

Colposcopic Proficiency-Disease Spectrum in a Single Family Practice Colposcopists' Clinic

Gregory L. Brotzman, MD; Kenneth G. Schellhase, MD, MPH

ABSTRACT

Purpose: We sought to assess colposcopic proficiency in a family practice teaching clinic.

Methods: Subjects were a prospective cohort of women age 13 to 68 who were colposcopy clinic attendees from 1991 to 2002. Data recorded on each subject included demographic variables, sexual history, history of sexually transmitted diseases, reason for referral to colposcopy, Pap smear results, colposcopic impression, colposcopic biopsy results, and diagnoses. The Kappa statistic was used to measure agreement between clinical colposcopic assessment and biopsy results.

Results: Eight hundred twenty-six patients were enrolled. Compared to biopsy, colposcopic impression overall correctly predicted normal cervical biopsy in 55.8% (95% CI: 45.8%, 65.8%) of cases, and predicted abnormal biopsy 84.9% (95% CI: 81.6%, 88.1%) of the time. Colposcopic impression of low-grade squamous intraepithelial lesion (LSIL) correctly predicted LSIL on biopsy in 64.6% of cases, and correctly predicted the absence of LSIL 74.2% of the time. Colposcopic impression of high-grade squamous intraepithelial lesion (HSIL) correctly predicted biopsy results of HSIL in 70.05% of cases (Kappa = 0.544, $P < .0001$). There was a 12.7% error rate in discriminating normal from LSIL (Kappa -.258, with $P < .0001$).

Conclusion: Family physicians perform colposcopy with good correlation between colposcopic impression and subsequent histology.

INTRODUCTION

Family physicians frequently provide colposcopic ser-

vices to their patients.¹ While most family practice residency programs report offering colposcopic training to their residents, very little is written on proficiency in colposcopic assessment.²⁻⁶ Colposcopic skills can vary widely amongst physicians.⁶ Assurance of accuracy in addition to technical competency is important for patient care both during residency and in practice.^{7,8} Our hypothesis is that expertise in colposcopy, like other procedural skills, is based on training (and maintenance of current competency) and is not specialty-specific. The objectives of this study were to describe the spectrum of disease seen and assess the proficiency of 1 family practice colposcopic educator in the areas of colposcopic impression and correlation with subsequent histology.

METHODS

Design

This was a prospective cohort of consecutively enrolled first-time attendees of a colposcopy clinic from 1991 to 2002. Data were gathered prospectively using (1) a patient questionnaire filled out at the time of the visit, and (2) clinical data gathered by the primary author at the time of the visit and shortly afterward when pathology results became available. The patient questionnaire gathered data on patient characteristics, including patient age, sexually transmitted disease history, lifetime number of sexual partners, smoking status, and pregnancy status. Clinical data collection included reason(s) for colposcopic referral and prior Pap smear result, history of prior pap abnormalities, and any prior treatment. Only patients seen by the specific provider were included in this analysis, and only data from the first colposcopy visit was analyzed.

Setting

The setting of the study was an urban family practice residency program with a referral-based colposcopic practice.

Subjects

Subjects for the study consisted of women ages 13 to

Doctor Brotzman is a professor and Doctor Schellhase is an assistant professor with the Department of Family and Community Medicine, Medical College of Wisconsin. Please address correspondence to: Gregory L. Brotzman, MD, CSM Family Health Center, 1121 E North Ave, Milwaukee, WI 53212; phone 414.267.6500; fax 414.267.3892; e-mail brotzman@mcw.edu.

68 referred for colposcopic assessment; referrals came from within the residency practice itself, as well as from the larger community.

Interventions

Colposcopy was performed in a systematic fashion by a single faculty physician (with or without a resident learner) on each patient, with repeat cytology and biopsies obtained as indicated. A Cabot MM6000 colposcope (Cabot Medical, Racine, Wis) was used along with a Denvu image capture system (Tucson Ariz). Initial visual impression, cytologic impression, and histologic impression were taken into account to arrive at a final assessment. The colposcopist received didactic and hands-on training in colposcopy in 1990 followed by procedural precepting in 1991, after which he began a colposcopy clinic at his residency site. The vast majority of the time a resident learner was observing or assisting with the colposcopic procedures along with the colposcopist. Colposcopy was performed in a systematic fashion as described elsewhere.³ All patients seen by the colposcopist since the inception of his colposcopic practice through 2002 were included in this analysis.

