edwards.
Leogane and were also validated for accuracy by an independent laboratory. For categorization of other variables, see Table 3.

Patient Age
A total of 1309 women reported their age at the time of their initial evaluation. Women were divided into four age groups: 19-30, 31-40, 41-50 and 51-80. The cut-offs for these groups was based on the recommendations of the World Health Organization for cervical cancer screening in areas where resources are limited. The main cervical cancer prevention documents state that screening is not recommended before age 30 as its prevalence is rare in this group of women. Current research indicates that women should receive at least one screening examination during the third, fourth and fifth decades of life. Other covariates assessed in the descriptive analyses include age of first coitus, number of lifetime sex partners, marital status, gravidity, current irregular menstrual cycle, presence of postcoital bleeding. These data were obtained from the medical records. Refer to Table 3 for specific categories.

Se and Sp Analysis
In order to evaluate screening and diagnostic accuracy of Pap smears, and see and treat colposcopy in this low resource setting, I performed two analyses. These analyses address the diagnostic accuracy as sensitivity and specificity by comparing each screening test to cervical biopsy as the diagnostic gold standard.

† Validity of Pap smear verified through personal communication with Dr. David Wahner, Chairman of FHM on 05/15/06 and 06/20/06. He reports that this evaluation was performed at Mercy Hospital of Pittsburgh, PA by Dr. Rosemary Edwards.
Both analyses involved calculation of sensitivity, specificity, positive predictive values and negative predictive values. The data was stripped of all identifiers and the Institutional Review Boards of UNC-Chapel Hill and Duke University determined that the study was exempt from IRB review\textsuperscript{2}.

I used Stata version 8.0 for all descriptive statistics and to create a bivariate two by two table for each analysis. The bivariate tables describe the number of patients that were positive or negative for disease and the test’s ability to correctly identify this in each analysis. Groups were labeled as a, b, c, d and corresponded to (See Tables 4A and 4B to view organization of bivariate tables):

\begin{align*}
  a &= \text{the numbers of true positives (diseased subjects with correct positive test results)} \\
  b &= \text{false negatives (disease, but negative test)} \\
  c &= \text{false positives (no disease, but positive test)} \\
  d &= \text{true negatives (no disease, negative test)}
\end{align*}

Sensitivity, specificity, positive predictive value, and negative predictive values were also calculated by Stata. Sensitivity was calculated as the proportion of women for whom the test correctly identified as positive for disease divided by the number of women that were correctly identified as positive plus those that tested negative but truly had the disease \[\frac{a}{a+b}\]. Specificity was determined as the proportion of women that were correctly identified as negative for disease divided by the number of women that were truly negative plus those that were

\textsuperscript{2} University of Chapel Hill study #06-0172
Duke University approval given by Dr. John Palletta on 01/31/06
falsely positive for the test \( \frac{d}{c+d} \). Statistical significance was calculated as 95% confidence intervals by Stata using the Wilson score method.

Positive predictive value (PPV) is the proportion of test positives that were correct \( \frac{a}{a+c} \) while PPN was the proportion to test negatives that was correct \( \frac{d}{b+d} \).

These calculations were calculated for each screening test as follows:

\[
PPV = \frac{\text{Sensitivity} \times \text{Prevalence}}{\text{Sensitivity} \times \text{Prevalence} + (1-\text{Sensitivity}) \times (1-\text{Prevalence})}
\]

\[
PPN = \frac{\text{Specificity} \times (1-\text{Prevalence})}{\text{Specificity} \times (1-\text{Prevalence}) + (1-\text{Specificity}) \times \text{Prevalence}}
\]

The prevalence of cervical cancer was calculated by Stata and based on the presence of carcinoma from biopsy results in this study population.

**Analysis One**

The first analysis addressed the issue of whether Pap smears can adequately diagnose cervical intraepithelial neoplasia grade I (CIN I) or worse lesions compared to biopsy from the 150 women that had both Pap smear and cervical biopsy. I performed this analysis by calculating the sensitivity and specificity of Pap smears against the diagnostic gold standard of cervical biopsy. Sensitivity, specificity, PPV and PPN with 95% confidence intervals were calculated by Stata using biopsy results to verify the true presence or absence of disease.
Analysis Two

The second analysis evaluates the ability of screening colposcopy to identify CIN I or worse lesions from the 145 women with results for both of these procedures. For this analysis, I compared colposcopy results to the results of the diagnostic gold standard, cervical biopsy. Sensitivity, specificity, PPV and PPN were calculated as described above with 95% confidence intervals for each of these calculations.

Results

Characteristics of the Study Populations

A total of 2,094 individual women (See Group A, Figure 1) had at least one analysis performed and data recorded, although there was missing data for certain variables in the descriptive statistics section and this is noted for each variable in Table 3. Missing data was reported for each variable and not included in the calculation of percentages. The average age for this entire population was 42.5 years (standard deviation/SD=10.5) ranging from 19 through 80 years. The predominant age categories for the group A population women was age 31-40 with 32% (n=422) of the population, and the 41-50 age group with 32% (n=419) of the population.

