Screening for Squamous Intraepithelial Lesions with Fluorescence Spectroscopy

Michele Follen Mitchell, MD, MS, Scott B. Cantor, PhD, Carrie Brookner, MS, Urs Utzinger, PhD, David Schottenfeld, MD, and Rebecca Richards-Kortum, PhD

Objective: To evaluate the accuracy of fluorescence spectroscopy in screening for squamous intraepithelial lesions (SILs) and to compare its performance with that of Papanicolaou smear screening, colposcopy, cervicography, and human papillomavirus (HPV) testing.

Data Sources: Receiver operating characteristic (ROC) curve analysis was used to analyze performance by fluorescence spectroscopy (primary data) and other methods (secondary data).

Methods of Study Selection: In our search, 275 articles were identified in MEDLINE (1966–1996). Articles were included if the investigators had studied a population in whom low disease prevalence was expected; used either Papanicolaou smear screening and colposcopy or colposcopically directed biopsy as a standard against which the screening technique was measured, and included enough data for recalibration of reported sensitivities and specificities.

Tabulation, Integration, and Results: Receiver operating characteristic curves for fluorescence spectroscopy were calculated using a Bayesian algorithm, and ROC curves for the other screening methods were constructed using meta-analytic techniques. Areas under the ROC curves and Q points were calculated. Screening colposcopy had the highest area under the curve (0.95), followed by screening cervicography (0.90), HPV testing (0.88), cervicography (0.85), fluorescence spectroscopy (0.76), and Papanicolaou smear screening (0.70).

Conclusion: In terms of screening for SILs, fluorescence spectroscopy performed better than the standard technique, Papanicolaou smear screening, and less well than screening colposcopy, cervicography, HPV testing, and cervicography. The promise of this research technique warrants further investigation. (Obstet Gynecol 1999;94:889–96. © 1999 by The American College of Obstetricians and Gynecologists.)
columnar epithelium, and 78% for the transformation zone.\(^5\) We compared the results of a clinical trial of fluorescence spectroscopy for diagnosis of SIL in a referral setting with results reported for other diagnostic techniques.\(^6\) In the present study, fluorescence spectroscopy was used in the screening setting.

In our prior work,\(^6\) we evaluated the discriminative ability of tests for cervical precancer using receiver operating characteristic (ROC) curve analysis, a method increasingly being used to evaluate medical tests by the Food and Drug Administration as new devices are developed.\(^7\)–\(^10\) In this study, we did ROC curve analysis using data collected in the screening setting. Fahey et al\(^11\) recently published a meta-analysis of the Papanicolaou smear technique in screening and diagnostic settings. They estimated a sensitivity of 58% and a specificity of 68% in the combined settings. In that study, diagnostic and screening populations were combined. In the current and our previous work, we separated diagnostic from screening populations for the meta-analysis.\(^6\)

### Data Sources

Two methods of data collection were required for this study. For fluorescence spectroscopy, we used primary data collected from women in the screening setting.\(^5\) Subjects in the clinical study were recruited using an advertisement offering a free screening Papanicolaou smear, cancer-screening gynecologic examination, colposcopic examination, and fluorescence spectroscopic measurement of the cervix. Women were scheduled for screening if they had no histories of abnormal Papanicolaou smears, had no current signs of vaginal infections, and were not pregnant. A history was obtained from each subject, and a gynecologic examination, Papanicolaou smear screening, and colposcopy were done as well. A research spectroscopic system incorporating a pulsed nitrogen laser, a fiberoptic probe, and an optical multichannel analyzer was used to record fluorescence spectra. The system measured tissue fluorescence at excitation of 337, 380, and 460 nm and has been