Analysis

Statistical analyses were performed using SPSS for Windows, Version 9.0. Because not all study subjects received cervical biopsy, there might be selection bias on the basis of who did or did not receive biopsy; this bias could affect the denominator in calculations of test accuracy. Therefore to avoid confusion we chose not to use the terms "sensitivity," "specificity," "positive predictive value," or "negative predictive value" when describing test accuracy. Measures of test accuracy were determined using 2x2 tables to calculate simple proportions. For example, the proportion of correctly predicted biopsy results was calculated like a positive predictive value ($a/a+b$ by convention in a 2x2 table). Proportions of correctly predicted negative results were similarly calculated using the formula for negative predictive value ($d/b+d$). Confidence intervals around proportions were determined using the formula:

$$\pm 1.96 \sqrt{[\text{proportion} (1-\text{proportion}) / N]}$$

Agreement between clinical colposcopy impression and biopsy results was also measured using the Kappa statistic.

Human Subjects Protection

The Medical College of Wisconsin Internal Review Board approved this study.

RESULTS

Included in this analysis were 826 consecutively enrolled

patients (see Table 1). Patients' mean age was 29 years (range 13 to 68). Ten percent were pregnant at the time of their visit and 36% were smokers. The mean number of lifetime sexual partners was 8.4, and 60.3% had a history of at least 1 sexually transmitted disease (STD), with the most common STDs being chlamydia (30.3%), trichomoniasis (20.6%), and gonorrhea (17.7%). The most common reasons for colposcopy referral were atypical squamous cells of uncertain significance (ASCUS) pap (33.1%), low-grade squamous intraepithelial lesion (LSIL) pap (29.5%), and high-grade squamous intraepithelial lesion (HSIL) pap (10.5%). Nearly three quarters of the patients had never had a prior Pap abnormality, and few had ever had prior treatment for cervical disease. Two hundred sixty-four patients (32%) had a final clinical impression of "normal," and 562 (68%) had a final impression of "abnormal." Only 10 patients (1.2%) were judged to have had an inadequate colposcopic examination; these patients were excluded from further analyses below. An examination was considered inadequate if the Squamocolumnar junction and limits of the lesion were not seen in their entirety. The majority of patients (69.1%, $n=564$) had cervical biopsy, and, of these, 22.0% were normal, 44.7% were LSIL, and 23.2% were HSIL (Figure 1).

To determine the overall accuracy of the colposcopic impression, we used the cervical biopsy result as the gold standard. We found the initial colposcopic impression of "normal" predicted a normal biopsy result 55.8% of the time (95% CI: 45.8%, 65.8%). Further, we found an initial colposcopic impression of "abnormal" correctly predicted an abnormal biopsy result 84.9% of the time (95% CI: 81.6%, 88.1%)

Table 2 describes our findings in determining the ability of the colposcopist to discriminate between normal and LSIL. First we compared colposcopic assessments of normal with biopsy results of low-grade disease ("under-reading" the colposcopy), and found a 12.7% error rate; that is, in 12.7% of cases where there was a biopsy result of low-grade disease, there was a colposcopic impression of normal. The Kappa statistic for this discrepancy between impression and biopsy result was $-.258$, with $P < .0001$. Considering error in the opposite direction ("over-reading" the colposcopy), we also compared colposcopic assessments of LSIL with biopsy results that were normal, excluding those patients who had no biopsy at all ($n = 252$), and found an error rate of 38.1% with $\text{Kappa} = -0.087$ ($P < .0001$). Overall, the colposcopic impression correctly predicted the presence of LSIL, compared with gold standard biopsy results, 64.3% of the time (95% CI: 58.7%, 69.9%), and cor-

rectly predicted the absence of LSIL 74.2% of the time (95% CI: 69.1%, 79.3%).