The majority of women in all populations, 59.8% (n=787), reported only one lifetime sexual partner and a very small proportion, 2.4% (n=31), of the combined population reported having more than 5 lifetime sex partners. This was similar to
the populations for each of the analyses (Table 3). Most women in all population groupings reported either living with their partner or being married, while fewer women were part of the single and/or divorced category. Few women (9.7%, n=127) from group A study reported that they had never been pregnant. Analysis 1 (groups C+D) and analysis 2 (groups B+C+D) had even fewer women reporting a gravidity of zero at 4.5% (n=9) and 4.1% (n=9) respectively. Most women reported a gravidity of 1-4. Refer to Table 3 for additional descriptive variables.

**Prevalence of Cancer by Diagnostic Test**

A total of 24.4% (n=54) women from analysis 2 (groups B+C+D) had gross cancer that only required visual inspection for diagnosis. This contrasts with the 48% (n=106) of women from the same group with dysplasia or cancer on biopsy, the gold standard. I included two different combinations of categories for Pap smear results. The first categorization included groupings for normal, abnormal but not dysplasia, dysplasia, and cancer. For these results, very few women had normal results on Pap in any of the groups, with the highest proportion at 2.0% for group A. A very large proportion of women had abnormal findings on Pap smear but not dysplasia (meaning inflammation/infection) for all analysis groups, including 93.2% (1207) of the women from group A.

A total of 13.5% (n=177) women had cancer or dysplasia on colposcopy for group A. Analysis 1 and 2 had high prevalence of cancer or dysplasia with colposcopy ranging from 66.7% (n=132) to 70.1% (n=155). Finally, the results for biopsy
were only relevant for the populations that had this evaluation performed.

Analysis 1 showed that 41.9% (n=83) of the women had dysplasia or cancer while analysis 2 had a percentage of 48.0% (n=106) with dysplasia or cancer on biopsy (Table 3).

Sensitivity and Specificity of Pap Smear vs. Biopsy

Analysis 1 entailed the comparison of Pap smear results to biopsy results from groups C+D (See Figure 1). The bivariate analyses with positive test results for the diagnostic test and test variables are included in Table 4A. The sensitivity of Pap smear compared to biopsy was 68.7% (CI 57.6-78.4) and specificity for Pap smear compared to biopsy was 96.5% (CI 91.0-99.0). Of those screened using Pap, 93.4% (CI 84.0-98.0) truly had disease whereas 81.0% (CI 73.0-87.0) correctly screened negative for disease with Pap.

Sensitivity and Specificity of Colposcopy vs. Biopsy

Analysis 2 included groups B+C+D in the study population and colposcopy and biopsy results were compared. The bivariate table with the number of test positives and negatives is available in Table 4B. Colposcopy was able to identify more cases than Pap with a sensitivity of 96.2% (CI 91.0-99.0) but colposcopy had less ability to rule out cases with a specificity of 53.9% (CI 44.0-63.0). Colposcopy correctly identified 65.8% (CI 58.0-73.0) women with disease while the negative predictive value was 81.0% (73.0-87.0) (Table 5).
Discussion

The US and Europe have significantly reduced the morbidity and mortality of cervical cancer programs that use serial Pap smear, colposcopy and biopsy as the major tools of screening in their respective environments\textsuperscript{11,34}. These methods have not been widely available in low resource settings and clinicians have encountered other difficulties in their use, such as the need for multiple visits from patients for screening, diagnostics and treating women. See and treat methods have gained popularity and validity in the past decade, although the tools for screening may be lacking in accuracy. This study focused on the determining the accuracy of screening colposcopy, as this method may be applied in a screen and treat protocol in a low resource program in Haiti. This method would require fewer laboratory services, and less frequent follow up appointments from women that experience difficulty returning to the clinic. For example, 34% of the women attending the clinic that had an initial Pap smear did not return for their results or further evaluation.

Summary of Key Findings

In this study, the prevalence of disease was not statistically different for the Pap smear +biopsy population than the colposcopy+ biopsy population. The results for sensitivity of Pap were low while specificity was good. Colposcopy was able to identify more cases than Pap, with a significantly higher sensitivity, although colposcopy had a significantly lower specificity.
The positive predictive value (PPV) of Pap was high with fairly narrow confidence intervals. The PPV of colposcopy demonstrated a fair percentage of correctly identified patients with dysplasia or cancer at 65.8% (CI 0.58-0.73), which is significantly different from the higher PPV for Pap smear (93.4% (0.84-0.98). The negative predictive value (PPN) for Pap and colposcopy was good and the results did not significantly differ, as the confidence intervals overlapped.

Interpretation of Findings

We expected that colposcopy would perform with less sensitivity and specificity than Pap smear when comparing each to the gold standard, biopsy. This was incorrect in the case of sensitivity, however, as colposcopy was able to identify positive cases with much higher capability than Pap smear. Our original hypotheses also predicted that colposcopy could not rule out disease as well as Pap smear, and this proved to be true. In addition, Pap smear was able to distinguish positive cases that actually had disease significantly better than colposcopy could. This suggests that Pap smear misses more disease while colposcopy is better in identifying cases. Both tests were able to identify women that were truly negative for disease at similar levels.

