Atrophic Vaginitis

Concordance and Interpretation of Slides in the College of American Pathologists Cervicovaginal Interlaboratory Comparison Program in Gynecologic Cytopathology

Barbara A. Crothers, DO; Christine N. Booth, MD; Teresa M. Darragh, MD; Marilee M. Means, PhD, SCT (ASCP), Ly Ma, MD; Rhona J. Souers, MS; Nicole Thomas, CT (ASCP), MPH; Ann T. Moriarty, MD

• ***Context.***—**Atrophic vaginitis is a commonly reported subset of Papanicolaou test results that are negative for intraepithelial lesion or malignancy, but interpretive criteria overlap with atrophic changes and other entities, hindering concordance among observers.**

Objectives.—**To report on the participant concordance from 2000 to 2009 in the College of American Pathologists Interlaboratory Comparison Program in Gynecologic Cytopathology, with a reference interpretation of atrophic vaginitis, and to investigate cytologic features of good and poorly performing slides to identify criteria useful in the interpretation of atrophic vaginitis.**

Design.—**We summarized 18 302 responses from the program for slides with a reference interpretation of atrophic vaginitis. We randomly selected 18 Papanicolaou test results (3 conventional, 4 SurePath, and 11 ThinPrep) from good and poor performers for prospective, blinded criteria scoring for the following features: abundance of neutrophils, more than 100 degenerating parabasal cells, more than 25% necrotic background, more than 100 pseudoparakeratotic cells, and the presence of stripped or streaked nuclei, histiocytes, and superficial or intermediate squamous cells.**

Results.—**Most Papanicolaou test results (>90%) with a specific reference interpretation of atrophic vaginitis were categorized as negative. Cytotechnologists are more likely than pathologists are to label it negative for intraepithelial lesion or malignancy (NILM) and are equally likely to mistake it for a high-grade lesion. Degenerating parabasal cells, pseudoparakeratosis, and necrotic background are associated with atrophic vaginitis ($P = .001$) on Papanicolaou. Abundant neutrophils (>100 per ×400 field) are also significantly correlated ($P = .01$).**

Conclusions.—**Exact concordance to atrophic vaginitis is less than 90%. Most of the discrepancies are negative results for intraepithelial lesion or malignancy. Advanced atrophic features are as significant as neutrophils are to the interpretation of atrophic vaginitis.**

(*Arch Pathol Lab Med.* 2012;136:1332–1338; doi: 10.5858/arpa.2011-0441-CP)

The interpretation of *atrophic vaginitis* from Papanicolaou (Pap) tests is seemingly subjective, despite descriptive terminology criteria published by the *Bethesda System for Reporting Cervical Cytology*,[1] as evidenced by the historically poor performance on these slides in the College of American Pathologists (CAP) Interlaboratory Comparison Program in Gynecologic Cytopathology (PAP Education). Despite several years of circulating slides that have been previously verified by 3 board-certified anatomic pathologists as representing atrophic vaginitis, overall exact concordance to the reference interpretation has never been 90% or greater, for either individuals or laboratories, regardless of specimen preparation type (Table 1). The College requires 90% concordance on educational slides as the minimum necessary result for field validation of Pap slides for entry into the Gynecologic Cytology Proficiency Testing Program. If criteria for atrophic vaginitis are not reproducible, perhaps it should not be offered as a unique interpretation separate from *negative for intraepithelial lesion or malignancy* (NILM) in an educational program. To further define the extent of the problem, we reviewed participant results for all slides with a reference diagnosis of atrophic vaginitis from 2000 to 2009. We wanted to investigate whether there

Accepted for publication February 9, 2012.

From the Department of Pathology and Area Laboratory Services, Walter Reed Army Medical Center, Washington, DC (Dr Crothers); the Department of Anatomic Pathology, Cleveland Clinic Foundation, Cleveland, Ohio (Dr Booth); the Pathology Cytology Laboratory, Mount Zion Medical Center Clinic, University of California, San Francisco (Dr Darragh); the Department of Cytopathology, University of Kansas Medical Center, Kansas City (Dr Means); the Department of Pathology, Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, New Brunswick (Dr Ma); the Departments of Biostatistics (Ms Souers) and Surveys (Ms Thomas), College of American Pathologists, Northfield, Illinois; and the Department of Pathology, AmeriPath Indiana, Indianapolis (Dr Moriarty).

Dr Crothers is now with the Department of Pathology and Area Laboratory Services, Walter Reed National Military Medical Center, Bethesda, Maryland.

The authors have no relevant financial interest in the products or companies described in this article.

The opinions or assertions contained herein are the private views of the authors and do not reflect the official policy of the Department of the Army, Department of Defense, or US government.

Reprints: Barbara A. Crothers, DO, Department of Pathology and Area Laboratory Services, Walter Reed National Military Medical Center, 8901 Wisconsin Ave, Bethesda, MD 20889-5600 (e-mail: Barbara.crothers@med.navy.mil).