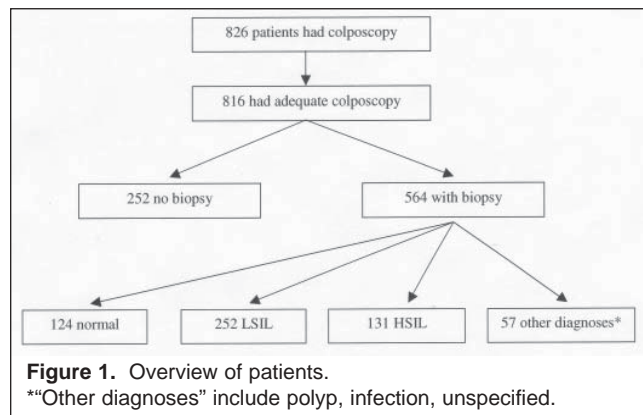
To determine the ability of the colposcopist to identify high-grade disease, we first compared colposcopy impression of “normal” with a biopsy result of high-grade disease and found a 3.1% error rate, Kappa = -0.238, $P < .0001$. Comparing colposcopy impression of low-grade disease with biopsy result of high grade disease, we found an error rate of 29.8%, Kappa = -0.054, $P = .074$. The colposcopic impression of HSIL, compared with biopsy results correctly predicted the presence of HSIL 70.0% of the time (95% CI: 61.7%, 78.3%), and correctly predicted the absence of HSIL 88.3% of the time (95% CI: 85.3%, 91.3%; Kappa = 0.544, $P < .0001$).

DISCUSSION

Colposcopy is a good diagnostic test, though far from perfect. We demonstrate that compared with the final histologic diagnosis, the colposcopic impression accurately predicted normal biopsy results 55.8% of the time, and abnormal results 84.9% of the time. This finding can be contrasted with Teale et al, who recently published a performance audit from a teaching hospital colposcopy clinic. Loop excision specimens from 589 patients were correlated with accuracy of diagnosis of high grade cervical intraepithelial neoplasia (CIN), yielding a correct diagnosis in 74.8% of the cases. However, this audit included only patients with known high-grade Pap smear results, thereby increasing the pre-test probability of finding high-grade results on colposcopy.⁷ In addition, Teale found that 21% of patients who underwent see-and-treat loop excision had no CIN in the loop specimen. This finding was present regardless of the level of experience of the colposcopist. Our results demonstrate a 12.7% error rate in discriminating between normal and LSIL, confirming that a significant portion of the time it can be a challenge to distinguish between normal and abnormal, particularly with low-grade findings. Homesley et al discuss the fact that accuracy may not improve as training progresses. Their study population was more comparable to ours, in that they compared the colposcopic and histologic diagnoses for 428 patients referred for colposcopy for a variety of reasons with a broad spectrum of Pap results. These subjects were evaluated by 23 residents, using a 5-component assessment system of vascular pattern, intercapillary distance, surface pattern, color of lesion, and borders. The overall sensitivity for diagnosis of disease was only 47% and specificity was 67%; these numbers did not improve through the 4 years of obstetric/gynecology (OB/GYN) residency training.⁸ Besides accuracy of colposcopic di-

Table 1. Cohort Demographic and Historical Information

Number of patients	826
Mean age (range, SD)	29 (13-68, 9.4)
Percent pregnant	10
Percent smokers	36
Mean # sexual partners (SD)	8.4 (38)
Percent having at least 1 prior STD	60.3
Percent with no prior abnormal Pap	73.5
Percent with no prior treatment	86.5
Reason for referral: ASCUS Pap (%)	33.1
Reason for referral: LSIL Pap (%)	29.5
Reason for referral: HSIL Pap (%)	10.5
Reason for referral: other reasons (%)	26.9



