Altered Recognition of Reparative Changes in ThinPrep Specimens in the College of American Pathologists Gynecologic Cytology Program

Tamela M. Snyder, MD; Andrew A. Renshaw, MD; Patricia E. Styer, PhD; Dina R. Mody, MD; Terence J. Colgan, MD; for the Cytopathology Resource Committee, College of American Pathologists

• ***Context.*—Previous studies have shown that the diagnosis of reparative changes in conventional smears in the College of American Pathologists Interlaboratory Comparison Program in Gynecologic Cytology is one of the least reproducible diagnoses. Indeed, the diagnosis of reparative changes consistently yields the highest false-positive rate of any negative for intraepithelial lesions and malignancy (NILM) cytodiagnostic category. It is unknown whether cytologists recognize reparative changes in ThinPrep specimens as well, or less often, as in conventional smears.**

***Objective.*—To assess and compare the ability of cytologists to recognize reparative changes in conventional and ThinPrep preparations.**

***Design.*—We compiled performance data from the College of American Pathologists Interlaboratory Comparison Program in Gynecologic Cytology from the 2000–2003 program years. More than 400 slides with a reference diagnosis of reparative changes met our study criteria, representing a total of 11 200 individual responses for conventional cases and 1155 individual responses for ThinPrep specimens. We evaluated the results of both individual and laboratory participants using 2 performance criteria: the false-positive discordancy rate and the exact match error rate (any response that does not exactly match the reference diagnosis of 120 [reparative changes]).**

***Results.*—Cases with a reference diagnosis of reparative changes made up 1.2% of all ThinPrep slides and 3.7% of all conventional slides in circulation. The false-positive discordancy rate of individual responses on educational slides for conventional smears was significantly higher than the corresponding false-positive discordancy rate for ThinPrep specimens (15.7% for conventional vs 7.1% for ThinPrep specimens, $P < .001$). Laboratory responses on educational conventional smears and ThinPrep slides showed a similar trend (14.2% for conventional smears vs 2.4% for ThinPrep slides, $P = .002$). The exact match error rate on educational conventional slides was 41.4% for individual responses, while on educational ThinPrep slides, the overall error rate was 57.5% ($P < .001$). For laboratory responses, the exact match error rate was 40.5% for educational conventional smears versus 58.9% for educational ThinPrep smears ($P < .001$). Characteristic features of reparative changes were identified in ThinPrep specimens.**

***Conclusions.*—In the College of American Pathologists Interlaboratory Comparison Program in Gynecologic Cytology, ThinPrep slides with a reference diagnosis of reparative changes have a lower false-positive discordancy rate than conventional slides. Responses to ThinPrep cases with a reference diagnosis of reparative change show a higher exact match error rate than conventional smears. Since reparative changes in gynecologic cytology are recognized as indicating an increased risk of significant lesions, the clinical significance of these altered patterns of recognition of reparative changes in ThinPrep specimens warrants further investigation.**

(*Arch Pathol Lab Med.* 2005;129:861–865)

Reparative changes, both typical and atypical, in conventional Papanicolaou tests are well described.[1–5] This cytodiagnosis is important, since studies have shown that repair and atypical repair are associated with an increased risk of squamous intraepithelial lesions as well as other clinically significant lesions.[6–9] Paradoxically, recent studies have also shown that the diagnosis of repair is one of the least reproducible interpretations in the Bethesda System.[10,11] In the College of American Pathologists (CAP) Interlaboratory Comparison Program in Gynecologic Cytology (PAP), the diagnosis of repair in conventional smears is consistently associated with the highest false-positive rate (lowest specificity) of any negative category in the program.

During a recent analysis of participant performance in the PAP program, it was suspected that, in contrast to reparative changes in conventional smears, responses to ThinPrep slides with reparative changes were associated

Accepted for publication March 15, 2005.

From the Department of Pathology, The Methodist Hospital, Houston, Tex (Drs Snyder and Mody); Department of Pathology, Baptist Hospital of Miami, Miami, Fla (Dr Renshaw); the College of American Pathologists, Northfield, Ill (Dr Styer); and the Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario (Dr Colgan).

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

Reprints: Terence J. Colgan, MD, Suite 600, Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Ave, Toronto, Ontario, Canada M5G 1X5 (e-mail: tcolgan@mtsinai.on.ca).

Table 1. Diagnostic Menu From the College of American Pathologists Interlaboratory Comparison Program in Cervicovaginal Cytology*

	Reference Diagnosis
000	Unsatisfactory
101	Negative for intraepithelial lesions and malignancy, not otherwise specified
111	Fungal organisms c/w *Candida*
113	*Trichomonas vaginalis*
115	Cellular changes c/w herpes
120	Reparative changes
121	Atrophic vaginitis
127	Follicular/lymphocytic cervicitis
201	LSIL
211	HSIL
221	Squamous cell carcinoma
225	Adenocarcinoma

* c/w indicates compatible/consistent with; LSIL, low-grade squamous intraepithelial lesion; and HSIL, high-grade squamous intraepithelial lesion.

with a relatively low false-positive rate (high specificity). To investigate this suspicion and the possible reasons for it, the data of respondent performance on slides of reparative changes in both conventional and ThinPrep specimens of the program for 4 years were reviewed.