### Table 1. Performance of Papanicolaou Smear Screening

<table>
<thead>
<tr>
<th>First author</th>
<th>Threshold</th>
<th>Standard</th>
<th>Pos</th>
<th>Neg</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>P (%)</th>
<th>100- Sp (%)</th>
<th>Se (%)</th>
<th>Sp (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson(^12)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Colpo and Pap, or bx</td>
<td>70</td>
<td>12</td>
<td>121</td>
<td>25</td>
<td>0.84</td>
<td>33</td>
<td>37</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>Bigrigg(^13)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Colpo and Pap, or bx</td>
<td>668</td>
<td>16</td>
<td>266</td>
<td>31</td>
<td>0.95</td>
<td>34</td>
<td>71</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Bolger(^14)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Colpo and Pap, or bx</td>
<td>25</td>
<td>37</td>
<td>11</td>
<td>18</td>
<td>0.40</td>
<td>67</td>
<td>68</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Chomet(^15)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Colpo and Pap, or bx</td>
<td>65</td>
<td>15</td>
<td>37</td>
<td>26</td>
<td>0.71</td>
<td>37</td>
<td>64</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>Engineer(^16)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>71</td>
<td>87</td>
<td>10</td>
<td>303</td>
<td>0.17</td>
<td>22</td>
<td>87</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>Frisch(^17)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>21</td>
<td>0.18</td>
<td>12</td>
<td>43</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>Giles(^18)</td>
<td>CIN 2</td>
<td>Bx</td>
<td>Colpo and Pap, or bx</td>
<td>5</td>
<td>9</td>
<td>3</td>
<td>182</td>
<td>0.04</td>
<td>5</td>
<td>60</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Gunderson(^19)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>4</td>
<td>2</td>
<td>16</td>
<td>31</td>
<td>0.38</td>
<td>9</td>
<td>23</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>Haddad(^4)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>87</td>
<td>13</td>
<td>12</td>
<td>9</td>
<td>0.82</td>
<td>58</td>
<td>87</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>Hellberg(^20)</td>
<td>CIN 2</td>
<td>Bx</td>
<td>Bx</td>
<td>15</td>
<td>3</td>
<td>65</td>
<td>15</td>
<td>0.82</td>
<td>20</td>
<td>19</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>Kashimura(^21)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0.82</td>
<td>17</td>
<td>95</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>Morrison(^22)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>23</td>
<td>50</td>
<td>10</td>
<td>44</td>
<td>0.26</td>
<td>53</td>
<td>69</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Nyirjesy(^23)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>83</td>
<td>26</td>
<td>24</td>
<td>0</td>
<td>0.80</td>
<td>98</td>
<td>77</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Okagaki(^24)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>1269</td>
<td>928</td>
<td>264</td>
<td>1084</td>
<td>0.43</td>
<td>46</td>
<td>83</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Parker(^25)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>154</td>
<td>30</td>
<td>20</td>
<td>237</td>
<td>0.39</td>
<td>12</td>
<td>88</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>Ramirez(^26)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>9</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>0.83</td>
<td>63</td>
<td>59</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>Reid(^27)</td>
<td>CIN 2</td>
<td>Bx</td>
<td>Bx</td>
<td>12</td>
<td>5</td>
<td>11</td>
<td>60</td>
<td>0.26</td>
<td>9</td>
<td>52</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>Soost(^28)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>1205</td>
<td>186</td>
<td>454</td>
<td>241</td>
<td>0.80</td>
<td>44</td>
<td>73</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>Soutter(^29)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>42</td>
<td>18</td>
<td>18</td>
<td>26</td>
<td>0.58</td>
<td>41</td>
<td>70</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>Soutter(^91)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>42</td>
<td>18</td>
<td>18</td>
<td>26</td>
<td>0.58</td>
<td>41</td>
<td>70</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>Staß(^30)</td>
<td>CIN 2</td>
<td>Bx</td>
<td>Bx</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>15</td>
<td>0.23</td>
<td>27</td>
<td>50</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>Szarewski(^31)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Colpo or bx</td>
<td>13</td>
<td>3</td>
<td>82</td>
<td>17</td>
<td>0.83</td>
<td>18</td>
<td>14</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>Tait(^32)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>26</td>
<td>14</td>
<td>25</td>
<td>0</td>
<td>0.78</td>
<td>97</td>
<td>51</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Tawa(^33)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>14</td>
<td>25</td>
<td>67</td>
<td>291</td>
<td>0.20</td>
<td>8</td>
<td>18</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>Upadhyay(^34)</td>
<td>CIN 1</td>
<td>Bx</td>
<td>Bx</td>
<td>238</td>
<td>52</td>
<td>2</td>
<td>16</td>
<td>0.78</td>
<td>76</td>
<td>99</td>
<td>24</td>
<td></td>
</tr>
</tbody>
</table>