Table 1. Participant Answers to the Exact Reference Interpretation of Atrophic Vaginitis for Years 2000–2009

Preparation Type (Total Responses)[a]	Participant Interpretation[b]			
	Atrophic Vaginitis, No. (%)	NILM, No. (%)	Unsatisfactory, No. (%)	HSIL+, No. (%)
Conventional (n = 5305)				
Pathologist	1932 (36.4)	553 (10.4)	61 (1.1)	102 (1.9)
Cytotechnologist	1556 (29.3)	751 (14.2)	44 (0.8)	60 (1.1)
ThinPrep (n = 11 064)				
Pathologist	3216 (29.1)	1230 (11.1)	263 (2.4)	209 (1.9)
Cytotechnologist	2794 (25.3)	2097 (19.0)	343 (3.1)	181 (1.6)
SurePath (n = 1933)				
Pathologist	518 (26.8)	202 (10.5)	8 (0.4)	16 (0.8)
Cytotechnologist	599 (31.0)	496 (25.7)	12 (0.6)	24 (1.2)

Abbreviations: HSIL+, high-grade squamous intraepithelial lesion or cancer; NILM, negative for intraepithelial lesion or malignancy.

[a] Not all responses are tabulated. Participants also chose other interpretations not identified in these categories.
[b] Percentages are based upon the total number of responses (n) for the reference interpretation.

were slide characteristics that defined better performance, so that we could refine selection of slides into the program. Furthermore, we wanted to determine what characteristics of atrophic vaginitis, if any, were statistically significant for making that interpretation.

Atrophic vaginitis has been recognized as a diagnostic and clinical entity since the inception of cervical cancer screening, which serendipitously emerged through the investigation of hormonal influences on the cervix and vaginal vault. When estrogen levels decrease in women, cervical and vaginal epithelium thins. Squamous epithelial maturation to superficial squamous cells decreases, and with prolonged low-estrogen states[2] (eg, certain birth control medication, pregnancy, postpartum period, menopause, or prepubertal state), the epithelium ceases to produce superficial and intermediate squamous cells, leaving only parabasal and basal cell populations lining the lower gynecologic tract. This progression is usually gradual. Depending on a woman's hormonal milieu, it may be years before the seminal features of ''vaginitis'' are imposed on the developing atrophic state, accounting for the spectrum of atrophic changes usually evident on Pap tests in these populations. With fewer superficial and intermediate squamous cells, there are more parabasal cells. Initially, those cells are exfoliated as single cells or small groups of metaplastic-appearing cells, but with advanced atrophy, the epithelium becomes so thin that large surface fragments of parabasal cells exfoliate and may predominate. Because the epithelium is thin, dry, and inactive, it may be more susceptible to injuries, such as abrasion and infection, which, in turn, incite inflammatory changes: neutrophils, histiocytes, and degenerated cells. These physiologic changes may result in clinical symptoms such as dyspareunia, vaginal dryness and itching, painful urination, or painful abrasions and mucosal bleeding. Traditionally, the presence of marked, acute inflammation has been suggested as the hallmark of vaginitis on Pap tests. However, this more likely represents a natural consequence of a progression of events that occur because of estrogen depletion, rather than a truly morbid disease state. Previous studies have been inconsistent in determining a correlation between atrophic vaginitis and symptoms of atrophy.[3] Cytologic criteria for *atrophy* include predominately parabasal and basal cells, with few or no superficial or intermediate squamous cells; background debris from cellular degeneration; nuclear pyknosis; cellular apoptosis; and possibly, histiocytes and giant cells. Nuclear features such as enlargement, hyperchromasia, and nuclear membrane irregularities can mimic dysplasia.[4] The Bethesda System (TBS) for reporting cervicovaginal cytology describes *atrophic vaginitis* as ''atrophy with inflammation,'' suggesting that the above criteria, with the addition of the presence of neutrophils, distinguishes atrophic vaginitis from atrophy.[1] We wondered whether the amount of inflammation or other features were the primary cytologic expression of atrophic vaginitis. If clinicians do not routinely treat a Pap test interpretation of atrophic vaginitis,[5,6] and if atrophy is cytologically indistinguishable from atrophic vaginitis, then perhaps, that interpretation should not be separately reported, particularly if the cytologic criteria for atrophic vaginitis are not reproducible.

MATERIALS AND METHODS

We reviewed 18 302 participant responses to 717 Pap slides from 2000 to 2009 with the reference interpretation of atrophic vaginitis. The slides, part of the PAP Education program, were all reference-validated by 3 board-certified, anatomic pathologists and one cytotechnologist in accordance with the guidelines for slide submission into the program before release for educational purposes. Laboratories that submit slides to the program for consideration are instructed to select slides of high quality (to include good staining, preparation, and cover-slipping), which clearly represent a single diagnostic entity. For purposes of selection into the program, a slide would not be submitted as both atrophic vaginitis and NILM; both the submitting laboratory and the reviewers would have to make a choice between those interpretations. Reviewers for the CAP Cytopathology Committee confirm slide quality and the viability of the reference interpretation using TBS criteria. We excluded data from the last quarter of 2005 because it was part of a pilot mailing for the Gynecologic Cytology Proficiency Test program, and all of those slides (n = 7) had been field validated with 90% or greater interpretive concordance. We summarized the participant interpretations to the exact reference interpretations of atrophic vaginitis. (Table 1). We ran a 2-part analysis of participant responses for 3 rates: (1) concordance to a general negative category (that includes NILM, atrophic vaginitis, infectious agents, and other benign interpretations) (Table 2); (2) NILM rate for a reference interpretation of atrophic vaginitis (Table 3); and (3) high-grade squamous intraepithelial lesion or cancer (HSIL+) rate (Table 4). We used the Pearson χ^2 test to test the rate differences by type of specimen preparation (ThinPrep [Hologic Inc, Bedford, Maryland], SurePath [Becton-Dickinson Diagnostics, Inc, Franklin Lakes, New Jersey], and conventional Pap tests)

