Review

Fluorescence spectroscopy for diagnosis of squamous intraepithelial lesions of the cervix

Michele Follen Mitchell, MD, MS, Scott B. Cantor, PhD, Nirmala Ramanujam, PhD, Guillermo Tortolero-Luna, MD, PhD, and Rebecca Richards-Kortum, PhD

Objective: **To calculate receiver operating characteristic (ROC) curves for fluorescence spectroscopy in order to measure its performance in the diagnosis of squamous intraepithelial lesions (SILs) and to compare these curves with those for other diagnostic methods: colposcopy, cervicography, speculoscopy, Papanicolaou smear screening, and human papillomavirus (HPV) testing.**

Data Sources: **Data from our previous clinical study were used to calculate ROC curves for fluorescence spectroscopy. Curves for other techniques were calculated from other investigators' reports. To identify these, a MEDLINE search for articles published from 1966 to 1996 was carried out, using the search terms "colposcopy," "cervicoscopy," "cervicography," "speculoscopy," "Papanicolaou smear," "HPV testing," "fluorescence spectroscopy," and "polar probe" in conjunction with the terms "diagnosis," "positive predictive value," "negative predictive value," and "receiver operating characteristic curve."**

Methods of Study Selection: **We found 270 articles, from which articles were selected if they reported results of studies involving high-disease-prevalence populations, reported findings of studies in which colposcopically directed biopsy was the criterion standard, and included sufficient data for recalculation of the reported sensitivities and specificities.**

Tabulation, Integration, and Results: **We calculated ROC curves for fluorescence spectroscopy using Bayesian and neural net algorithms. A meta-analytic approach was used to calculate ROC curves for the other techniques. Areas under the curves were calculated. Fluorescence spectroscopy using the neural net algorithm had the highest area under the ROC curve, followed by fluorescence spectroscopy using the Bayesian algorithm, followed by colposcopy, the standard diagnostic technique. Cervicography, Papanicolaou smear screening, and HPV testing performed comparably with each other but not as well as fluorescence spectroscopy and colposcopy.**

Conclusion: **Fluorescence spectroscopy performs better than colposcopy and other techniques in the diagnosis of SILs. Because it also permits real-time diagnosis and has the potential of being used by inexperienced health care personnel, this technology holds bright promise. (Obstet Gynecol 1999;93:462–70. © 1999 by The American College of Obstetricians and Gynecologists.)**

Cervical cancer is the second most frequent cancer in women and the leading cause of cancer mortality in women worldwide.[1] The introduction of comprehensive screening and detection programs using the Papanicolaou smear has led to a substantial decrease in the mortality of cervical cancer over the last 50 years.[2,3] Most patients with abnormal Papanicolaou smears are evaluated with colposcopy, which is an accurate diagnostic method but is expensive and requires considerable skill. New strategies that lower costs are needed.

Laser-induced fluorescence spectroscopy, a noninvasive real-time technique for evaluating neoplasia, measures the autofluorescence of tissue based on the amounts of naturally occurring fluorophores present.[4] By modeling measurements of pure fluorophores, Ramanujam et al[5] demonstrated how fluorophores might change in concentration in preneoplastic and neoplastic tissues, accounting for the differences in autofluorescence that are seen in each grade of squamous intraepithelial lesions (SILs). With fluorescence spectroscopy, diagnostic algorithms can be derived that allow reasonable sensitivity and specificity for the diagnosis of SILs and work without *a priori* information about the abnormalities of the cervix.[6–8] The use of this new technology for the diagnosis of SILs has been reported.[6] Brookner et al (personal communication) found sensitivities of 87% for squamous epithelium, 96% for columnar epithelium, and 78% for the transformation zone. These sensitivities are comparable to the sensitivity of colposcopy performed by an experienced colposcopist.[9]

New medical technologies can be evaluated using several measures, including sensitivity, specificity, positive and negative predictive values, receiver operating characteristic (ROC) curves, and areas under ROC curves. The ROC curve has the advantage of comparing test performance over several thresholds and can be

From the Department of Gynecologic Oncology and the Section of General Internal Medicine, Department of Medical Specialties, University of Texas M. D. Anderson Cancer Center, Houston; and the Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, Texas.

Supported by the National Science Foundation and the Whitaker Foundation.

Table 1. Performance of Colposcopy for Diagnosis of Squamous Intraepithelial Lesions

First author	Threshold	Criterion standard		TP	FP	FN	TN	Prev	100-Sp	Se	Sp
		Pos	Neg								
Benedet[14]	CIN I	Bx	Bx	434	53	2	60	0.79	0.47	0.99	0.53
Benedet[15]	CIN I	Bx	Bx	2284	467	131	370	0.74	0.56	0.95	0.44
Cristoforoni[16]	HPV	Bx	Bx	127	38	3	20	0.69	0.65	0.97	0.35
Edebiri[17]	HPV*	Bx	Bx	113	30	17	62	0.59	0.33	0.87	0.67
Ferris18	CIN I	Bx	Bx	112	69	3	21	0.56	0.76	0.97	0.24
Javaheri[19]	CIN I	Bx	Bx	680	28	1	194	0.75	0.13	1.00†	0.87
Lozowski[20]	CIN I	Bx	Bx	109	27	4	11	0.75	0.71	0.96	0.29
Seshadri[21]	CIN I	Bx	Bx	61	54	9	28	0.46	0.66	0.87	0.34
Stafl[22]	CIN I	Bx	Bx	493	118	6	42	0.76	0.74	0.99	0.26
Unweighted mean										0.95	0.44
Weighted mean										0.96	0.48

Threshold = threshold for diagnosis of abnormality; Pos = criterion used to indicate positive for disease; Neg = criterion used to indicate negative for disease; TP = true positive; FP = false positive; FN = false negative; TN = true negative; Prev = prevalence; Sp = specificity; Se = sensitivity; CIN = cervical intraepithelial neoplasia; Bx = biopsy; HPV = human papillomavirus.