Current knowledge on the accuracy of screening colposcopy has shown mixed results depending on multiple factors determined by the setting in which studies are carried out. The Schneider et al (2000) study was performed in Germany, where screening colposcopy is used in standard practice as a component of
cervical evaluation and screening\textsuperscript{34}. They found that screening colposcopy had a very low sensitivity at 13.2 (95% CI 7.3-19.6) for moderate dysplasia or worse lesions. They also found that the specificity of screening colposcopy was good at 99.2% (95% CI 98.9-99.4). The sensitivity for cytology was still poor for Pap smear, and not significantly higher at 18.4 (11.6–25.7) and there was no significant difference in Pap smear specificity, which was 99.0% (98.7-99.3).

The results of my study demonstrated the opposite effect from the Schneider study, in which sensitivity for colposcopy was significantly higher than Pap at 96.5% (CI 0.91-0.99) while the specificity was fair, although significantly lower than Pap at 53.9% (0.44-0.63). The PPV for colposcopy in the Schneider study was poor at 28.8% (95% CI 17.0-41.5) and the PPN was good at 97.9 (CI 97.5-98.3). This study found that the PPV for colposcopy was significantly higher than colposcopy in Schneider et al’s study, where the PPV was 65.9% (CI 0.58-0.73) and a PPN of 93.7% (0.85-0.98). Overall, the results for my study reveal a better sensitivity and positive predictive value for colposcopy, worse specificity, while the negative predictive value was not significantly different than that found in the Schneider study. Also, my results for Pap smear sensitivity were a minimum of 3.18 times higher than found in that study, while specificity for Pap was high and not significantly different.

Various issues may contribute to the differences found between these two studies. The populations screened in each of these studies demonstrate different
background characteristics such as age. The Schneider et al study reported a median age of 35, whereas the mean age for my study was 44.5 (SD 10.5). In addition, the prevalence of CIN II, CIN III and cancer on biopsy in the German study population was much lower at 2.4%. The difference in age may have affected the prevalence of disease as the highest incidence of cervical cancer is found between the ages of 30 and 60. The prevalence of dysplasia or cancer on biopsy for the larger study population (analysis 2) in my study also included CIN I and was 48.0%, and therefore including another level of dysplasia would also presumably increase prevalence in my study. Despite the inclusion of CIN I patients in my study, the huge difference in prevalence suggests that dysplasia and cancers are more prevalent in the Haitian population and this may contribute to the higher sensitivity of colposcopy in this study. Finally, the study size of the Schneider study included 5,455 women and therefore the results are likely a closer estimation to the truth, whereas my study populations around 200 women may have allowed for sufficient random error to make the results inaccurate.

The study done by Wu et al (2005) demonstrated a sensitivity of colposcopy was 55.6%, which is 35.4 percentage points less than the lowest end of the confidence interval sensitivity I calculated in my study. The specificity for colposcopy in the Wu study was significantly higher than what I found in my study at 79.45%. The Youden index score for colposcopy was 0.35, which suggests that this screening method was not ideal in the setting in which their study was carried out. ThinPrep cytology was more accurate than conventional Pap with a 72.2% sensitivity, a
98.6% specificity and Youden index of 0.7, which was higher than that for colposcopy\textsuperscript{19}. Overall, their cytology test did not perform significantly different from Pap smear sensitivity and specificity in my study. The sensitivity of colposcopy in my study was significantly higher than their results while specificity between the studies was not significantly different. Colposcopy use in Haiti is able to identify more cases than the Wu study, yet the ability to identify patients without disease was similar between studies. Thus colposcopy as used in Haiti was an overall better screening test than in the Wu et al. study.

One disadvantage to comparing the results of the Wu et al study, is the fact that they did not calculate confidence intervals for their diagnostic accuracy values. In this sense it is impossible to truly determine whether the range of confidence intervals were significantly different from their study, although it seems as though the differences are large enough that intervals would not overlap in most cases. Both studies had fairly small populations as well and thus may be subject to random error more than large study populations.

The results of my study fall into the range of the results found in the meta-analysis for colposcopy performed by Mitchel et al\textsuperscript{20}. They reported that the sensitivity of colposcopy ranged from 64 to 99% and the specificity from 30 to 93%. My sensitivity result was at the higher end of their range, and my specificity was at the lower end of the range for specificity of colposcopy values from that meta-analysis. This supports that the accuracy colposcopy in Haiti falls
into the known range of results. Additionally, Mitchell's study included studies of high prevalence as only women referred for colposcopy because of abnormal Pap results and therefore prevalence of disease would be higher in their population. The study populations in those studies are similar to my study population in having a high prevalence of dysplasia and/or cancer.