Table 2. Participant Concordance to the Negative Category for a Reference Interpretation of Atrophic Vaginitis			
Factor	**Responses, No.**	**Concordance Rate, No., %**	***P* Value**
Preparation type[a]			
SurePath	1933	1869 (96.7)	<.001
Conventional	5305	5003 (94.3)	
ThinPrep	11 064	9980 (90.2)	
Reader type			
Cytotechnologist	9313	8596 (92.3)	.44
Pathologist	8989	8261 (91.9)	

[a] Preparation type comparison: SurePath–conventional, $P < .001$; SurePath–ThinPrep, $P < .001$; conventional–ThinPrep, $P < .001$.

and participant type (pathologist versus cytotechnologist). Factors significant at the $P = .01$ level were included in a nonlinear mixed model with a repeated measures component to model the slide factor correlation structure. We also included the interaction terms between those factors. A P value of .05 was used as the significance level for the nonlinear mixed model. We used least-square means with a Bonferroni adjustment to compare multiple specimen types (conventional slides, ThinPrep, and SurePath slides) with one another for concordance to the negative category and to NILM and HSIL$^+$ participant interpretations to determine whether there was a significant difference between preparation types for these interpretations.

To investigate the significance of specific cytologic features of atrophic vaginitis in good and poorly performing slides, 2 of us (R.J.S. and N.T.) randomly selected 9 good performers and 9 poor performers from available slides from the set of atrophic vaginitis slides for blinded review. The selected slides had at least 10 participant responses from the program. We defined *good performers* as those slides with greater than 90% concordance for overall performance in the program and *poor performers* as slides that had less than 18% concordance. The rest of the set was in circulation as educational slides. Of the 18 selected slides, 3 (17%) were conventional, 4 (22%) were SurePath, and 11 (61%) were ThinPrep. Three board-certified pathologists with additional qualifications in cytopathology (C.N.B., B.A.C., and T.M.D.), one senior pathology resident (L.M.), and one experienced cytotechnologist (M.M.M.) independently reviewed and graded each slide of the 18 for the presence of the following characteristics: any histiocytes, stripped nuclei or nuclear streaks, intermediate and/or superficial squamous cells, neutrophils (0–100 per ×400 field, 101–1000 per ×400 field, or >1000 per ×400 field), degenerating parabasal cells greater than 100, background debris involving more than 25% of the cellular area, and pseudoparakeratotic cells greater than 100. We arbitrarily selected the numerical criteria as being generally easy to quantify at the magnifications specified because the objective field can be broken into 4 quadrants, allowing cellular components to be counted. Criteria for atrophic vaginitis were chosen as described in *The Bethesda System for Reporting Cervical Cytology*[1] and included features of atrophy with the addition of acute inflammation to define *vaginitis*. We subdivided acute inflammation into 3 categories to investigate whether the amount of inflammation influenced the interpretation. The reviewers did not discuss the criteria for morphologic features before review but were instructed to follow the guidelines given for evaluating the presence, absence, or amount of each feature and to record those data on the worksheet. Each reviewer was blinded to the results of other reviewers and to the historic performance of the slide but was able to recognize the preparation type during the review. All reviewers were aware that the reference interpretation for all slides was *atrophic vaginitis*. There was no adjudication process for disagreements in rating slides. We combined and evaluated the graded results using Fisher exact test, with a significance level of .05 for these tests, as shown in Table 5.

RESULTS

In the PAP Education program, both cytotechnologists and pathologists appropriately placed atrophic vaginitis into a negative category (Table 2) without a significant difference in their categorization. The negative category includes atrophic vaginitis, follicular cervicitis, reparative changes, and infectious agents, as well as NILM Pap tests. Cytotechnologists were significantly more likely to designate an atrophic vaginitis slide as NILM than were pathologists (Table 3), but there was no difference between the 2 groups for overcalling HSIL$^+$ on an atrophic vaginitis slide (Table 4). Preparation type had an influence on a mistaken HSIL$^+$ interpretation (Table 4) ($P = .05$), but that was significant only when comparing SurePath with ThinPrep ($P = .001$) and not when comparing either liquid-based preparation with a conventional Pap test. ThinPrep slides of atrophic vaginitis were more likely to be overcalled HSIL$^+$ (3.5%) than were SurePath slides of atrophic vaginitis (2.1%; Table 4). For concordance to a negative category and participant interpretation of NILM, there was a significant difference between all the preparation types ($P = .001$ for each comparison), implying that the specimen preparation influences those interpretations (Table 2). ThinPrep Pap slides were more often designated unsatisfactory than were the other 2 preparations (Table 1). For the reference interpretation of atrophic