Unweighted mean 62
Weighted mean 73
Unweighted mean 60
Weighted mean 63

Threshold = threshold for diagnosis of abnormality; Pos = technique used to determine presence of disease; Neg = technique(s) used to determine absence of disease; TP = number of true-positive; FP = number of false-positive; FN = number of false-negative; TN = number of true-negative; P = prevalence of disease; Sp = specificity; Se = sensitivity; CIN = cervical intraepithelial neoplasia; Bx = biopsy; Colpo = colposcopy.


† Papanicolaou smears were sent to two different laboratories.
described in detail elsewhere.2–5 On average, spectra were collected from two normal areas of squamous epithelium, two normal areas of columnar epithelium, and one area of the transformation zone. If detected, colposcopically abnormal sites were also measured. Approximately 10% of screening Papanicolaou smear results in the United States are abnormal.11

Fifty-five women were screened in this clinical trial. Spectroscopy data from one patient were lost, leaving data from 54 for analysis. All 54 women had Papanicolaou smears adequate for assessment; 50 had normal Papanicolaou smear results and four (10%) had abnormal results. One woman had a high-grade SIL, one had atypical cells of uncertain significance favoring dysplasia, and the other two had atypical cells of uncertain significance favoring human papillomavirus (HPV). The four women with abnormalities were referred for colposcopically directed biopsies. The woman with high-grade SIL was treated in our clinic with the loop electrosurgical excision procedure.

For all other screening techniques, we analyzed data from published reports identified in a MEDLINE search covering the period of 1966–1996. We used the search terms “Papanicolaou smear,” “colposcopy,” “cervicography,” “HPV testing,” “fluorescence spectroscopy,” and “polar probe.” Each term was combined with “screening,” “sensitivity,” “specificity,” “positive predictive value,” “negative predictive value,” and “receiver operating characteristic curve.”

Studies were selected using three criteria. The intent of the test had to be a screening in a low-disease-prevalence setting; reports of studies in which women were referred with abnormal Papanicolaou smear results were excluded because we assumed that such populations would have higher disease prevalence. The standard against which the technique was measured had to be either Papanicolaou smear screening and colposcopy or colposcopically directed biopsy, and the sensitivity and specificity calculations had to be reproducible from data in the report. The first two criteria were chosen to ensure selection of studies that were similar to our fluorescence spectroscopy study in terms of population prevalence characteristics and data analysis. In all selected studies, colposcopically normal areas were not biopsied.