The results of the Massad and Collins study may reflect underlying causes for the fair specificity found in my study. They reported that colposcopy had poor capability to predict the exact result of biopsy, however results within one histologic grade agreed in 75% of cases. The sensitivity for finding CIN II/III lesions for their study was 56% while specificity to rule out any atypia or worse lesions was 52%. The sensitivity for my study was significantly higher, while specificity was not significantly different than their study. This suggests that screening colposcopy as used in Haiti is a better screening tool than in urban areas in the US. This may be attributed to the higher prevalence of disease in Haiti. The sensitivity in the Massad and Collins study may be lower because they did not include low-grade dysplasia. Their study population had a lower mean age of 33.5 years which is a decade younger than the mean age for my populations. Although they state that their population largely included a higher risk group with African America and Hispanic women, the prevalence of cervical cancer and dysplasia even in these populations is not as high as it was in the Haitian population.
This analysis demonstrates that colposcopy is carried out with high levels of sensitivity and specificity in Haiti. In a population with a high incidence of cervical cancer and dysplasia it is important to use a screening method that is able to identify a high proportion of disease, while also retaining ability to rule out disease. The use of colposcopy in Haiti does not have a high level of specificity, which means that incorporation into a one day screen and treat protocol would lead to unnecessary treatment for a proportion of women. This issue must be carefully weighed and evaluated as disease prevalence and mortality from cervical cancer in Haiti is high. A treatment strategy that minimizes morbidity in patients without disease may be a small cost with the large benefit of saving a large proportion of lives. The high loss to follow up (34%) and the high prevalence of dysplasia in the Haitian population speak to the need for a screen and treat method. The low sensitivity of Pap smear in this population signals the need for a test that identifies cases more accurately. Colposcopy is a good candidate that fulfils these requirements; however one must consider the morbidity associated with any treatment method incorporated into a one day protocol. Cryotherapy is a viable option over conization or Loop Electrosurgical Excision Procedure (LEEP) as several studies have demonstrated little effect on pregnancy outcomes in women that have undergone this treatment$^{37,38}$. One issue with this treatment modality lies in the high percentage of women with inflammation on Pap smear in this Haitian population (93.2%). While one study performed in Greece demonstrated that 86% of women with chronic cervicitis treated for dysplasia with cryotherapy made full recoveries, environmental factors such as poverty may
influence recovery differently in the Haitian population\textsuperscript{39}. This suggests that colposcopy may be used in conjunction with a screening method that has high specificity for cervical dysplasia, thus preventing over treatment and the risks that may arise therein.

\textit{Limitations of This Study}

One limitation of this study lies in the study design. Because this study was based on medical records from routine clinical care and not a designed scientific study, randomization to various methods of cervical evaluation was not possible, or ethical. The likely effect of this was probably increased selection of the women that received biopsies because of a higher index of suspicion after abnormal Pap and/or colposcopy results. For example women with more obvious clinical signs may have been more likely to receive biopsy than those with mild dysplasia. This would increase the prevalence of disease and also increase the value for sensitivity. Also, because of the study design, I can only hypothesize that this screening tool would work in a see and treat approach. Basic accuracy data on sensitivity and specificity does not necessarily translate into the effectiveness of a screening program to decrease mortality from cervical cancer.

A second limitation was that this study only delineates between all levels of dysplasia versus normal/abnormal. Because mild dysplasia (CIN I) often resolves, other studies include only moderate dysplasia (CIN II) or worse lesions as a “positive” test result. Thus, this study would have a higher prevalence of
disease and lower threshold for diagnosis. If these accuracy results were to be used in a see and treat protocol, there is a risk of over-treating the population.

A third limitation is that bias may have been introduced in several areas. As noted previously, selection of the biopsy populations may have occurred due to a combination of interacting factors. These factors may have lead to selection bias in terms of which women from the larger Haitian population attended the clinic. Because of the large number of abnormal Pap results, these patients may have had infection or symptoms that lead them to the only gynecologist in the area. As such, these patients may have different risk factors from other women in low resource areas.

A fourth limitation is that all demographic information was given by patient report and the questions asked by the nurse are not standardized or validated. Upon visiting the clinic I discovered that there may be more than one patient in the room during a patient's clinic appointment and it is possible that patients answer sensitive questions based on the expected or moral answers. This may be the case with questions regarding the number of lifetime sexual partners, for instance.

Finally, there may be an issue with the nurse's understanding of some of the questions that she asked and this was apparent in the gravidity and parity statistics. For example, in the combined study population 53.1% of the women
reported 1-4 pregnancies whereas 60.4% reported that they had 1-4 live births. It is impossible to have more live births than pregnancies unless there are problems with the question or its comprehension. This highlights the importance and difficulty of working across languages but signals an important issue that can now be addressed in future data collection.