* Called "warty atypia" in this study.

† 99.8.

used both in diagnostic settings, in which the prevalence of disease is high, and in screening settings, in which the prevalence of disease is low.[10–13] In this article, we describe ROC curves for fluorescence spectroscopy, which we generated from measurements made in the diagnostic setting, and compare them with ROC curves for other diagnostic methods (colposcopy, Papanicolaou smear screening, cervicography, speculoscopy, and HPV testing), which we calculated from published reports.

Data Sources

To construct ROC curves for fluorescence spectroscopy, we used data from a clinical trial we described previously.[5–8] The trial included 95 women referred to the University of Texas M. D. Anderson Cancer Center Colposcopy Clinic between 1992 and 1994 because of abnormal Papanicolaou smear results. Briefly, a research device was used to measure fluorescence spectra at excitation wavelengths of 337, 380, and 460 nm. On average, spectra were collected in each patient from two normal and two abnormal sites, which were identified by colposcopy. Colposcopically directed biopsies then were performed of the abnormal sites that had been measured spectroscopically.

For all other diagnostic techniques, we constructed ROC curves using data from other investigators' published research reports. To identify these, a MEDLINE search for articles published from 1966 to 1996 was carried out, using the search terms "Papanicolaou smear," "colposcopy," "cervicoscopy," "speculoscopy," "cervicography," "HPV testing," "fluorescence spectroscopy," and "polar probe" in conjunction with the key words "diagnosis," "positive predictive value," "negative predictive value," and "receiver operating characteristic curve."

The colposcopy ROC curve has been reported previously.[9]

Methods of Study Selection

Two hundred seventy articles were identified in the MEDLINE search. Studies were selected if the particular test was used as a diagnostic measure in a high-disease-prevalence setting, if the criterion standard for presence of disease was colposcopically directed biopsy, and if enough data were provided for reproduction of the sensitivity and specificity calculations. All patients were referred with abnormal Papanicolaou smears. Articles were excluded if the test was used for screening (ie, a general group of patients was screened to find abnormalities and thus the disease prevalence was expected to be low). The rationale for excluding articles without clear indications of high disease prevalence or with mixed low and high disease prevalence and for excluding articles that used standards other than colposcopically directed biopsy for abnormalities was to facilitate comparison with our fluorescence spectroscopy patient population. In the patients in the selected studies, as in our patients, colposcopically normal areas were not biopsied.

Sixty-six of the 270 articles identified were excluded because they were review articles without data for analysis. There were 86 articles about colposcopy; as reported previously,[9] in nine articles,[14–22] colposcopy was used for diagnosis and there was sufficient information to recalculate sensitivities and specificities (Ta-

Table 2. Performance of Cervicography for Diagnosis of Squamous Intraepithelial Lesions

First author	Threshold	Criterion standard		TP	FP	FN	TN	Prev	100-Sp	Se	Sp
		Pos	Neg								
August[23]	HPV	Bx	Bx	206	127	45	208	0.43	0.38	0.82	0.62
Baldauf[24]	CIN I	Bx	Colpo negative or bx negative	51	135	11	127	0.19	0.52	0.82	0.48
Cecchini[25]*	HPV	Bx	Colpo negative or bx negative	133	170	30	232	0.29	0.42	0.81	0.58
Cecchini[25]*	HPV	Bx	Colpo negative or bx negative	133	222	29	170	0.29	0.57	0.82	0.43
Coibion[26]	CIN I	Bx	Colpo negative or bx negative	106	34	17	6	0.75	0.84	0.86	0.16
Ferris[27]	CIN I	Bx	Bx	127	44	40	13	0.75	0.77	0.76	0.23
Hall[28]	CIN I	Bx	Bx	28	4	25	14	0.75	0.24	0.53	0.76
Jones[29]	CIN I	Bx	Bx	47	59	5	90	0.26	0.40	0.90	0.60
Spitzer[30]	HPV	Bx	Bx	43	13	6	4	0.74	0.75	0.87	0.25
Spitzer[30]	HPV	Bx	Cerv and colpo negative or bx negative	43	13	6	20	0.60	0.40	0.87	0.60
Stafl[31]	CIN I	Bx	Cerv and colpo negative or bx negative	128	14	3	104	0.53	0.12	0.97	0.88
Tawa[32]	CIN I	Bx	Pap and cerv negative or bx negative	72	301	9	15	0.20	0.95	0.88	0.05
Unweighted mean										0.83	0.47
Weighted mean										0.83	0.47

Colpo = colposcopy; Cerv = cervicography; Pap = Papanicolaou smear; all other abbreviations as in Table 1.
*Data obtained by two different observers are listed separately.

ble 1). There were 35 articles about cervicography, of which 10[23–32] had sufficient detail to be used for the diagnostic analysis (Table 2). There were four articles about speculoscopy, two[33,34] of which could be used in a diagnostic analysis (Table 3). Fifty-nine articles were reviewed in the meta-analysis of Papanicolaou smear screening by Fahey et al[35]; 28 were articles in which the Papanicolaou smear was used for screening in a low-disease-prevalence setting, and these articles were excluded. In 31 articles, Papanicolaou smear screening was used for diagnosis; 25[20,29,30,36–57] were suitable for this analysis (Table 4). No articles published after the meta-analysis by Fahey et al[35] were suitable for our analysis. Twenty papers concerned HPV testing using ViraPap (Digene Corp., Beltsville, MD), Hybrid Capture (Digene Corp.) and polymerase chain reaction assays; because there would have been insufficient data for an ROC curve if the analysis had been limited to one type of HPV testing, articles using any of the three tests were selected. Eleven of the 20 articles[58–68] were used in a diagnostic setting and were suitable for analysis (Table 5). One article[69] concerned the polar probe; it was excluded from our analysis because biopsy was not the criterion standard.