Two hundred seventy-five articles were identified in the MEDLINE search, 66 of which were review articles about tests without data for analysis. Of the 59 articles reviewed by Fahey et al in their meta-analysis of Papanicolaou smear screening, 28 were reports of studies in which Papanicolaou smears were used for screening in a low-disease-prevalence setting, and that technique was compared with biopsy (disease presence or absence was determined with biopsy). Twenty-four (Haddad NG, Hussein IY, Livingstone JR, Smart GE. Colposcopy in teenagers [letter]. BMJ 1988;297:29–30) of those 28 were suitable for this analysis (Table 1). In six articles about colposcopy, the authors reported that colposcopy had been used for screening and sufficient information was included to permit recalculation of sensitivities and specificities against the standards (Table 2). In those, the standard was biopsy and disease absence was demonstrated by negative colposcopic and Papanicolaou smear results. Of the eight articles on cervicography in the screening setting, three were used for this analysis (Table 2). For those studies, disease presence was determined with biopsy, whereas disease absence was demonstrated by negative biopsy findings, negative colposcopic results, or a combination of other tests. In five articles about cervicography, the authors reported that cervicography had been used for screening and sufficient detail was included for our analysis (Table 4). Disease presence was determined with biopsy, and disease absence was demonstrated by negative Papanicolaou smear or biopsy results or negative findings by cervicography, cervicography or colposcopy. There were 20 articles about HPV testing using ViraPap (Digene Corp., Beltsville, MD), Hybrid Capture (Digene Corp.), or PCR. There would have been too few data for an ROC curve if the analysis had been limited to one type of HPV testing, so articles on

---

**Table 2. Performance of Colposcopy**

<table>
<thead>
<tr>
<th>First author</th>
<th>Threshold</th>
<th>Standard</th>
<th>Pos</th>
<th>Neg</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>P (%)</th>
<th>100-Sp (%)</th>
<th>Se (%)</th>
<th>Sp (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cecchini35</td>
<td>CIN 2</td>
<td>Bx or LEEP Colpo and Pap, or LEEP or bx</td>
<td>18</td>
<td>591 0</td>
<td>2391</td>
<td>0.01</td>
<td>20</td>
<td>97</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cecchini36</td>
<td>CIN 2</td>
<td>Bx or LEEP Colpo and Pap, or LEEP or bx</td>
<td>19</td>
<td>284 0</td>
<td>571</td>
<td>0.02</td>
<td>33</td>
<td>97</td>
<td>67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Davison27</td>
<td>CIN 1</td>
<td>Bx Colpo and Pap, or bx</td>
<td>25</td>
<td>6  8</td>
<td>157</td>
<td>0.17</td>
<td>4</td>
<td>75</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giles38</td>
<td>CIN 1</td>
<td>Bx Pap or bx</td>
<td>15</td>
<td>5  7</td>
<td>173</td>
<td>0.11</td>
<td>3</td>
<td>67</td>
<td>97</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hockstad39</td>
<td>CIN 1</td>
<td>Bx Colpo and Pap, or bx</td>
<td>3</td>
<td>20  4</td>
<td>43</td>
<td>0.10</td>
<td>32</td>
<td>44</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olatunbosun39</td>
<td>CIN 1</td>
<td>Bx Colpo and Pap, or bx</td>
<td>39</td>
<td>41  0</td>
<td>1469</td>
<td>0.03</td>
<td>3</td>
<td>99</td>
<td>97</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pap = Papanicolaou smear screening; LEEP = loop electrosurgical excision procedure; other abbreviations as in Table 1.
any of these three techniques were considered. Three articles were on techniques used in a screening setting and were suitable for analysis (Table 5). Disease presence was determined with Papanicolaou smear screening or biopsy; disease absence was demonstrated by negative Papanicolaou smear or biopsy findings or negative findings by cervicography or colposcopy. The one article found on use of the polar probe was on use of that technique in the diagnostic setting.

Tabulation and Integration

Bayesian statistical methods were used to classify primary data collected with fluorescence spectroscopy. Details of the algorithm have been reported elsewhere. The results of the algorithm were used to determine an ROC curve and calculate the area under the curve using the Excel software program (Microsoft Corp., Redmond, WA) following the method of Metz and Moses using Excel software. The method of meta-analysis is described in detail by Mitchell et al.