**Strengths of This Study**

There were, however, several strengths in this study. This study was important to provide feedback on the current state of cervical cancer screening in Haiti. While it does not establish a screening method that is better than Pap overall, it does at least establish a role for colposcopy in this clinic. This will provide a foundation from which future development of methods may be developed and improved upon. For instance, colposcopy was more sensitive in identifying cases, although Pap was able to rule out disease with more accuracy. One reason that Pap smear may have such a low sensitivity in this setting may lie in the fact that there was a large population of women that had abnormal results due to inflammation or infection. Haitian physicians have reported the difficulty of using Pap smear previously as many results are clouded by obscuring inflammation. This was evident in the study population characteristics presented in Table 3, where the Pap smear and biopsy group (analysis 1) demonstrated that 68.2% had abnormal results (not dysplasia or cancer) on Pap smear. Alternatively, a problem with correct cytological diagnostics may be the cause of this problem yet the fact that these services exist at all is a benefit available in very few areas in Haiti. This
problem further supports the need for a screening program that can identify cases better than Pap smear in this population of women.

Another strength of this study was the fact that that the clinical protocol and data collection for this study were carried out with good standards in a low resource setting. The use of File Maker Pro to record results was likely to ensure that the demographic data as well as test results were correct for each patient in this dataset. Furthermore, because these were the medical records used daily by the Haitian staff, it was in their interest to ensure correct information for good patient care. The tools used to evaluate cervical findings were the best available in the community in which these methods were carried out, and far above the standard of care for the nation.

A third strength of this study was that measurement bias was limited in at least one respect. Inter-observer variation in the evaluation results was controlled as the same physician determined all of the colposcopy results, the same cytologist diagnosed all Pap smears and the same pathologist determined all biopsy results. On the other hand, it is difficult to ensure that results for pathology services were valid. Although Pap smear results were reported to be validated by a Pittsburg laboratory at a previous date, the fact that the results of that analysis are not available also puts the validity of Pap smear results into question. These are the tools that are available in this setting, however, and their use and evaluation is important in any studies seeking to analyze screening methods in Leogane.
The population included in the study demonstrated certain characteristics that increase the risk of cervical cancer. The mean age of the population was typical of that worldwide at 44.5 (standard deviation 10.5) as it is the third, fourth, and fifth decades in which cervical cancer incidence is highest\(^3\). The majority of women reported only one lifetime sexual partner, for instance 59.8\% (n=787/1315) of the combined study population reported this, while the majority of the remainder of the populations reported having 2-4 lifetime sex partners. Despite this, the prevalence of dysplasia or cancer in this population appears to be high, for instance the prevalence in the Pap smear and biopsy group (analysis 1) was 4.8\% (n=62). This suggests that other factors may contribute to the development of cervical cancer in this population and future research may investigate the association of population specific risk factors with development of cervical cancer. More research is required to determine what these risk factors are and how they may be different from other populations in low resource settings.

**Public Health Implications**

This study is important in the consideration of cervical cancer screening methods in low resource nations. Because of challenges presented to both clinicians and patients in these settings, a method that decreases the number of follow-up appointments while also sufficiently decreasing the number of cases is desired. For this reason many researchers have been working toward the development of see and treat methods internationally. Using colposcopy in Haiti seemed to be
appropriate as a local program had acquired the clinical knowledge and skill as well as resources to carry this method out along with the use of routine screening and diagnostic procedures. Because it is a relatively rapid screening process, and results are immediately available, it would allow for treatment of cervical lesions on the same day a patient attended her first visit. While this test demonstrated an excellent ability to identify cases, its low specificity would not rule disease out in many women that did not have dysplasia. This means that this test has the potential to decrease disease in this population however the costs of over treatment remain to be determined. Future study should incorporate a test with high specificity such as Pap smear or HPV DNA testing to prevent treatment in patients without disease. Simultaneously, morbidities such as continued inflammation and/or infection should be recorded from women with true disease for whom treatment is indicated.

In this study I also characterized the accuracy of the Pap smear as used in Leogane and established a baseline of the accuracy this screening method provides. At present, the sensitivity is fair and the specificity was excellent. The overall results found herein emphasize that other methods need to be explored to address the issue of prevention of cervical cancer in this and other low-resource settings as disease is not reliably identified by this screening method. Other methods including colposcopy, HPV DNA testing or the use of Visual Inspection with Acetic acid appear to have adequate accuracy while also addressing issues of cost and difficulty of follow-up.
Conclusions

The use of colposcopy over Pap smear for cervical screening in Leogane, Haiti remains a viable option because of many characteristics of the test that would allow for easy treatment. The accuracy of the test itself as used in the Cervix Clinic was significantly higher than Pap smear yet specificity was lower than Pap at 53.9% (0.44-0.63). These findings fit within the ranges of results of most other studies performed in high-resource settings with this screening method. At the same time, Pap smear accuracy was not good, although better than results found in Germany, Zimbabwe, and other nations with a sensitivity of 68.7% (0.58-0.78) and specificity of 96.5% (0.91-0.99). These results demonstrate that the worldwide screening gold standard varies by prevalence of disease and does not yield good accuracy at sites where resources are limited. Importantly, the prevalence of cervical cancer among patients with biopsies in this study ranged from 41.9% (n=198) to 48.0% (n=221), which is extremely high. The prevalence of cancers and dysplasia detected when Pap smear is assumed to be the gold standard was 4.81% (0.04-0.06) (n=2090), and this prevalence is still high enough to merit serious revamping of a cervical screening program. Colposcopy may serve in a screening program in two capacities: 1. It may simply aid a physician in the estimation of the grade of a cervical lesion in conjunction with another screening method because of its poor ability to rule out disease 2. It may be used as the primary screening tool in a screen and treat protocol if the benefits in lives saved outweighs the costs incurred by the treatment method.
Reference List