Table 3. Participant Response of Negative for Intraepithelial Lesion or Malignancy (NILM) for a Reference Interpretation of Atrophic Vaginitis			
Factor	**Responses, No.**	**NILM Rate**	***P* Value**
Preparation type[a]			
SurePath	1933	698 (36.1)	<.001
Conventional	5305	1305 (24.6)	
ThinPrep	11 064	3330 (30.1)	
Reader type			
Cytotechnologist	9313	3343 (35.9)	.001
Pathologist	8989	1987 (22.1)	

[a] Preparation type comparison: SurePath–conventional, $P < .001$; SurePath–ThinPrep, $P < .001$; conventional–ThinPrep, $P < .001$.

Table 4. Participant Response of High-Grade Squamous Intraepithelial Lesion or Cancer (HSIL+) for a Reference Interpretation of Atrophic Vaginitis

Factor	Responses, No.	HSIL+ Rate	*P* Value
Preparation type[a]			
SurePath	1933	41 (2.1)	.05
Conventional	5305	165 (3.1)	
ThinPrep	11 064	387 (3.5)	
Reader type			
Cytotechnologist	9313	261 (2.8)	.08
Pathologist	8989	324 (3.6)	

[a] Preparation type comparison: SurePath–conventional, $P = .13$; SurePath–ThinPrep, $P = .001$; conventional–ThinPrep, $P = .43$.

vaginitis, there was a significant difference between slide performance for preparation types in each participant interpretative category ($P < .001$) but not between pathologists and cytotechnologists. Slides that performed well were different from poor performers in that they contained more degenerating parabasal cells, necrotic background, and pseudoparakeratotic cells, as well as had more inflammation and stripped nuclei and/or nuclear streaks (Table 5). The presence of intermediate and/or superficial squamous cells or histiocytes and/or giant cells (representing 2 ends of the spectrum of atrophy) did not show statistical significance in determining good slide performance.

The amount of acute inflammation (neutrophils) present was significant to the interpretation of atrophic vaginitis ($P = .01$), but both good and poorly performing slides showed few (<100 per ×400 field) neutrophils. Marked, acute inflammation, as defined by more than 1000 neutrophils in a ×400 field, was uncommon in both good and poor slide performers (11.1% and 2.2%, respectively). Poorly performing slides had fewer neutrophils present (24.4% of slides had >100 in a ×400 field) than did good performers (53.3% had >100 in a ×400 field). We showed a statistically significant difference for the number of neutrophils with poor performers having fewer neutrophils (75.6% versus 46.7% for 0–100 neutrophils per ×400). The presence of degenerating parabasal cells, necrotic background, and pseudoparakeratosis were equally important in arriving at a correct interpretation of atrophic vaginitis ($P < .001$).

COMMENT

For several years, we have observed that slides representing atrophic vaginitis in the CAP PAP Education program did not perform well when correlated to the exact reference interpretation, although most of these slides were correctly placed in the negative reference category by participants. We designed this study to investigate whether the presence of certain cytologic features, especially acute inflammation, resulted in a more accurate interpretation of atrophic vaginitis, recognizing that the features of atrophy and atrophic vaginitis overlap. The Bethesda System suggests that it is primarily the presence of acute inflammation that distinguishes atrophic vaginitis as separate from atrophy, but our data show that although neutrophils are important to the interpretation, they are not always present in large amounts in either good or poorly performing slides. The gold standard used to determine the reference interpretation was validation of the original submitter's interpretation of atrophic vaginitis by 3 board-certified anatomic pathologists and 1 cytotechnologist, the CAP Cytopathology Committee's usual method for approving slide entry into the educational program. Submitting laboratories are asked to select only one interpretation for submitted slides and must choose between NILM and atrophic vaginitis in the negative reference category. Slide reviewers are instructed to reject slides from the program if they do not represent classic examples of the interpretation in accordance with TBS criteria. Additionally, program participants are expected to choose between NILM and atrophic vaginitis as choices listed under the negative category, along with other benign entities. In the PAP Educational program, these entities include NILM, *Trichomonas vaginalis*, herpes simplex virus, *Candida* species, atrophic vaginitis, follicular cervicitis, and repair. The instructions for the program do not specifically tell the participants to choose the ''best'' answer. However, only one choice is permitted in the negative category, including NILM, so participants must determine if the cytologic features on the slide warrant a more specific interpretation. Participants are given credit for the interpretation if it