The thresholds of abnormality varied among studies. In some studies, the presence of normal tissue was distinguished from all abnormalities (atypia, low-grade SILs, high-grade SILs, and cancer), and in other studies, the presence of normal tissue and atypia were distinguished from low-grade SILs, high-grade SILs, and cancer. Tests were considered positive if they indicated low-grade SILs, high-grade SILs, or cancer. Those thresholds were accounted for by the meta-analytic method of Littenberg and Moses. In those studies, SILs were termed “cervical intraepithelial neoplasia” (CIN) and were classified as grade 1, 2, or 3. Low-grade SILs correspond to HPV or CIN 1, and high-grade SILs correspond to CIN 2 or CIN 3. Many of the studies were analyzed before the institution of the Bethesda system, so grades of CIN used by those authors are used in our tables as well.

Table 3. Performance of Cervicoscopy

<table>
<thead>
<tr>
<th>First author</th>
<th>Threshold</th>
<th>Standard</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>P (%)</th>
<th>100-SP (%)</th>
<th>Se (%)</th>
<th>Sp (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cecchini40</td>
<td>CIN2 Bx</td>
<td>Cyt, cervicog, and cervicosc, or colpo or bx</td>
<td>7</td>
<td>334</td>
<td>1</td>
<td>1694</td>
<td>0.004</td>
<td>16</td>
<td>83</td>
<td>84</td>
</tr>
<tr>
<td>Megevand41</td>
<td>HPV Bx</td>
<td>Pap and cervicosc, or colpo or bx</td>
<td>55</td>
<td>21</td>
<td>229</td>
<td>2121</td>
<td>0.12</td>
<td>1</td>
<td>19</td>
<td>99</td>
</tr>
<tr>
<td>Slawson42</td>
<td>HPV Bx</td>
<td>Cyt and cervicog, or colpo or bx</td>
<td>47</td>
<td>38</td>
<td>93</td>
<td>2375</td>
<td>0.05</td>
<td>2</td>
<td>34</td>
<td>98</td>
</tr>
<tr>
<td>Weighted mean</td>
<td></td>
<td></td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For the other screening techniques, data from the published studies were used to reproduce the reported calculations of sensitivity and specificity. Receiver operating characteristic curves and respective areas under the curves were calculated using the formula described by Littenberg and Moses using Excel software. The method of meta-analysis is described in detail by Mitchell et al.

Table 4. Performance of Cervicography

<table>
<thead>
<tr>
<th>First author</th>
<th>Threshold</th>
<th>Standard</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>P (%)</th>
<th>100-SP (%)</th>
<th>Se (%)</th>
<th>Sp (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baldauf43</td>
<td>CIN 1 Bx or ECC</td>
<td>Colpo or bx</td>
<td>51</td>
<td>135</td>
<td>11</td>
<td>1149</td>
<td>0.05</td>
<td>11</td>
<td>82</td>
<td>89</td>
</tr>
<tr>
<td>Baldauf43</td>
<td>CIN 1 Bx or ECC</td>
<td>Colpo or bx</td>
<td>33</td>
<td>39</td>
<td>29</td>
<td>1245</td>
<td>0.05</td>
<td>3</td>
<td>53</td>
<td>97</td>
</tr>
<tr>
<td>Cecchini44</td>
<td>HGSIL Bx</td>
<td>Pap, cervicog, and cervicosc, or colpo or bx</td>
<td>5</td>
<td>280</td>
<td>3</td>
<td>1748</td>
<td>0.004</td>
<td>14</td>
<td>61</td>
<td>86</td>
</tr>
<tr>
<td>Coibon45</td>
<td>CIN 1 Bx</td>
<td>Pap and cervicog, or bx</td>
<td>106</td>
<td>34</td>
<td>17</td>
<td>3858</td>
<td>0.03</td>
<td>1</td>
<td>86</td>
<td>99</td>
</tr>
<tr>
<td>Schneider46</td>
<td>CIN 2 Bx</td>
<td>Pap and cervicog, or colpo or bx</td>
<td>17</td>
<td>82</td>
<td>21</td>
<td>847</td>
<td>0.04</td>
<td>9</td>
<td>45</td>
<td>92</td>
</tr>
<tr>
<td>Tawa33</td>
<td>CIN 1 Bx</td>
<td>Pap and cervicog, or colpo or bx</td>
<td>72</td>
<td>301</td>
<td>9</td>
<td>2889</td>
<td>0.02</td>
<td>9</td>
<td>88</td>
<td>91</td>
</tr>
<tr>
<td>Weighted mean</td>
<td></td>
<td></td>
<td>76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ECC = endocervical curettage; HGSIL = high-grade squamous intraepithelial lesions; Pap = Papanicolaou smear screening; cervicosc = cervicography; cervicog = cervicoscopy; other abbreviations as in Table 1.