Tabulation and Integration

Preliminary results of the fluorescence spectroscopy measurements, including sensitivity (the probability that test results are positive in the presence of disease) and specificity (the probability that test results are negative in the absence of disease) for one wavelength of measurement[7] and for three wavelengths,[6] have been reported. For this analysis, two statistical methods were used to classify these data: Bayesian and neural net algorithms. The Bayesian algorithm involves two steps: first, reduction of data into their principal components; second, Bayesian classification, in which prior probabilities are used to compute posterior probabilities. The neural net algorithm involves a multilayer radial basis function neural network that takes intensities at 15 excitation/emission wavelength pairs and develops coefficients by dividing the data into a training set and a testing set. The details of the algorithms are reported elsewhere.[6,8] The results of the algorithms were used to calculate two ROC curves, using Excel software (Microsoft Corp., Redmond, WA) following the methodology outlined by Metz[10] and Moses et al.[13] The areas under the curves were calculated.

Table 3. Performance of Speculoscopy for Diagnosis of Squamous Intraepithelial Lesions

First author	Threshold	Criterion Standard		TP	FP	FN	TN	Prev	100-Sp	Se	Sp
		Pos	Neg								
Lonky[33]	HPV	Bx	Bx	187	72	40	69	0.62	0.51	0.82	0.49
Massad[34]	HPV	Bx	Bx	33	21	10	28	0.47	0.43	0.76	0.57

Abbreviations as in Table 1.

Table 4. Performance of the Papanicolaou Smear for Diagnosis of Squamous Intraepithelial Lesions

First author	Threshold	Criterion standard: Pos	Criterion standard: Neg	TP	FP	FN	TN	Prev	100-Sp	Se	Sp
Anderson[36]	CIN I	Cone	Cone	65	10	6	6	0.82	0.62	0.92	0.38
Anderson[37]	CIN I	Cone	Cone	20	3	15	4	0.83	0.43	0.57	0.57
Andrews[38]	CIN I	Bx	Bx negative or colpo negative	35	92	20	156	0.18	0.37	0.64	0.63
Byrne[39]	CIN I	Bx	Bx	38	28	17	37	0.46	0.43	0.69	0.57
Giles[40]	CIN I	Bx	Bx	38	0	29	45	0.60	0.0	0.57	1.00
Hirschowitz[41]	CIN II	Bx	Bx	76	12	11	12	0.78	0.50	0.87	0.50
Jones[29]	CIN I	Bx	Bx	27	11	28	77	0.38	0.13	0.49	0.87
Jones[42]	CIN I	Bx	Bx	4	1	54	177	0.25	0.0	0.0	0.99
Kealy[43]	CIN I	Bx	Bx	80	25	13	182	0.31	0.12	0.86	0.88
Koonings[44]*	CIN I	Bx	Bx	61	20	27	35	0.62	0.36	0.69	0.64
Koonings[44]*	CIN I	Bx	Bx	62	20	16	49	0.53	0.29	0.79	0.71
Kwikkel[45]	CIN I	Bx	Bx	284	31	68	68	0.78	0.32	0.81	0.68
Lozowski[20]	CIN I	Bx	Bx	107	20	8	20	0.74	0.50	0.93	0.50
Maggi[46]	CIN I	Bx	Bx	40	43	12	47	0.37	0.48	0.77	0.52
Morrison[47]	CIN I	Bx	Bx	11	1	1	2	0.80	0.33	0.92	0.67
Oyer[48]	CIN I	Bx	Bx	223	22	21	72	0.72	0.23	0.91	0.77
Pearlstone[49]	CIN I	Bx	Bx	6	2	12	81	0.18	0.0	0.33	0.98
Robertson[50]	CIN I	Bx	Bx	348	41	212	103	0.80	0.28	0.62	0.72
Shaw[51]	CIN I	Bx	Bx	12	2	6	0	0.90	1.00	0.67	0.0
Skehan[52]	CIN I	Cone	Cone	43	15	15	14	0.67	0.52	0.74	0.48
Smith[53]	CIN I	Bx	Bx	71	13	20	18	0.75	0.42	0.78	0.58
Spitzer[30]	CIN I	Bx	Bx	10	31	5	32	0.19	0.49	0.67	0.51
Syrjanen[54]	CIN I	Bx	Bx	118	40	44	183	0.42	0.18	0.73	0.82
Tay[55]	CIN I	Bx	Bx	21	15	12	6	0.61	0.71	0.64	0.29
Walker[56]	CIN I	Bx	Bx	140	15	42	17	0.85	0.47	0.77	0.53
Wetrich[57]	CIN I	Bx	Bx	954	143	221	289	0.73	0.33	0.81	0.67
Unweighted mean										0.70	0.63
Weighted mean										0.75	0.73

Cone = cone biopsy, Colpo = colposcopy; all other abbreviations as in Table 1.
*Data for the cytobrush and cotton brush are listed separately.

For the other diagnostic techniques, data from the published studies were used for calculation of sensitivity and specificity. Using methods of meta-analysis[11] and Excel software, we estimated summary ROC curves from the independent reports of sensitivity and specificity of each test. Briefly, the logistic transforms of the true-positive ratio (Se) and the false-positive ratio (1-Sp) were calculated, a linear regression between the sum and difference of these transforms was performed, and the resulting regression line was reverse-transformed to yield the summary ROC curves. Areas under the ROC curves were calculated, using the formula described by Littenberg and Moses.[11]

There is no standard method for statistical comparison of areas under the curves. However, Moses et al[13] devised a method for comparing curves by comparing the Q point, the point at which sensitivity equals specificity. As described earlier, for all studies the criterion standard for a diagnosis positive for abnormality was a biopsy. Because biopsies were not performed unless abnormal areas were found, the standard for a diagnosis of negative or normal varied (negative cervical biopsy, negative cone biopsy, negative cervicographic, or negative colposcopic and negative cytologic findings were used).