Figure 1  Study Population Selection Process  
Cervical Cancer Prevention Program  
Leogane, Haiti 2000-2005

Attended Cervix Clinic \( n = 2106 \)

\[ \Downarrow \] \[ \text{No evaluation recorded } n = 82 \]

\( \text{Group A} \)
Have an Evaluation Recorded  
\( n = 2024 \)

\[ \Downarrow \]

Had Pap smear as Part of Evaluation  
\( N = 2001 \)

\[ \Downarrow \]

Asked to Return for Further Evaluation  
\( n = 1843 \)

\[ \Downarrow \]

Did not Return  
Had pap and no further evaluation  
\( n = 689 \)

\[ \Downarrow \]

Returned for Further Evaluation  
\( n = 1154 \)

\( \text{Group B} \)
No pap smear  
Had Colpo and Biopsy  
\( n = 23 \)

\[ \Downarrow \]

Had Same Day Screening and Pap smear  
\( n = 158 \)

\[ \Downarrow \]

Pap Only  
\( n = 20 \)

\[ \Downarrow \]

Pap + Colposcopy  
\( n = 127 \)

\[ \Downarrow \]

Pap + Biopsy + Colposcopy  
\( n = 11 \)

\( \text{Group D} \)
Pap & Colposcopy  
\( n = 187 \)

\[ \Downarrow \]

Pap & Colposcopy  
\( n = 967 \)
<table>
<thead>
<tr>
<th>Author &amp; Year</th>
<th>Screening Method</th>
<th>Study Type</th>
<th>Study population</th>
<th>Outcome Measure</th>
<th>Results</th>
<th>Study Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denny et al 2005</td>
<td>VIA</td>
<td>RCT</td>
<td>n=6555</td>
<td>High grade cervical lesions on biopsy (CIN2) after screening and treatment for each method and control group at 6 months and 12 months</td>
<td>-6 months after VIA screen and tx: 2.23% of women with positive VIA had CIN2 or higher at 6 months, p=0.02</td>
<td>Excellent</td>
</tr>
<tr>
<td></td>
<td>HPV DNA Testing</td>
<td></td>
<td></td>
<td></td>
<td>-6 months after HPV screen and tx: 0.80% HPV positive women had CIN2 or higher at 6 months, p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-For delayed tx group: 3.55% had CIN2 or higher</td>
<td></td>
</tr>
<tr>
<td>Mitchell et al 1998</td>
<td>Colposcopy</td>
<td>Meta-Analysis</td>
<td>n=5978, 9 studies</td>
<td>LSIL or HSIL</td>
<td>Weighted average sensitivity 96%, Weighted average specificity 48%</td>
<td>Good</td>
</tr>
<tr>
<td>Massad &amp; Collins</td>
<td>Colposcopy</td>
<td>Cross Sectional</td>
<td>n=2112</td>
<td>See results</td>
<td>Sensitivity 56% for CIN II/III, Specificity 52% for atypia and worse lesions</td>
<td>Good</td>
</tr>
<tr>
<td>Hutchinson et al 1999</td>
<td>LBC</td>
<td>Cross sectional comparative study</td>
<td>n=8,636</td>
<td>ASCUS or HSIL</td>
<td>Detected 12.7% ASCUS, Sensitivity for HSIL, 92.9%, p&lt;0.001</td>
<td>Good</td>
</tr>
<tr>
<td></td>
<td>Pap smear</td>
<td></td>
<td></td>
<td>ASCUS or HSIL</td>
<td>Detected 6.7% ASCUS, Sensitivity 77.8% for HSIL, p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>JHPIEGO Cervical Cancer Project, Zimbabwe 1999</td>
<td>Pap smear</td>
<td>Cross sectional comparative study</td>
<td>n=10,934</td>
<td>Mild Dysplasia or worse lesions</td>
<td>Sens: 29.6% (95% CI 25.6-33.8), Spec: 92.3% (90.9-93.6)</td>
<td>Fair-good</td>
</tr>
<tr>
<td></td>
<td>VIA</td>
<td></td>
<td></td>
<td>Mild Dysplasia or worse lesions</td>
<td>Sens: 63.5% (95% CI 59.1-67.7), Spec: 63.7% (95% CI 65.0-69.6)</td>
<td></td>
</tr>
<tr>
<td>Sankaranarayanan et al 2005</td>
<td>VIA</td>
<td>Cluster RCT</td>
<td>n=142 701</td>
<td>-Screening test positivity</td>
<td>Detection: 14.0%, High grade: 0.7% (p=0.06 for all detection results)</td>
<td>Good</td>
</tr>
<tr>
<td>Study</td>
<td>Method</td>
<td>Design</td>
<td>Endpoint</td>
<td>Detection: 10.3%</td>
<td>Specificity 99.0 (95% CI 98.7-99.3)</td>
<td>Sensitivity 18.4 (11.6-25.7)</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------------------</td>
<td>------------------------------</td>
<td>------------------------------------------------------</td>
<td>-----------------</td>
<td>-----------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td><strong>Schneider et al 2000</strong></td>
<td><strong>Colposcopy</strong></td>
<td>Cross sectional comparative</td>
<td>CIN 2/3 and cancer</td>
<td>Sensitivity 13.2 (95% CI 7.3-19.6)</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n=5,455</td>
<td></td>
<td>Specificity 99.2 (95% CI 98.9-99.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Conventional cytology</strong></td>
<td></td>
<td></td>
<td></td>
<td>Sensitivity 18.4 (11.6-25.7)</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td><strong>Wu et. al 2005</strong></td>
<td><strong>Colposcopy as screening method</strong></td>
<td>Cross sectional comparative</td>
<td>Presence of cervical dysplasia</td>
<td>Sensitivity 13.2 (95% CI 7.3-19.6)</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>n=273</td>
<td></td>
<td>Specificity 99.2 (95% CI 98.9-99.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>VIA</strong></td>
<td>n=450</td>
<td></td>
<td>Sensitivity 13.2 (95% CI 7.3-19.6)</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>LBC</strong></td>
<td>n=273</td>
<td></td>
<td>Sensitivity 13.2 (95% CI 7.3-19.6)</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>HPV DNA</strong></td>
<td>n=450</td>
<td></td>
<td>Sensitivity 13.2 (95% CI 7.3-19.6)</td>
<td>Good</td>
<td></td>
</tr>
</tbody>
</table>