* Baldauf et al43 reported data using both the original reporting criteria (negative findings, positive findings, technically defective cervigram) and the new criteria (negative findings, atypical findings, positive findings, technically defective cervigram).
**Results**

The ROC curve calculated for fluorescence spectroscopy using Bayesian statistical methods is presented in Figure 1. The area under the curve was 0.76. The ROC curves for Papanicolaou smear screening, colposcopy, cervicoscopy, cervicography, and HPV testing are shown individually in Figure 2. The curves for all screening techniques are superimposed in Figure 3. The areas under the curves were 0.70 for Papanicolaou smear screening, 0.95 for colposcopy, 0.85 for cervicoscopy, 0.90 for cervicography, and 0.88 for HPV testing. Fluorescence spectroscopy compared favorably with the other tests but outperformed Papanicolaou smear screening, the current standard screening technique.

Q points are the uppermost points in ROC curves, at which sensitivity equals specificity. They are preferred by some meta-analysts for comparisons of ROC curves because confidence intervals can be obtained. The corresponding Q points and standard errors were as follows: Papanicolaou smear screening, 0.65 (0.04); colposcopy, 0.85 (0.05); cervicoscopy, 0.83 (0.05); cervicography, 0.83 (0.05); and HPV testing, 0.81 (0.07). A statistic could not be calculated for fluorescence spectroscopy because the ROC curve generated was from data from one study. The Q points for all other techniques were statistically significantly different from that of Papanicolaou smear screening: colposcopy, $P < .005$; cervicoscopy, $P < .05$; cervicography, $P < .005$; and HPV testing, $P < .005$.

**Discussion**

Cervical cancer is a disease for which screening is suitable because it is a serious disease for which early treatment is beneficial. Good screening tests should be easy to administer, be inexpensive, and cause minimal discomfort. Papanicolaou smear screening meets those

---

**Table 5. Performance of Human Papillomavirus Testing**

<table>
<thead>
<tr>
<th>First author</th>
<th>Test(s)</th>
<th>Threshold</th>
<th>Pos</th>
<th>Neg</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>P (%)</th>
<th>100-Sp (%)</th>
<th>Se (%)</th>
<th>Sp (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cuzick⁴⁷</td>
<td>PCR</td>
<td>CIN 1</td>
<td>Cyt</td>
<td>Cyt</td>
<td>43</td>
<td>97</td>
<td>85</td>
<td>1754</td>
<td>0.06</td>
<td>5</td>
<td>34</td>
<td>95</td>
</tr>
<tr>
<td>Schneider⁴⁶</td>
<td>Hybrid Capture</td>
<td>HGSIL</td>
<td>Bx</td>
<td>Pap and cervicogr, or colpo or bx</td>
<td>19</td>
<td>41</td>
<td>19</td>
<td>888</td>
<td>0.04</td>
<td>4</td>
<td>50</td>
<td>96</td>
</tr>
<tr>
<td>Zazove⁴⁸</td>
<td>ViraPap, PCR</td>
<td>LGSIL</td>
<td>Bx</td>
<td>Bx</td>
<td>164</td>
<td>2</td>
<td>16</td>
<td>18</td>
<td>0.90</td>
<td>12</td>
<td>91</td>
<td>88</td>
</tr>
</tbody>
</table>

Unweighted mean
Weighted mean

Unweighted mean
Weighted mean

Unweighted mean
Weighted mean

| PCR = polymerase chain reaction; Cyt = cytology; HGSIL = high-grade squamous intraepithelial lesions; Pap = Papanicolaou smear screening; cervicogr = cervicography; LGSIL = low-grade squamous intraepithelial lesions; other abbreviations as in Table 1. |

---

**Figure 1.** Receiver operating characteristic (ROC) curve for screening fluorescence spectroscopy. Dots are data points and the line is a fitted ROC curve.