The thresholds for classification of tissue as abnormal varied by study; for some studies, the threshold was normal tissue versus all abnormalities (atypia, low-grade SIL, high-grade SIL, and cancer), and for other studies, it was normal tissue and atypia versus low-grade SIL, high-grade SIL, and cancer. These variations are accounted for by the meta-analytic method of Littenberg and Moses.[11]

Results

The ROC curves calculated for fluorescence spectroscopy are presented in Figure 1. The neural net algorithm outperformed the Bayesian algorithm. Both the areas under the curve—0.87 for the neural net algorithm and 0.82 for the Bayesian algorithm—and the Q points—0.80 (standard error [SE] 0.01) for the neural net algorithm and 0.75 (SE 0.01) for the Bayesian algorithm—show this difference.

Tables 1–5 show the true-positive, false-positive, false-negative, and true-negative rates for the other

Table 5. Performance of Human Papillomavirus Testing for Diagnosis of Squamous Intraepithelial Lesions

First author	Test type	Threshold	Criterion Standard		TP	FP	FN	TN	Prev	100-Sp	Se	Sp
			Pos	Neg								
Becker[58]	ViraPap*	CIN	Bx	Cyt and colpo negative	133	47	67	290	0.37	0.14	0.66	0.86
Becker[58]	PCR	CIN	Bx	Cyt and colpo negative	165	131	11	180	0.36	0.42	0.94	0.68
Brown[59]	Hybrid Caputre*	LGSIL	Bx	Bx	15	3	8	8	0.68	0.29	0.65	0.71
Burger[60]	PCR	CIN I	Bx	Bx	86	12	37	22	0.78	0.36	0.70	0.64
Cox[61]	Hybrid Caputre*	CIN I	Bx	Colpo negative or box negative	43	48	7	119	0.23	0.29	0.86	0.71
Cox[62]	ViraPap*	LSGIL	Bx	Colpo negative or bx negative	85	56	52	289	0.28	0.16	0.62	0.84
Farthing[63]	Hybrid Capture*	CIN I	Bx	Bx	40	12	17	26	0.60	0.32	0.70	0.68
Ferenczy[64]	Hybrid Capture*	LGSIL	Bx	Bx	123	40	63	138	0.51	0.23	0.66	0.77
Holman[65]	ViraPap*	LGSIL	Bx	Colpo	27	1	80	133	0.44	0.0	0.25	0.99
Kaufman[66]	ViraPap*	CIN II	Bx	Bx	152	283	121	519	0.25	0.35	0.56	0.65
Nuovo[67]	ViraPap*	SIL	Bx	Colpo negative or bx negative	18	28	8	55	0.24	0.34	0.69	0.66
Sigurdsson[68]	ViraPap*	CIN I	Bx	Bx	54	8	28	10	0.82	0.45	0.66	0.55
Unweighted mean											0.66	0.72
Weighted mean											0.65	0.73

Cyt = cytology; Colpo = colposcopy; PCR = polymerase chain reaction; LGSIL = low-grade squamous intraepithelial lesion; SIL = squamous intraepithelial lesion; all other abbreviations as in Table 1.

* Digene Corp., Beltsville, MD.

diagnostic techniques for each study, and the calculated prevalences of disease in these populations are presented. For colposcopy (Table 1), the prevalence of disease varied from 0.46 to 0.79 and the mean weighted sensitivity and specificity were 96 and 48%, respectively. For cervicography (Table 2), the prevalence of disease varied from 0.19 to 0.75 and the mean weighted sensitivity and specificity were 83 and 47%, respectively. For speculoscopy (Table 3), the prevalences of disease were 0.47 and 0.62; no mean weighted sensitivity and specificity were calculated because there were only two studies. For Papanicolaou smear screening (Table 4), the prevalence of disease varied from 0.18 to 0.90 and the mean weighted sensitivity and specificity were 75 and 73%, respectively. For HPV testing (Table 5), the prevalence varied from 0.23 to 0.82 and the mean weighted sensitivity and specificity were 65 and 73%, respectively. The prevalence of disease varied widely, despite the fact that all patients were referred for colposcopy with abnormal Papanicolaou smears.

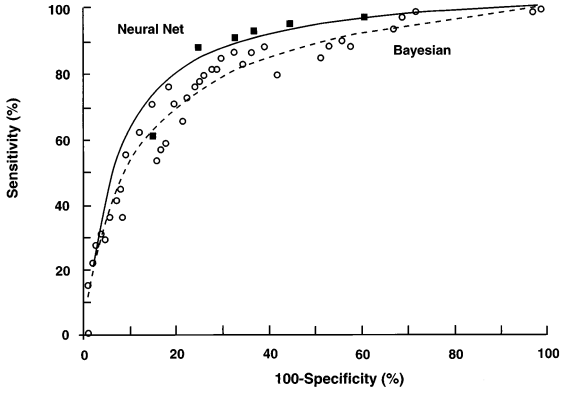

Figure 1. Receiver operating characteristic (ROC) curves for diagnostic fluorescence spectroscopy as analyzed by the Bayesian and neural net algorithms. The lines are fitted ROC curves.

The ROC curves for diagnostic colposcopy, cervicography, speculoscopy, Papanicolaou smear screening, and HPV testing are shown in Figure 2. The curve for speculoscopy is based on only two points and therefore should be considered speculative. The ROC curves for all techniques are superimposed in Figure 3. The areas under the curves were 0.84 for diagnostic colposcopy, 0.71 for cervicography, 0.76 for Papanicolaou smear screening, 0.75 for HPV testing, and 0.72 for speculoscopy (there were only two points). Fluorescence spectroscopy outperformed the other tests but, most important, compared favorably with colposcopy, the current standard diagnostic technique.