VIA: Visual Inspection with Acetic Acid; RCT: Randomized Controlled Trial; LBC=Liquir Based Cytology; LSIL: Low Grade Squamous Intraepithelial Neoplasia; HSIL= High Grade Squamous Intraepithelial Neoplasia CI = Confidence interval; ASCUS= Atypical Squamous Cells of Undetermined Significance
Table 2  Final Study Populations for Three Diagnostic Accuracy Analyses in Leogane, Haiti 2000-2005

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Boxes from Figure 1</th>
<th>Total Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis #1: Sensitivity and Specificity of Pap Smear compared to Cervical Biopsy</td>
<td>C+D</td>
<td>198</td>
</tr>
<tr>
<td>Analysis #2: Sensitivity and Specificity of Colposcopy compared to Cervical Biopsy</td>
<td>B+C+D</td>
<td>221</td>
</tr>
</tbody>
</table>
Table 3: Sociodemographic Characteristics and Risk Factors for Cervical Cancer by Study Population Characteristics Reported at Initial Evaluation
Leogane, Haiti 2000-2005

<table>
<thead>
<tr>
<th></th>
<th>Group A n=2094</th>
<th>Pap + Biopsy Analysis 1 (Groups C+D) n=198</th>
<th>Colposcopy + Biopsy Analysis 2 (Groups C+D+E) n=221</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age mean, yrs (SD)</td>
<td>44.5 (10.5)</td>
<td>44.4 (9.2)</td>
<td>45.4 (9.9)</td>
</tr>
<tr>
<td>Missing</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;30</td>
<td>267 (13.2)</td>
<td>13 (6.6)</td>
<td>13 (5.9)</td>
</tr>
<tr>
<td>31-40</td>
<td>648 (32.0)</td>
<td>53 (26.8)</td>
<td>55 (24.9)</td>
</tr>
<tr>
<td>41-50</td>
<td>669 (33.0)</td>
<td>86 (43.4)</td>
<td>95 (44.0)</td>
</tr>
<tr>
<td>&gt;51</td>
<td>440 (21.7)</td>
<td>46 (23.2)</td>
<td>58 (26.2)</td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Number of Partners in Lifetime</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1211 (59.8)</td>
<td>107 (54.0)</td>
<td>116 (52.5)</td>
</tr>
<tr>
<td>2-4</td>
<td>760 (37.6)</td>
<td>86 (43.4)</td>
<td>97 (43.9)</td>
</tr>
<tr>
<td>&gt;5</td>
<td>53 (2.4)</td>
<td>5 (2.5)</td>
<td>8 (3.6)</td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single or Divorced</td>
<td>42 (2.1)</td>
<td>5 (2.6)</td>
<td>5 (2.4)</td>
</tr>
<tr>
<td>Living with Partner</td>
<td>961 (47.5)</td>
<td>105 (55.0)</td>
<td>116 (56.3)</td>
</tr>
<tr>
<td>Married</td>
<td>957 (47.3)</td>
<td>81 (42.4)</td>
<td>85 (41.2)</td>
</tr>
<tr>
<td>Unknown</td>
<td>(64)</td>
<td>(7)</td>
<td>(15)</td>
</tr>
<tr>
<td>Age First Coitus &lt; 16</td>
<td>268 (13.2)</td>
<td>30 (15.2)</td>
<td>34 (15.4)</td>
</tr>
<tr>
<td>Missing</td>
<td>(80)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gravidity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>184 (9.1)</td>
<td>9 (4.5)</td>
<td>9 (4.1)</td>
</tr>
<tr>
<td>1-4</td>
<td>1074 (53.1)</td>
<td>94 (47.5)</td>
<td>101 (45.7)</td>
</tr>
<tr>
<td>5-15</td>
<td>764 (37.8)</td>
<td>95 (48.0)</td>
<td>111 (50.2)</td>
</tr>
<tr>
<td>Missing</td>
<td>72</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Irregular Menses Currently</td>
<td>99 (4.9)</td>
<td>8 (4.0)</td>
<td>8 (3.7)</td>
</tr>
<tr>
<td>Unknown</td>