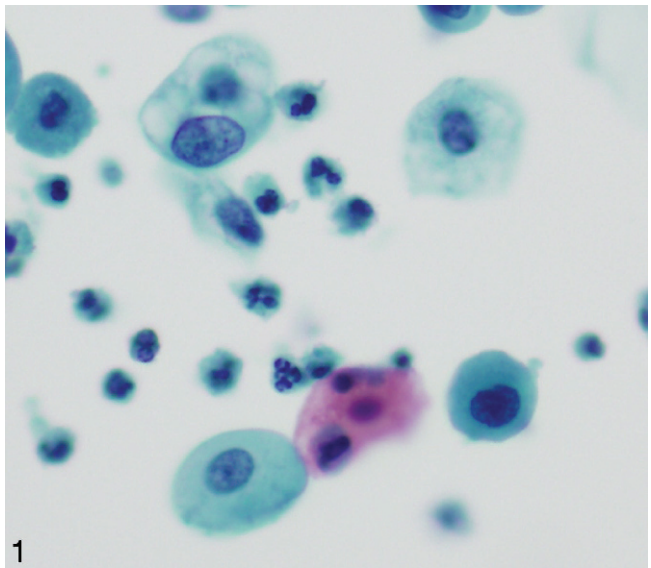
Table 5. Cytologic Characteristics of Good and Poor Slide Performance with a Reference Interpretation of Atrophic Vaginitis

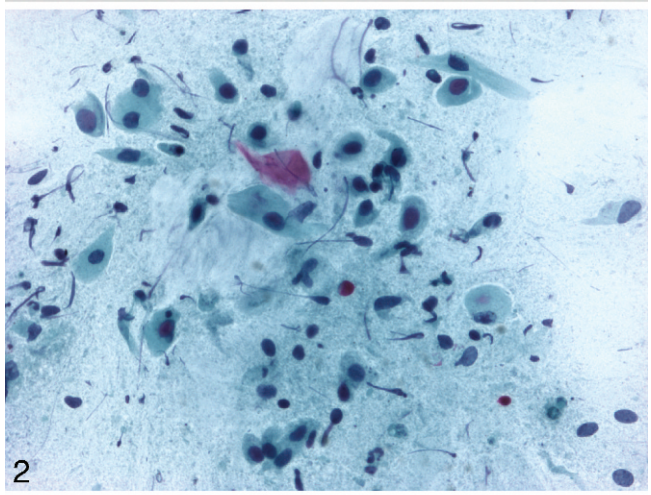
Attribute	Slide Performance Type		*P* Value
	Good, (n = 45)	Poor, (n = 45)	
Neutrophils, per ×400 field			.01
0–100	21 (46.7)	34 (75.6)	
101–1000	19 (42.2)	10 (22.2)	
>1000	5 (11.1)	1 (2.2)	
Degenerating parabasal cells >100	44 (97.8)	21 (46.7)	<.001
Necrotic background >25%	36 (80.0)	18 (40.0)	<.001
Pseudoparakeratosis >100 cells	28 (62.2)	9 (20.0)	<.001
Histiocytes/giant cells present	24 (53.3)	14 (31.1)	.05
Stripped nuclei/nuclear streaks present	30 (66.7)	17 (37.8)	.01
Intermediate/superficial squamous cells present	12 (26.7)	20 (44.4)	.12

falls within the same general category, even if they do not select the target interpretation.

Our findings show that ThinPrep atrophic vaginitis slides are significantly more likely to be deemed unsatisfactory. Atrophic epithelium, because it is thin and composed of parabasal cells, may not shed many single cells. If the epithelium is inflamed, neutrophils may be the primary type of cell procured for a Pap test, resulting in scant squamous cellularity and, hence, an unsatisfactory Pap test. In a 2006 CAP survey[7] of reporting rates for unsatisfactory gynecologic slides, conventional slides had the highest 95th percentile (5.9%) and ThinPrep had the highest median (1.1%) of unsatisfactory slides. In the same study,[7] 79% used minimum cellularity criteria, as outlined by TBS, to evaluate atrophic slides. It is not clear from our study why participants placed more ThinPrep slides with atrophic vaginitis into the unsatisfactory category than they did other preparations. Certain types of lubricant used during a speculum examination can clog filter pores during preparation of ThinPrep slides, and postmenopausal women often require more lubricant during an examination than do premenopausal women, thereby resulting in reduced cellularity. Conventional slides in our study may have performed better regarding cellularity because paucicellular conventional slides would be rejected from entry into the program as a reference interpretation for atrophic vaginitis.