The Q points (SE) were 0.77 (0.07) for colposcopy, 0.66 (0.05) for cervicography, 0.70 (0.02) for Papanicolaou smear screening, 0.69 (0.08) for HPV testing, and 0.67 (0.06) for speculoscopy (Figure 4). By the Q point calculation, colposcopy was significantly better than cervicography, Papanicolaou smear screening, and

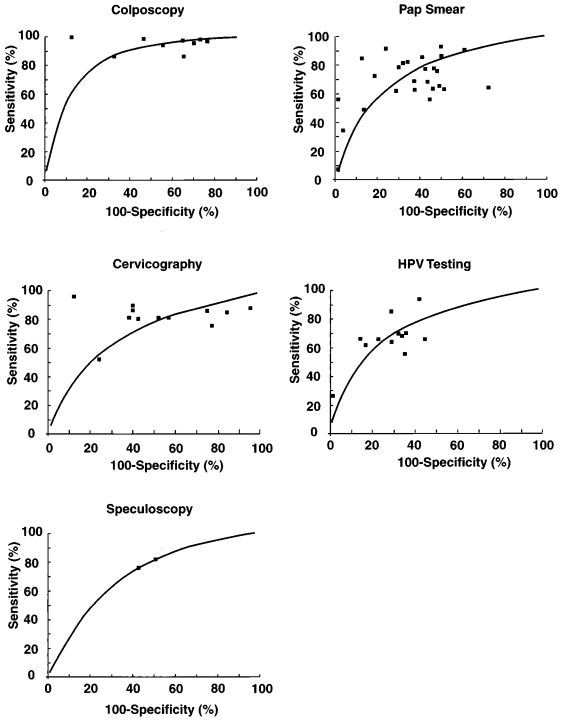

Figure 2. Receiver operating characteristic (ROC) curves for diagnostic colposcopy, cervicography, speculoscopy, Papanicolaou smear screening (Pap Smear), and human papillomavirus (HPV) testing. Data points represent mean values for each study listed in Tables 1–5, and the lines are fitted ROC curves.

HPV testing. There was no statistically significant difference between colposcopy and speculoscopy, but only two studies were reported for speculoscopy. Using this methodology, we could not compare fluorescence spectroscopy with the other techniques statistically because the denominator for calculating the Q point is the number of studies included in the meta-analysis and the

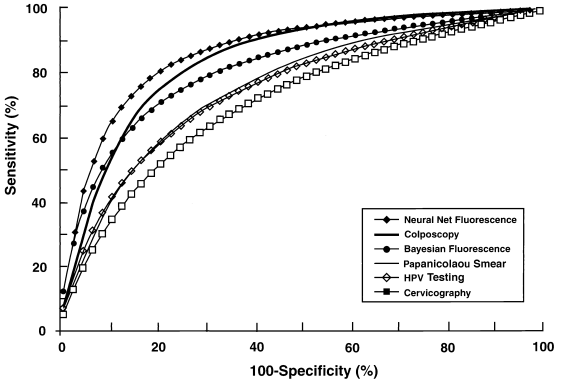

Figure 3. Receiver operating characteristic curves for colposcopy, cervicography, speculoscopy, Papanicolaou smear screening, human papillomavirus (HPV) testing, and fluorescence spectroscopy.

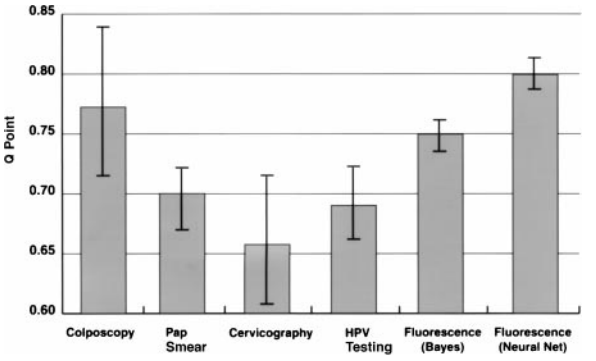

Figure 4. Q points (values at which sensitivity equals specificity) for the receiver operating characteristic curves. Speculoscopy is not included because only two studies were available for analysis. Pap Smear = Papanicolaou smear testing.

fluorescence spectroscopy curve was generated from only one study.

Comment

Diagnosis of SILs will continue to be an important part of the effort to decrease the morbidity and mortality of cervical cancer. In developing new technologies, decreasing cost and streamlining the process for the patient are important objectives. Techniques that add value are those that 1) allow us to identify patients at the highest risk for high-grade SILs, with the possibility of progression to malignancy, and 2) lead to immediate diagnosis, so evaluation and treatment can be done in a single outpatient visit. Avoiding second visits would lower health care costs and possibly patient anxiety.

For diagnosis, new technologies should be compared rigorously with colposcopy, the standard technique. Colposcopy works well when performed by an experienced colposcopist; our mean weighted sensitivity was 96% and our mean weighted specificity 48%. Specificity is lower at the expense of sensitivity; identifying lesions as cancer is more important than avoiding overcalling of lesions. The advantage of high sensitivity is offset somewhat by the level of expertise needed to perform colposcopy adequately. Studies of colposcopy that is performed by less experienced colposcopists are difficult to find. The studies included in our analysis involved clinicians who were experienced with the technique. Training in colposcopy requires a high volume of patients and good interactive correlation with cytologists and histopathologists, to ensure that what is seen colposcopically correlates well with the final cytologic and histopathologic diagnosis. The amount of time required for training depends on volume and the visual-recognition skills of the student. A second disadvantage of colposcopy is that results are not available until a week or two after the visit.

In our analysis, diagnostic Papanicolaou smear screening, cervicography, and HPV testing appear to have equal areas under the ROC curve, and these are slightly lower than the area under the ROC curve for colposcopy. Data for HPV testing might have indicated better performance if the analysis had been limited to one type of testing. Larger published studies in the literature also may show improved performance as the HPV-testing technology advances. The potential advantage of diagnostic Papanicolaou smear screening, cervicography, and HPV testing is that they are easy to perform and can be done by less experienced practitioners. Also, they might add value by helping to identify who is at higher risk for high-grade lesions. The disadvantage is that, as for colposcopy, results are not available until a week later and thus the patient must return for follow-up and treatment. At our institution, cervicography and speculoscopy would be no less expensive than colposcopy.