agnosis compared to biopsy, variations exist between different colposcopists viewing the same cervical lesion. Hopman et al presented 23 experienced colposcopists with 11 cervical images and asked them to grade the lesions and select the site to biopsy. The same individuals were asked to repeat this exercise 2-3 months later. The intraobserver concordance was 66.7%, showing a significant lack of colposcopic diagnostic consistency.⁹ Since our study looked at the impression and subsequent histologic finding of 1 colposcopist, we were not able to determine differences between him and others performing the procedure. Other studies show very good correlation between colposcopic assessment and subsequent histology. Ferris and Miller prospectively evaluated the colposcopic impression in 282 patients and compared them to the histology. They used mild, moderate, and severe dysplasia classifications. Using a somewhat more relaxed criterion of accuracy than we used (agreement within 1 histologic grade) they found that the colposcopic impression agreed within 1 histologic grade in 91.7% of cases. Of the 16 cases with a normal impression on colposcopy, none had normal biopsies.¹⁰ They subsequently recommended a minimal proficiency level of 80% for colposcopic accuracy to show proof of

Table 2. Error in Discriminating Between Normal, Low-Grade, and High-Grade Disease

	# With Colposcopy Impression	Total # With Biopsy Result	Error Rate (95% CI)
Colposcopic impression normal with biopsy LSIL	32	252	12.7% (8.6%, 16.8%)
Colposcopic impression LSIL vs. biopsy normal	48	124	39.7% (31.1%, 48.3%)
Colposcopic impression normal vs. biopsy HSIL	4	131	3.1% (.1%, 6.1%)
Colposcopic impression LSIL vs. biopsy HSIL	39	131	29.8% (22.0%, 37.6%)

colposcopic competency. Staffl and Mattingly used the same criteria for accuracy (within 1 histologic grade) in a review of 1410 colposcopic assessments. Among those patients with biopsies (659), they found an 85% correlation between impression and subsequent histology. The histology was more advanced in 3.3% of cases and less advanced in 11.7%.¹¹

Skehan et al reviewed 118 cone specimens and compared them to the preceding directed punch biopsy results. While not reported directly by the authors, these results show a positive predictive value of 33% and negative predictive value of 76% for colposcopy impression of normal vs. biopsy result. Indeed another area of concern is the accuracy of a punch biopsy itself. Skehan demonstrated a 54% false negative rate for the punch biopsy, with neither of the 2 microinvasive cancers identified on punch biopsy.¹² One must wonder if this is a result of inadequacy of the punch biopsy or a reflection on the inaccuracy of colposcopic detection of the worst lesion.

Spitzer et al surveyed family practice and obstetrician-gynecologist residency programs to evaluate the type of lesions seen and number of colposcopic procedures performed. Only 34% of family practice (FP) residency programs estimated having an adequate clinical volume of high-grade lesions, compared to 84% for OB/GYN programs. Both types of programs saw low-grade lesions, high grade lesions, and cancer, with the OB/GYN programs seeing 10 times more invasive disease than the FP programs.² Vigilance is therefore required to be sure cancer is not missed, since this is an infrequent occurrence, especially for the average family physician colposcopist.

Gordon and Hatch surveyed obstetrician/gynecologists in Arizona to determine who performed colposcopy and at what frequency. They found 98% reported performing colposcopy, with 60% of these individuals performing 2 or 3 colposcopies a week, and 20% performing less than 1 a week.¹³ Colposcopy, like other procedures, requires numerous exposures to develop competency. The previous studies mentioned emphasize the fact that training in a procedural skill does not always indicate proficiency in that procedure.

Colposcopy is a pattern recognition skill that requires experience to gain and maintain competency. Specific numbers necessary to obtain and subsequently maintain competency are somewhat arbitrary, but have been reported to be 25-50 precepted procedures to acquire competency.^{2,3} It is even less clear how many procedures need to be done on an ongoing basis to maintain proficiency. Colposcopy is both a skill and an art, with diagnostic accuracy improving, in general, as exposure to disease increases. Clinicians desiring to perform colposcopy well, regardless of their specialty, should be sure they are in a practice setting that will provide the necessary clinical material to allow for ongoing sharpening of their diagnostic skills.