Approximately one-third of participants designate slides with the reference interpretation of *atrophic vaginitis* as *NILM*, and the type of preparation influenced that decision, with conventional slides having the lowest (24.6%) and SurePath slides having the highest (36.1%) rates of discordant categorization. Perhaps this is not surprising if participants rely heavily upon the presence of neutrophils to recognize atrophic vaginitis as required by TBS. Of the preparation types, we expect conventional slides to contain the most neutrophils because health care providers collect material from the cervix and smear it directly onto a slide. ThinPrep Pap tests use a filter preparation method that allows for the passage of most neutrophils through the membrane but leaves clumps of neutrophils on the slide. SurePath Pap tests rely on gravity sedimentation of cells, so that heavier cells, such as squamous and glandular cells, settle onto the slide, but lighter hematopoietic cells (the ''buffy coat'') are removed from the slide surface, leaving primarily epithelial cells. The results of our study somewhat support that assumption because slides with fewer than 100 neutrophils performed poorly when compared with slides that had more than 100 neutrophils in a ×400 field. However, we rated most of the slides in our study (both good and poor performers) as having fewer than 1000 neutrophils per ×400 field, and nearly half of the good performing slides had fewer than 100 neutrophils per ×400 field, suggesting that participants rely on more than just acute inflammation to make the interpretation of atrophic vaginitis. Indeed, the presence of degenerating parabasal cells ($P = .001$), necrotic background ($P = .001$), and pseudoparakeratosis ($P = .001$) were also significant in recognizing atrophic vaginitis. The last 3 features are well-documented cytologic features of advanced atrophy (Figures 1 through 3), and these findings suggest that participants recognize features of severe atrophy as critical to the diagnosis of atrophic vaginitis, even in the absence of abundant neutrophils. Less than half of poor performing





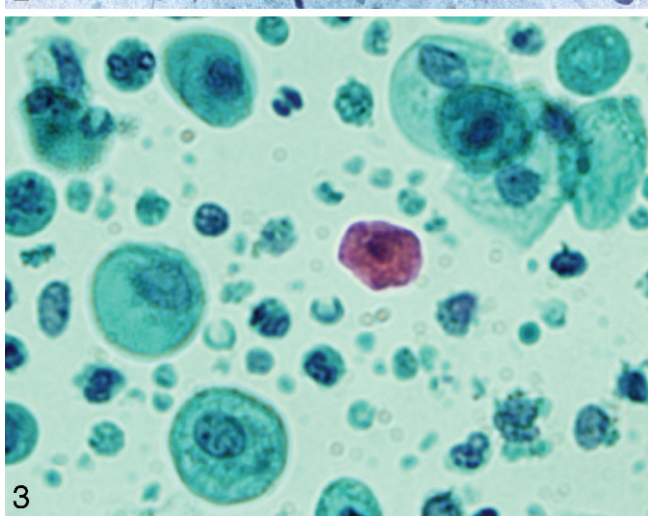


Figure 1. *Degenerating parabasal cells with pseudoparakeratosis. ThinPrep Papanicolaou test (Papanicolaou stain, original magnification ×400).*

Figure 2. *Atrophic vaginitis showing typical necrotic background and streaked nuclei. Conventional Papanicolaou test (Papanicolaou stain, original magnification ×200).*

Figure 3. *Pseudoparakeratosis. SurePath Papanicolaou test (Papanicolaou stain, original magnification ×400).*

slides showed these 3 features. Of interest, the presence of stripped parabasal nuclei and/or nuclear streaks was as significant ($P = .01$) to the interpretation of atrophic vaginitis as was the presence of neutrophils. Both of these are also features of advanced atrophy. Our findings indicate that the cytologic features of advanced atrophy are good predictors of slide performance for atrophic vaginitis in the CAP Pap Education program.

There are several limitations of this study. First, TBS is silent on the number of neutrophils necessary for an interpretation of atrophic vaginitis, making that criterion subjective. We arbitrarily selected categories for few, moderate, and abundant neutrophils, but our ranges may have been too wide, so that few slides were rated as having abundant neutrophils. Some participants may regard advanced atrophy as atrophic vaginitis regardless of the number of neutrophils present and do not make the interpretation without features of advanced atrophy. After all, atrophic vaginitis, by definition, must include features of atrophy. Additionally, there is no reliable gold standard for the interpretation of atrophic vaginitis. Benign Pap tests do not have to be biopsy-proven for submission into the program. The CAP Cytopathology Committee does allow atrophic slides into the program under the reference interpretation of NILM, but it is not possible to identify those slides because *atrophy* is not a choice on the slide submission or participant's menu. During the many years of gamesmanship in the program, participants may conclude that only slides with advanced atrophic features are regarded as atrophic vaginitis. The program is an artificial practice environment that does not allow all the choices made in actual practice, such as *atypical squamous cells of undetermined significance*. In the presence of uncertainty about criteria that constitute an interpretation, participants may choose *NILM*, as opposed to a more-precise interpretation, to be correct. One possibility for the many slides of atrophic vaginitis being assigned to the more-generic NILM interpretation might be that participant's laboratories do not report atrophic vaginitis separately, and the participants are following their institution's laboratory procedures for reporting Pap tests. Even though all of the choices in the benign category are accepted as NILM in actual practice, many laboratories choose to separately report only infectious agents. Participants may also identify slides as atrophic vaginitis more often than they would in actual practice because it is a choice on the menu. It is not clear whether all participants consider atrophy and atrophic vaginitis as 2 separate interpretations—some individuals may label all slides showing atrophy as atrophic vaginitis. Participants may decide that recognizing that a slide is negative for dysplasia is the primary purpose of the educational exercise because atrophic changes mimic squamous intraepithelial lesions.