Speculoscopy appears to have a lower area under the ROC curve and thus theoretically adds less value than do diagnostic Papanicolaou smear screening, cervicography, and HPV testing. However, final judgments about speculoscopy should wait for the appearance of larger published diagnostic studies of good quality. Speculoscopy appears to have the disadvantage of requiring training and experience not unlike that needed for colposcopy. Because results depend on the biopsy targeted by speculoscopy, a diagnosis is not available for a week and the patient must return for follow-up and treatment. Therefore, at this time, speculoscopy has no theoretical advantage over colposcopy.

Our results show that fluorescence spectroscopy, currently performed with a research device, has a higher area under the curve than does colposcopy. It also provides an immediate diagnosis, allowing evaluation and treatment at a single visit. In this study, fluorescence spectroscopy was used to study areas identified by colposcopy. The algorithm now developed works without *a priori* information concerning what is normal or abnormal about the cervix and requires only that a probe be placed on the cervix. Any practitioner able to perform Papanicolaou smear screening could one day use fluorescence spectroscopy. Papanicolaou smears are obtained by nurse practitioners in some settings, by registered nurses in others, and by nondegreed health care workers in developing countries. The ease of use of fluorescence spectroscopy would permit a large number of trained personnel to use the device.

Although promising, this algorithm needs to be tested in the context of a large clinical trial comparing a prototype device that meets Food and Drug Administration standards with colposcopy in experienced hands. Once good results are confirmed, a large multicenter trial with the prototype will need to be performed. Adequate sample size and centralized consensus-derived histopathologic review will be crucial. A challenge to industry will be to make a prototype of low cost so that the savings realized by real-time diagnosis translate into savings in health care dollars.[71]

In addition, all of these strategies for diagnosis of SILs need to be subjected to a cost-effectiveness analysis. The cost of a return visit (medical costs, child-care costs, costs of time off work, and parking costs) will need to be considered in this analysis. Methods for measuring these variables will need to be developed. Although difficult to quantify, the anxiety of waiting for a diagnosis—as well as possible anxiety associated with the use of new technology—also must be measured.

The potential of emerging diagnostic technologies to help patients and lower health costs is great. Further studies are needed, but our results suggest that fluorescence spectroscopy may play a key role in SIL diagnosis in the future.

References

1. Mitchell MF, Hittelman WN, Hong WK, Schottenfeld D. The natural history of cervical intraepithelial neoplasia: An argument for intermediate endpoint biomarkers. Cancer Epidemiol Biomarkers Prev 1994;3:619–26.
2. Koss LG. The Papanicolaou test for cervical cancer detection: A triumph and a tragedy. JAMA 1989;261:737–43.
3. Kurman RJ, Henson DE, Herbst AL, Noller KL, Schiffman MH. Interim guidelines for management of abnormal cervical cytology. JAMA 1994;271:1866–9.
4. Richards-Kortum R, Mitchell MF, Ramanujam N, Mahadevan A, Thomsen S. *In vivo* fluorescence spectroscopy: Potential for noninvasive, automated diagnosis of cervical intraepithelial neoplasia and use as a surrogate endpoint biomarker. J Cell Biochem Suppl 1994;19:111–9.
5. Ramanujam N, Mitchell MF, Mahadevan A, Warren S, Thomsen S, Silva E, et al. *In vivo* diagnosis of cervical intraepithelial neoplasia using 337-nm-excited laser-induced fluorescence. Proc Natl Acad Sci U S A 1994;91:10193–7.
6. Ramanujam N, Mitchell MF, Mahadevan-Jansen A, Thomsen SL, Staerkel G, Malpica A, et al. Cervical pre-cancer detection using a multivariate statistical algorithm based on laser induced fluorescence spectra at multiple excitation wavelengths. Photochem Photobiol 1996;64:720–35.
7. Ramanujam N, Mitchell MF, Mahadevan A, Thomsen S, Silva E, Richards-Kortum R. Fluorescence spectroscopy: A diagnostic tool for cervical intraepithelial neoplasia (CIN). Gynecol Oncol 1994;52:31–8.
8. Tumer K, Ramanujam N, Ghosh J, Richards-Kortum R. Ensembles of radial basis function networks for spectroscopic detection of cervical precancer. IEEE Trans Biomed Eng 1998;45:953–61.
9. Mitchell MF, Schottenfeld D, Tortolero-Luna G, Cantor SB, Richards-Kortum R. Colposcopy for the diagnosis of squamous intraepithelial lesions: A meta-analysis. Obstet Gynecol 1998;91:626–31.
10. Metz CE. Basic principles of ROC analysis. Semin Nucl Med 1978;8:283–98.
11. Littenberg B, Moses LE. Estimating diagnostic accuracy from