<td>33</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Postcoital Bleeding</td>
<td>18 (0.89)</td>
<td>5 (1.6)</td>
<td>9 (4.3)</td>
</tr>
<tr>
<td>missing</td>
<td>(38)</td>
<td>(5)</td>
<td>(9)</td>
</tr>
<tr>
<td>Pap Results in 4 categories</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>26 (2.0)</td>
<td>2 (1.0)</td>
<td>2 (1.0)</td>
</tr>
<tr>
<td>Abnormal/not dysplasia</td>
<td>1207 (93.2)</td>
<td>135 (68.2)</td>
<td>135 (68.2)</td>
</tr>
<tr>
<td>Dysplasia</td>
<td>30 (2.3)</td>
<td>29 (14.6)</td>
<td>29 (14.6)</td>
</tr>
<tr>
<td>Cancer</td>
<td>32 (2.5)</td>
<td>32 (16.2)</td>
<td>32 (16.2)</td>
</tr>
<tr>
<td>missing</td>
<td>729</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>Frank Cancer with Visual</td>
<td>54 (2.68)</td>
<td>32 (16.6)</td>
<td>54 (24.4)</td>
</tr>
<tr>
<td></td>
<td>Inspection (no equipment)</td>
<td>Pap Results</td>
<td>Colposcopy Result</td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------------------</td>
<td>-------------</td>
<td>-------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dysplasia or cancer</td>
<td>62 (4.8)</td>
<td>61 (30.8)</td>
<td>61 (30.8)</td>
</tr>
<tr>
<td>missing</td>
<td>24</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>Dysplasia or cancer</td>
<td>177 (13.5)</td>
<td>132 (66.7)</td>
<td>155 (70.1)</td>
</tr>
<tr>
<td>missing</td>
<td>(711)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dysplasia or cancer</td>
<td>83 (41.9)</td>
<td>106 (48.0)</td>
<td></td>
</tr>
</tbody>
</table>
Table 4A:
Bivariate Table for Diagnostic Accuracy of Pap Smear Compared to Cervical Biopsy
Leogane, Haiti 2000-2005

<table>
<thead>
<tr>
<th>Biopsy Results</th>
<th>Pap Smear Results</th>
<th>Negative</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive*</td>
<td>57</td>
<td>26</td>
<td>83</td>
</tr>
<tr>
<td>Negative</td>
<td>4</td>
<td>111</td>
<td>115</td>
</tr>
<tr>
<td>Total</td>
<td>61</td>
<td>137</td>
<td>198</td>
</tr>
</tbody>
</table>

A positive results means that the test demonstrated the presence of dysplasia or cancer.

Table 4B:
Bivariate Table for Diagnostic Accuracy of Colposcopy Compared to Cervical Biopsy
Leogane, Haiti 2000-2005

<table>
<thead>
<tr>
<th>Biopsy Results</th>
<th>Colposcopy Results</th>
<th>Negative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive*</td>
<td>102</td>
<td>4</td>
<td>106</td>
</tr>
<tr>
<td>Negative</td>
<td>53</td>
<td>62</td>
<td>115</td>
</tr>
<tr>
<td>Total</td>
<td>155</td>
<td>66</td>
<td>221</td>
</tr>
<tr>
<td>Analysis #1: Pap Smear compared to Cervical Biopsy</td>
<td>Prevalence (CI)</td>
<td>Sensitivity (CI)</td>
<td>Specificity (CI)</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>------------------</td>
</tr>
<tr>
<td></td>
<td>41.9 % (0.35-0.49)</td>
<td>68.7% (0.58-0.78)</td>
<td>96.5% (0.91-0.99)</td>
</tr>
<tr>
<td>Analysis #2: Colposcopy compared to Cervical Biopsy</td>
<td>48.0% (0.41-0.55)</td>
<td>96.2% (0.91-0.99)</td>
<td>53.9% (0.44-0.63)</td>
</tr>
</tbody>
</table>

CI - 95% Confidence Intervals  
PPV - Positive Predictive Value  
PPN - Negative Predictive Value