This study has limitations. This study assesses the skills of only 1 faculty physician, whose skills may not be representative of other family physicians. Thus the results may not be generalizable to other family physicians who do not have a particular interest in colposcopy, or to other specialties. Also, because biopsies were not done on all patients, we were not able to calculate true sensitivity, specificity, or predictive values. We minimized the impact of this limitation by excluding those patients without biopsy results from calculations of colposcopic accuracy. However, our results still could be affected by this limitation, particularly when considering the ability in the initial colposcopic impression to differentiate normal from low-grade disease. It is more likely that low-grade disease may have a normal appearance and thus would more likely be missed on biopsy. Thus our estimates of ability to discriminate between normal and low-grade disease might be artificially high.

CONCLUSION

Our results indicate that family physicians can see a wide spectrum of colposcopic disease, similar to that seen by gynecologists, and can competently perform colposcopy with good correlation between colposcopic impression and histology.

ACKNOWLEDGMENTS

We would like to thank the Center for Practice Based Research for providing feedback on study design.

REFERENCES

1. Gordon P. Colposcopy training in family practice residency programs. *Fam Med.* 1991;23:27-30.
2. Spitzer M, Apgar B, Brotzman G, Krumholz B. Residency training in colposcopy: survey of program directors in obstetrics and gynecology and family practice. *Am J Obstet Gynecol.* 2001;507-513.
3. Brotzman G, Apgar B. Assessing colposcopic skills: the instructor's handbook. *Fam Med.* 1998;30:350-355.
4. Jordan J. International standards for training in colposcopy. *Clin Obstet Gynecol.* 1995;38:662-665.
5. Benedet JL, Anderson GH, Matistic JP, Miller DM. A quality control program for colposcopic practice. *Obstet Gynecol.* 1991;78:872-875.
6. Etherington IJ, Luesley DM, Shafi MI, Dunn J, Hiller L, Jordan JA. Observer variability among colposcopists from the West Midlands region. *Br J Obstet Gynaecol.* 1997;104:1380-1384.
7. Teale G, Etherington L, Luesley D, Jordan J. An audit of standards and quality in a teaching hospital colposcopy clinic. *Br J Obstet Gynaecol.* 1999;106:83-86.
8. Homesley H, Wolff J, Reish R, Jobson V. Evaluating the acquisition of colposcopy skills in an obstetric-gynecologic residency program. *J Reprod Med.* 1985;30:911-914.
9. Hopman E, Voorhorst F, Kenemans P, Meyer C, Helmersorst T. Observer agreement on interpreting colposcopic images of CIN. *Gynecol Oncol.* 1995;58:206-209.
10. Ferris D, Miller M. Colposcopic accuracy in a residency training program: defining competency and proficiency. *J Fam Pract.* 1993;36:515-520.
11. Stafil A, Marringly RF. Colposcopic diagnosis of cervical neoplasia. *Obstet Gynecol.* 1973;41:168-176.
12. Skehan M, Soutter W, Lim K, Krausz T, Pryse-Davies J. Reliability of colposcopy and directed punch biopsy. *Br J Obstet Gynaecol.* 1990;97:811-816.
13. Gordon P, Hatch K. Survey of colposcopy practices by obstetrician/gynecologists. *J Reprod Med.* 1992;37:861-863.

Wisconsin Medical Journal

The mission of the *Wisconsin Medical Journal* is to provide a vehicle for professional communication and continuing education of Wisconsin physicians.

The *Wisconsin Medical Journal* (ISSN 1098-1861) is the official publication of the Wisconsin Medical Society and is devoted to the interests of the medical profession and health care in Wisconsin. The managing editor is responsible for overseeing the production, business operation and contents of *Wisconsin Medical Journal*. The editorial board, chaired by the medical editor, solicits and peer reviews all scientific articles; it does not screen public health, socioeconomic or organizational articles. Although letters to the editor are reviewed by the medical editor, all signed expressions of opinion belong to the author(s) for which neither the *Wisconsin Medical Journal* nor the Society take responsibility. The *Wisconsin Medical Journal* is indexed in Index Medicus, Hospital Literature Index and Cambridge Scientific Abstracts.

For reprints of this article, contact the *Wisconsin Medical Journal* at 866.442.3800 or e-mail wmj@wismed.org.

© 2004 Wisconsin Medical Society