multiple conflicting reports: A new meta-analytic method. Med Decis Making 1993;13:313–21.
12. Swets JA. Measuring the accuracy of diagnostic systems. Science 1988;240:1285–93.
13. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: Data-analytic approaches and some additional considerations. Stat Med 1993;12: 1293–316.
14. Benedet JL, Boyes DA, Nichols TM, Millner A. Colposcopic evaluation of patients with abnormal cervical cytology. Br J Obstet Gynaecol 1976;83:177–82.
15. Benedet JL, Anderson GH, Matisic JP, Miller DM. A quality-control program for colposcopic practice. Obstet Gynecol 1991;78:872–5.
16. Cristoforoni PM, Gerbaldo D, Perino A, Piccoli R, Montz FJ, Capitanio GL. Computerized colposcopy: Results of a pilot study and analysis of its clinical relevance. Obstet Gynecol 1995;85: 1011–6.
17. Edebiri AA. The relative significance of colposcopic descriptive appearances in the diagnosis of cervical intraepithelial neoplasia. Int J Gynaecol Obstet 1990;33:23–9.
18. Ferris DG, Miller MD. Colposcopic accuracy in a residency training program: Defining competency and proficiency. J Fam Pract 1993; 36:515–20.
19. Javaheri G, Fejgin MD. Diagnostic value of colposcopy in the investigation of cervical neoplasia. Am J Obstet Gynecol 1980;137: 588–94.
20. Lozowski MS, Mishriki Y, Talebian F, Solitare G. The combined use of cytology and colposcopy in enhancing diagnostic accuracy in preclinical lesions of the uterine cervix. Acta Cytol 1982;26:285–91.
21. Seshadri L, Jairaj P, Krishnaswami H. Colposcopy in the diagnosis of cervical neoplasia. Indian J Cancer 1990;27:180–6.
22. Stafl A, Mattingly RF. Colposcopic diagnosis of cervical neoplasia. Obstet Gynecol 1973;41:168–76.
23. August N. Cervicography for evaluating the "atypical" Papanicolaou smear. J Reprod Med 1991;36:89–94.
24. Baldauf JJ, Dreyfus M, Lehmann M, Ritter J, Philippe E. Cervical cancer screening with cervicography and cytology. Eur J Obstet Gynecol Reprod Biol 1995;58:33–9.
25. Cecchini S, Iossa A, Bonardi R, Gustavino C, Ciatto S. Evaluation of the sensitivity of cervicography in a consecutive colposcopic series. Tumori 1992;78:211–3.
26. Coibion M, Autier P, Vandam P, Delovelle A, Huet F, Hertens D, et al. Is there a role for cervicography in the detection of premalignant lesions of the cervix uteri? Br J Cancer 1994;70:125–8.
27. Ferris DG, Payne P, Frisch LE. Cervicography: An intermediate triage test for the evaluation of cervical atypia. J Fam Pract 1993;37:463–8.
28. Hall JB, McGee JA, Marroum M, Dee L. Evaluation of the cervicoscope as a screening instrument. Gynecol Oncol 1985;20:17–22.
29. Jones DE, Creasman WT, Dombroski RA, Lentz SS, Waeltz JL. Evaluation of the atypical Papanicolaou smear. Am J Obstet Gynecol 1987;157:544–9.
30. Spitzer M, Krumholz BA, Chernys AE, Seltzer V, Lightman AR. Comparative utility of repeat Papanicolaou smears, cervicography, and colposcopy in the evaluation of atypical Papanicolaou smears. Obstet Gynecol 1987;69:731–5.
31. Stafl A. Cervicography: A new method for cervical cancer detection. Am J Obstet Gynecol 1981;139:815–25.
32. Tawa K, Forsythe A, Cove JK, Saltz A, Peters HW, Watring WG. A comparison of the Papanicolaou smear and the cervicogram: Sensitivity, specificity and cost analysis. Obstet Gynecol 1988;71: 229–35.
33. Lonky NM, Mann WJ, Massad LS, Mutch DG, Blanco JS, Vasilev SA, et al. Ability of visual tests to predict underlying cervical neoplasia: Colposcopy and speculoscopy. J Reprod Med 1995;40: 530–6.
34. Massad LS, Lonky NM, Mutch DG, Mann WJ, Blanco JS, Vasilev SA, et al. Use of speculoscopy in the evaluation of women with atypical Papanicolaou smears: Improved cost effectiveness by selective colposcopy. J Reprod Med 1993;38:163–9.
35. Fahey MT, Irwig L, Macaskill P. Meta-analysis of Pap test accuracy. Am J Epidemiol 1995;141:680–9.
36. Anderson W, Frierson H, Barber S, Tabbarah S, Taylor P, Underwood P. Sensitivity and specificity of endocervical curettage and the endocervical brush for the evaluation of the endocervical canal. Am J Obstet Gynecol 1988;159:702–6.
37. Anderson DJM, Strachan F, Parkin DE. Cone biopsy: Has endocervical sampling a role? Br J Obstet Gynaecol 1992;99:668–70.
38. Andrews S, Hernandez E, Miyazawa K. Paired Papanicolaou smears in the evaluation of atypical squamous cells. Obstet Gynecol 1989;73:747–50.
39. Byrne P, Jordan J, Williams D, Woodman C. Importance of negative result of cervical biopsy directed by colposcopy. BMJ 1988;296:172.
40. Giles JA, Deery A, Crow J, Walker P. The accuracy of repeat cytology in women with mildly dyskaryotic smears. Br J Obstet Gynaecol 1989;96:1067–70.
41. Hirschowitz L, Raffle AE, Mackenzie EF, Hughes AO. Long term follow-up of women with borderline cervical smear test results: Effects of age and viral infection on progression to high grade dyskaryosis. BMJ 1992;304:1209–12.
42. Jones MH, Jenkins D, Cuzick J, Wolfendale MR, Jones JJ, Balogun-Lynch C, et al. Mild cervical dyskaryosis: Safety of cytologic surveillance. Lancet 1992;339:1440–3.
43. Kealy WF. Correlation of cervical cytodiagnosis and histopathology—An exercise in quality control. Ir J Med Sci 1986;155:381–8.
44. Koonings PP, Dickinson K, d'Ablaing G, Schlaerth JB. A randomized clinical trial comparing the cytobrush and cotton brush for Papanicolaou smears. Obstet Gynecol 1992;80:241–5.
45. Kwikkel HJ, Quaak MJ, de With C. Predictive value of the abnormal Pap smear: A retrospective analysis of error rates. Eur J Obstet Gynecol Reprod Biol 1986;21:101–12.
46. Maggi R, Zannoni E, Giorda G, Biraghi P, Sideri M. Comparison of repeat smear, colposcopy, and colposcopically directed biopsy in the evaluation of the mildly abnormal smear. Gynecol Oncol 1989;35:294–6.
47. Morrison EAB, Goldberg GL, Hagan RJ, Kadish AS, Burk RD. Self-administered home cervicovaginal lavage: A novel tool for the clinical-epidemiologic investigation of genital human papillomavirus infections. Am J Obstet Gynecol 1992;167:104–7.
48. Oyer R, Hanjani P. Endocervical curettage: Does it contribute to the management of patients with abnormal cervical cytology? Gynecol Oncol 1986;25:204–11.
49. Pearlstone AC, Grigsby PW, Mutch DG. High rates of atypical cervical cytology: Occurrence and clinical significance. Obstet Gynecol 1992;80:191–5.
50. Robertson JH, Woodend BE, Crozier EH, Hutchinson J. Risk of cervical cancer associated with mild dyskaryosis. BMJ 1988;297: 18–21.
51. Shaw CT. Colposcopic examination of patients with class II Papanicolaou smears. J Am Osteopath Assoc 1986;86:355–9.
52. Skehan M, Soutter WP, Lim K. Reliability of colposcopy and directed punch biopsy. Br J Obstet Gynaecol 1990;97:811–6.
53. Smith JJ, Cowan L, Hemstreet G 3d, Smith ML, Pippitt CH Jr, Doggett R, et al. Automated quantitative fluorescent image analysis of cervical cytology. Gynecol Oncol 1987;28:241–54.
54. Syrjanen KJ, Mantyjarvi R, Vayrynen M, Yliskoski M, Syrjanen SM,