**Figure 2.** Receiver operating characteristic (ROC) curves for Papanicolaou (Pap) smear screening, colposcopy, cervicoscopy, cervicography, and human papillomavirus (HPV) testing. Dots are data points and lines are fitted ROC curves.
cost-effective adjuvant to Papanicolaou smear screening so that it may be determined whether it is a useful and needed. Screening fluorescence spectroscopy will need rescence spectroscopy using biopsy as a standard are techniques. Studies of screening colposcopy and fluo-
test is one strategy; improving sensitivity and specific-
in detecting the precursor stage. In the United States, the threshold for evaluation has been reset so that women with almost any form of abnormal Papanicolaou smear are now referred for colposcopy, which has increased the cost of cervical cancer screening.

The incidence and prevalence of cervical cancer are decreasing. Lesions are being detected in the precursor stage. In the United States, the threshold for evaluation has been reset so that women with almost any form of abnormal Papanicolaou smear are now referred for colposcopy, which has increased the cost of cervical cancer screening.

In our analysis, screening colposcopy was an excellent discriminator, outperforming fluorescence spectroscopy, cervicography, HPV testing, cervicoscopy, and Papanicolaou smear screening. Colposcopy, however, is expensive, requires training, and is labor-intensive and therefore probably should not be considered as a screening test. Cervicography and HPV testing also perform well in the screening setting. Cervicography, like colposcopy, is highly dependent on good visual skills, is expensive, and is labor- and equipment-intensive. Human papillomavirus testing is not dependent on good visual skills, but it might not be cost-effective.

In our trial, in which Papanicolaou smear screening and colposcopy were used as the standard, fluorescence spectroscopy outperformed Papanicolaou smear screening in screening for SILs. This work is preliminary and is based on only one study, but the screening technique shows promise when compared with existing techniques. Studies of screening colposcopy and fluorescence spectroscopy using biopsy as a standard are needed. Screening fluorescence spectroscopy will need to be subjected to randomized multicenter clinical trials so that it may be determined whether it is a useful and cost-effective adjuvant to Papanicolaou smear screen-

requirements, although as assessed by Fahey et al,\textsuperscript{11} it has a sensitivity of 58% and specificity of 68%. Given these low levels, strategies that increase sensitivity and specificity may be called for. Adding a second screening test is one strategy; improving sensitivity and specificity of Papanicolaou smear screening is another.

The incidence and prevalence of cervical cancer are decreasing. Lesions are being detected in the precursor stage. In the United States, the threshold for evaluation has been reset so that women with almost any form of abnormal Papanicolaou smear are now referred for colposcopy, which has increased the cost of cervical cancer screening.

In our analysis, screening colposcopy was an excellent discriminator, outperforming fluorescence spectroscopy, cervicography, HPV testing, cervicoscopy, and Papanicolaou smear screening. Colposcopy, however, is expensive, requires training, and is labor-intensive and therefore probably should not be considered as a screening test. Cervicography and HPV testing also perform well in the screening setting. Cervicography, like colposcopy, is highly dependent on good visual skills, is expensive, and is labor- and equipment-intensive. Human papillomavirus testing is not dependent on good visual skills, but it might not be cost-effective.