In this study, some participants confused HSIL$^+$ with atrophic vaginitis, most often on ThinPrep slides and least often on SurePath slides. We hypothesize that this may be due to the increased tendency of groups in ThinPrep slides to round up, overlap, and shrink, making evaluation of cell size and individual nuclei more difficult. Additionally, both atrophic parabasal cells and HSIL involving endocervical glands show the presence of inconspicuous nucleoli. Some participants may interpret hyperchromatic, crowded, atrophic parabasal cells as the hyperchromatic, crowded groups of HSIL. Renshaw et al[8] demonstrated that HSIL in CAP educational ThinPrep Pap slides is mistaken for a glandular lesion by 75% of expert cytopathologists from the Cytopathology Resource Committee. Notably, participants in the program do not have the choice of an atypical interpretation, such as atypical squamous cells, cannot exclude high-grade squamous intraepithelial lesion. This forces participants to choose between a negative and high-grade interpretation, when they might have preferred an atypical interpretation. A recent study by Patton et al[9] determined that age is a key factor for determining the positive predictive value of a diagnosis of atypical squamous cells, cannot exclude high-grade squamous intraepithelial lesion, with a low predictive value in postmenopausal women as opposed to other women with low-estrogenic states (postpartum, contraceptive use, and pregnancy). They also recognized that cytologic criteria alone do not appear to be reliable in distinguishing between atrophic changes and possible HSIL. Selvaggi[10] determined that the background necrosis in atrophic Pap tests could not be reliably distinguished from that of squamous cell carcinoma—only the detection of malignant cells was reliable for differentiating between them. Cytologists recognize that atrophic changes can be difficult to distinguish from significant lesions because of the lack of epithelial maturation. Even though the proportion of interpretations as atypical squamous cells of undetermined significance is usually lower in postmenopausal women, those patients may have a higher atypical squamous cell to squamous intraepithelial lesion ratio and, subsequently, more false-positive Pap tests, highlighting the difficulty of reporting squamous intraepithelial lesions in atrophy.[11]

A surprising finding was that marked acute inflammation (as defined by >1000 neutrophils in a ×400 field) was uncommon, even for atrophic vaginitis slides that performed well. By definition, *vaginitis* suggests inflammation of the vaginal vault. We should disclose that, to our knowledge, none of these slides was purposefully collected from the lateral vaginal wall, and we did not distinguish between vaginal and cervical Pap tests in this study. However, cytologists commonly use the term *atrophic vaginitis* for cervical Pap tests (as opposed to the more correct term, *atrophic cervicitis*), and it is applied to both cervical and vaginal Pap tests. If accurate identification of *atrophic vaginitis* does not depend on the presence of abundant neutrophils, then it may be reasonable to exclude this Pap test interpretation from the menu and require that participants identify *atrophic change*, as opposed to *atrophic vaginitis*. This may result in more reproducible interpretations. More to the point, does a health care provider rely on the Pap test interpretation of atrophic vaginitis or atrophic changes to direct patient therapy? Capewell et al[3] investigated the clinical usefulness of vaginal smears in the diagnosis of atrophy in elderly women and found that inflammation was present in only 44% of smears and that the degree of inflammation was unrelated to clinical symptoms of atrophy or cytologic atrophy. Furthermore, although most Pap smears (60%) showed severe atrophic changes, there was no association with clinical symptoms or age of the patient. They concluded that vaginal dryness, low parity, and low body weight were better predictors of clinically significant atrophic vaginitis than were Pap smear findings. Greendale et al[12] also found no correlation between the degree of inflammation on Pap smear or on physical examination

and the presence of clinically significant vaginal atrophy. Repse-Fokler et al[13] support vaginal atrophy as correlating best with lean body mass and low estradiol, as opposed to age or time lapse since menopause.

Few health care providers use Pap tests for the evaluation of vaginal atrophy. Pap tests for hormonal status fell out of favor clinically more than 20 years ago with the advent of more reliable tests, such as serum hormone levels. Review of the gynecologic literature shows little evidence that clinicians distinguish between atrophic vaginitis and atrophic changes on Pap tests. Pap tests are only infrequently used in the evaluation of a woman with suspected vaginal atrophy, an evaluation that relies primarily on physical examination and clinical symptoms.[14] Measurement of vaginal pH and serum estradiol levels are also sometimes used. Clinical relief of symptoms is the usual measure of the effectiveness of therapy. Even though several studies have demonstrated epithelial maturation on Pap tests when women are treated with vaginal or oral estrogen preparations,[15–17] there are no recommendations to repeat Pap tests to follow the effectiveness of therapy. A major consideration by the cytopathology community for identifying advanced atrophy or atrophic vaginitis is to document that the cellular features observed (changes that can mimic more severe diseases, such as squamous intraepithelial lesion and cancer) are interpreted as benign changes because of atrophy. The Clinical Laboratory Improvement Amendments of 1988 requires a hierarchic review of reactive and reparative changes.[18] Laboratories using the term *atrophic vaginitis* might be expected to secondarily review these slides before finalizing the report because the presence of inflammation suggests reactive changes, whereas laboratories using NILM for atrophic slides would not. However, our study and others would suggest that it is unnecessary and unreliable to discriminate between atrophic vaginitis and atrophic change on Pap tests. These changes are part of a spectrum of atrophy that has no clinical significance. It might be preferable for individuals finalizing Pap tests to either report these slides as NILM or simply indicate that an atrophic pattern is present if hierarchic review is part of the laboratory policy for these slides. If atrophic vaginitis slides in the CAP Pap Education program were reassigned to a new reference interpretation of atrophic change, and participants were asked to distinguish between them and NILM in the future, they might show greater interpretive concordance than indicated by our investigation. Including atrophic vaginitis, or atrophic changes, as a separate entity in the CAP Pap Education program continues to have value to educate participants on the features of atrophy that can mimic HSIL.