Saarikoski S, et al. Cervical smears in the assessment of the natural history of human papillomavirus infections in prospectively followed women. Acta Cytol 1987;31:855–65.
55. Tay SK, Jenkins D, Singer P. Management of squamous atypia (borderline nuclear abnormalities): Repeat cytology or colposcopy? Aust N Z J Obstet Gynaecol 1987;27:140–1.
56. Walker EM, Dodgson J, Duncan ID. Does mild atypia on a cervical smear warrant further investigation? Lancet 1986;2:672–3.
57. Wetrich DW. An analysis of the factors involved in the colposcopic evaluation of 2194 patients with abnormal smears. Am J Obstet Gynecol 1986;154:1339–49.
58. Becker TM, Wheeler CM, McGough NS, Parmenter CA, Jordan SW, Stidley CA, et al. Sexually transmitted diseases and other risk factors for cervical dysplasia among Southwestern Hispanic and non-Hispanic white women. JAMA 1994;271:1181–8.
59. Brown DR, Rawlings K, Handy V, Fife KH, Bryan JT, Cramer H, et al. Human papillomavirus detection by hybrid capture in paired cervicovaginal lavage and cervical biopsy specimens. J Med Virol 1996;48:210–4.
60. Burger MPM, Hollema H, Peiters WJLM, Quint WGV. Predictive value of human papillomavirus type for histological diagnosis of women with cervical cytological abnormalities. BMJ 1995;310:94–5.
61. Cox JT, Lorincz AT, Schiffman MH, Sherman ME, Cullen A, Kurman RJ. Human papillomavirus testing by hybrid capture appears to be useful in triage in women with a cytologic diagnosis of atypical squamous cells of undetermined significance. Am J Obstet Gynecol 1995;172:946–53.
62. Cox JT, Schiffman MH, Winzelberg AJ, Patterson JM. An evaluation of human papillomavirus testing as part of referral to colposcopy clinics. Obstet Gynecol 1992;80:389–95.
63. Farthing A, Masterson P, Mason WP, Vousden KH. Human papillomavirus detection by hybrid capture and its possible clinical use. J Clin Pathol 1994;47:649–52.
64. Ferenczy A, Franco E, Arseneau J, Wright TC, Richart RM. Diagnostic performance of Hybrid Capture human papillomavirus deoxyribonucleic acid assay combined with liquid-based cytologic study. Am J Obstet Gynecol 1996;175:651–6.
65. Holman JR. Detection of HPV DNA in patients referred to a family practice colposcopy clinic. J Am Board Fam Pract 1996;9:162–6.
66. Kaufman RH, Adam E, Icenogle J, Lawson H, Lee N, Reeves KO, et al. Relevance of HPV screening in management of CIN. Am J Obstet Gynecol 1997;176:87–92.
67. Nuovo GJ, Moritz J, Walsh LL, MacConnell P, Koulos J. Predictive value of human papillomavirus DNA detection by filter hybridization of polymerase chain reaction in women with negative results of colposcopic examination. Am J Clin Pathol 1992;98:489–92.
68. Sigurdsson K, Arnadottir M, Snorradottir K, Benediksdottir K, Saemundsson H. Human papillomavirus (HPV) in an Icelandic population: Correlation of HPV DNA to cyto- and histopathologic lesions and evaluation of treatment strategies. Int J Gynaecol Cancer 1996;6:175–82.
69. Coppleson M, Reid BL, Skladnev VN, Dalrymple JC. An electronic approach to the detection of pre-cancer and cancer of the uterine cervix: A preliminary evaluation of polar probe. Int J Gynecol Cancer 1994;4:79–83.
70. Cantor SB, Mitchell MF, Tortolero-Luna G, Bratka CS, Bodurka DC, Richards-Kortum R. Cost-effectiveness analysis of diagnosis and management of cervical squamous intraepithelial lesions. Obstet Gynecol 1998;91:270–7.

Address reprint requests to:
Michele Follen Mitchell, MD, MS
Department of Gynecologic Oncology, Box 67
University of Texas
M. D. Anderson Cancer Center
Houston, TX 77030
E-mail: follen_mitchell@gyn.mda.uth.tmc.edu

Received February 11, 1998.
Received in revised form May 18, 1998.
Accepted June 19, 1998.