In our trial, in which Papanicolaou smear screening and colposcopy were used as the standard, fluorescence spectroscopy outperformed Papanicolaou smear screening in screening for SILs. This work is preliminary and is based on only one study, but the screening technique shows promise when compared with existing techniques. Studies of screening colposcopy and fluorescence spectroscopy using biopsy as a standard are needed. Screening fluorescence spectroscopy will need to be subjected to randomized multicenter clinical trials so that it may be determined whether it is a useful and cost-effective adjuvant to Papanicolaou smear screen-

A potential criticism of the clinical measurements in this screening study is that each woman underwent colposcopy before fluorescence spectroscopy, which provided assurance that the cervix of a woman with negative Papanicolaou smear findings truly was clinically disease free. Our probe is the size of a pencil and measures a 2-mm area. Our goal was to test our algorithm, so we wanted to be sure we were placing the probe on normal tissue.

Fluorescence spectroscopy makes real-time diagnosis possible at the time of patient screening. Women then can be treated at screening visits, saving women and the health care system money and time that would have been spent on return visits for colposcopy and colposcopically directed biopsies and for treatment, and saving women the 1- to 2-week anxiety-producing period in which results are awaited, as well as costs of time off work, child care, and parking. Our cost-effectiveness analysis showed that in the diagnostic setting, fluorescence spectroscopy could save the United States $625 million annually.\textsuperscript{49} We are pursuing a similar cost-effectiveness analysis of fluorescence spectroscopy in the screening setting.

Another advantage of fluorescence spectroscopy is its ease of use. In the United States, most Papanicolaou smear screening is done by physicians, physician assistants, and nurse practitioners. In rural areas and some underserved urban settings, screening is done by registered nurses. In developing countries, much of the screening is done by nondegreeed trained health care workers. In some settings, visual inspection of the cervix is the only affordable option. In other settings, health care workers do Papanicolaou smear screening. Fluorescence spectroscopy involves simply the placement of a probe on the cervix, so it is technically easier to do than Papanicolaou smear screening and could be used by less trained members of the health care team, further reducing costs. We found that with fluorescence spectroscopy, a hospital aide can obtain the same results as a physician or nurse practitioner can (unpublished data). The algorithms work without a priori knowledge of the presence of SILs, and thus expertise in recognizing lesions is not necessary.\textsuperscript{5}

Although our data suggest that fluorescence spectroscopy might outperform Papanicolaou smear screening in detecting SILs, Papanicolaou smear screening does have an advantage in that it permits sampling of the endocervical canal. The flat fluorescence spectroscopy probe used in this study cannot—a major limitation to the use of the technology in the screening setting. Our team is working on strategies for assessing the endocervical canal spectroscopically. Also, the current probe

Figure 3. Superimposed receiver operating characteristic curves for all six screening techniques studied. HPV = human papillomavirus testing; Pap = Papanicolaou smear screening.
measures only a small area of the cervix, another barrier for use in screening. The probe should be capable of measuring the surface of many different sizes and shapes of cervices. Those issues will have to be explored extensively before the technology is used in larger screening trials.

The research-level fluorescence spectroscopy device performs well, but the design of a commercial prototype that performs as well and meets the stringent requirements of the Food and Drug Administration is a challenge to industry. Making probes that can be sterilized readily and that can sample the whole cervix adequately will be equally challenging. For society to benefit maximally from this technology, the device must be priced so the savings in health care dollars from the fewer clinic visits are realized. Our group estimates that devices could be made for as little as $3600.

With high value placed on streamlining procedures and conserving resources, exciting opportunities lie ahead for women and physicians willing to explore the benefit optical technologies might bring.

References


Address reprint requests to:
Michele Follen Mitchell, MD, MS
Department of Gynecologic Oncology
University of Texas M. D. Anderson Cancer Center
1515 Holcombe Boulevard, Box 67
Houston, TX 77030
E-mail: mfollem@mdanderson.org

Received March 17, 1999.
Received in revised form April 30, 1999.
Accepted May 13, 1999.

Copyright © 1999 by The American College of Obstetricians and Gynecologists. Published by Elsevier Science Inc.