References

1. Solomon D, Nayar R, eds. *The Bethesda System for Reporting Cervical Cytology: Definitions, Criteria, and Explanatory Notes*. 2nd ed. New York, NY: Springer-Verlag; 2004.
2. Pandit L, Ouslander JG. Postmenopausal vaginal atrophy and atrophic vaginitis. *Am J Med Sci*. 1997;314(4):228–231.
3. Capewell AE, McIntyre MA, Elton RA. Post-menopausal atrophy in elderly women: is a vaginal smear necessary for diagnosis? *Age Ageing*. 1992;21(2):117–120.
4. DeMay RM. *The Art and Science of Cytopathology, Volume I—Exfoliative Cytology*. Hong Kong, China: ASCP Press; 1996.
5. Castelo-Branco C, Cancelo MJ, Villero J, Nohales F, Juliá MD. Management of post-menopausal vaginal atrophy and atrophic vaginitis. *Maturitas*. 2005; 52(suppl 1): S46–S52.
6. Lynch C. Vaginal estrogen therapy for the treatment of atrophic vaginitis. *J Women's Health (Larchmt)*. 2009;18(10):1595–1606.
7. Moriarty AT, Clayton AC, Zaleski S, et al. Unsatisfactory reporting rates: 2006 practices of participants in the College of American Pathologists Interlaboratory Comparison Program in Gynecologic Cytology. *Arch Pathol Lab Med*. 2009;133(12):1912–1916.
8. Renshaw AA, Mody DR, Wang E, et al. Hyperchromatic crowded groups in cervical cytology—differing appearances and interpretations in conventional and ThinPrep preparations: a study from the College of American Pathologists Interlaboratory Comparison Program in Cervicovaginal Cytology. *Arch Pathol Lab Med*. 2006;130(3):332–336.
9. Patton AL, Duncan L, Bloom L, Phaneuf G, Zafar N. Atypical squamous cells, cannot exclude a high-grade squamous intraepithelial lesion and its clinical significance in postmenopausal, pregnant, postpartum, and contraceptive-use patients. *Cancer*. 2008;114(6):481–488.
10. Selvaggi SM. Atrophic vaginitis versus invasive squamous cell carcinoma on ThinPrep cytology: can the background be reliably distinguished? *Diagn Cytopathol*. 2002;27(6):362–364.
11. Keating JT, Wang HH. Significance of a diagnosis of atypical squamous cells of undetermined significance for Papanicolaou smears in perimenopausal and postmenopausal women. *Cancer*. 2001;93(2):100–105.
12. Greendale GA, Zibecchi L, Petersen L, Ouslander JG, Kahn B, Ganz PA. Development and validation of a physical examination scale to assess vaginal atrophy and inflammation. *Climacteric*. 1999;2(3):197–204.
13. Repse-Fokter A, Takac I, Fokter SK. Postmenopausal vaginal atrophy correlates with decreased estradiol and body mass index and does not depend on the time since menopause. *Gynecol Endocrinol*. 2008;24(7):399–404.
14. Society of Obstetricians and Gynaecologists of Canada. SOGC clinical practice guidelines: the detection and management of vaginal atrophy. *Int J Gynaecol Obstet*. 2005;88(2):222–228.
15. Bateson DJ, Weisberg E. An open-label randomized trial to determine the most effective regimen of vaginal estrogen to reduce the presence of atrophic changes reported in postmenopausal cervical smears. *Menopause*. 2009;16(4): 765–769.
16. Bachmann G, Lobo RA, Gut R, Nachtigall L, Notelovitz M. Efficacy of low-dose estradiol vaginal tablets in the treatment of atrophic vaginitis: a randomized control trial. *Obstet Gynecol*. 2008;111(1):67–76.
17. Gupta S, Kumar N, Singhal N, Manektala U, Jain S, Sodhani P. Cytohormonal and morphological alterations in cervicovaginal smears of postmenopausal women on hormone replacement therapy. *Diagn Cytopathol*. 2006;34(10):676–681.
18. Centers for Disease Control and Prevention, Centers for Medicare and Medicaid Services, Department of Health and Human Services. Medicare, Medicaid, and CLIA programs; laboratory requirements relating to quality systems and certain personnel qualifications: final rule. *Fed Regist*. 2003; 68(16):3711. To be codified at 42 CFR §493.1274(e)(1).