MATERIALS AND METHODS

The PAP program is a quarterly, mailed, glass slide quality improvement program. The CAP Laboratory Accreditation Program requires that all laboratories evaluating gynecologic cytology enroll in the PAP program or an equivalent glass slide program. Cytology laboratories of all types participate, with the largest number (approximately 60%) being hospital laboratories. In addition, independent laboratories, federal and state government laboratories, university laboratories, and others (eg, those associated with a group practice or physician's office) participate.

Participants generously contribute slides to the program. All conventional smears are original preparations, whereas ThinPrep slides may be either original preparations or duplicate preparations. Submitted slides with a diagnosis of low-grade squamous intraepithelial lesion or higher must be biopsy confirmed. After receipt and accessioning into the program, the slides are reviewed by at least 3 experienced cytopathologists from the CAP Cytopathology Resource Committee. Before acceptance into the program, each slide must be judged to be of good technical quality and an excellent example of the reference diagnosis. All 3 reviewers must agree on the exact target diagnosis, and this diagnosis must agree with the submitted and biopsy diagnosis prior to accepting a slide for circulation into an educational set.

The PAP program consists of 5 glass slides of cervicovaginal material mailed 4 times per year to participating laboratories. The coded answer sheets have diagnostic menus using terminology modified from the Bethesda System. Referenced slides are placed into one of 3 selection series: the 000 for unsatisfactory slides; the 100 series for normal, infections, and reparative conditions; and the 200 series for epithelial cell abnormalities and carcinoma (Table 1).

Following acceptance by the Cytopathology Resource Committee, all slides commence circulation as educational (or unvalidated) slides. Subsequently, validation status is designated for slides that have met specific performance requirements. For all validated slides, the cases must have been reviewed by at least 20 participants and achieved a 90% level of agreement to the correct selection series (be it 000 or 100 or 200), with a minimum of 20 correct responses. The standard of error of this percentage must be, at most, 5%. During the study period, additional criteria were added for low-grade squamous intraepithelial lesion and 100 series slides in order to develop the best pool of validated slides. For all 100 series slides, cases must achieve at least a 50% assignment to the exact reference diagnosis; for a slide with a reference diagnosis of reparative changes, this would mean that 50% of the respondents select 120 (reparative changes) as their answer. For low-grade squamous intraepithelial lesion slides, cases must achieve at least a 70% concordance to the exact reference diagnosis.

Data from the 2000–2003 PAP program were used for this analysis. There were relatively few validated ThinPrep cases with a diagnosis of reparative change during this time period, so comparisons were performed using data from educational slides only for statistical analysis. The analysis included slides only if there were at least 5 responses per slide, tallied separately for individual responses and laboratory responses. A summary of the slide data for these responses to slides with a reference diagnosis of reparative changes are shown in Table 2.

Responses were analyzed at 2 levels of agreement with the reference diagnosis of the slide. In the first analysis, participant responses were examined with respect to their discordancy from the 100 series for the reference diagnosis of reparative changes. Hence, a discordant response placed a slide with a reference diagnosis of reparative changes slide in the 200 series (''false-positive''). (No responses in the 000 series were recorded.) In the second agreement analysis, the proportion of exact matches (ie, responses correctly identifying a slide to the exact reference diagnosis of reparative changes) was determined. In this exact match analysis, a response that did not exactly identify the slide with a reference diagnosis of reparative changes was labeled an exact match error. The distribution of responses for slides with a reference diagnosis of reparative changes for both types of preparations was identified.

Statistical analysis was performed using generalized linear models for binomial responses to compare differences in error rates between educational conventional and ThinPrep slides. This methodology was used to account for varying numbers of observations per slide. *P* values of all significance tests are reported in Tables 3 and 4.

Because the sample sizes were sometimes quite small (n = 24) for laboratory responses on ThinPrep slides (Tables 3 and 4), an approximate power calculation was performed using linear models fit to the logits of the error proportions. This calculation specified the observed difference in means as the size of the expected

Table 2. Summary of Data for Slides With a Reference Diagnosis of Reparative Changes for 2000–2003 College of American Pathologists Program Years

Slide Type*	No. of Slides Meeting Criteria for Laboratory Responses†	No. of Slides Meeting Criteria for Individual Responses‡	Average No. of Laboratory Responses per Slide	Average No. of Individual Responses per Slide
Conventional	311	394	9.5	28.4
ThinPrep	25	50	8.3	23.1

* Includes both educational and validated slides.
† Slides must have at least 5 laboratory responses to be included in study for laboratory responses.
‡ Slides must have at least 5 responses from either cytotechnologists or pathologists to be included in study for individual responses.

Table 3. Summary of False-Positive Discordancy Rates (ie, False-Positive Responses) for Both Conventional and ThinPrep Slides With a Reference Diagnosis of Reparative Changes

Type of Response	False-Positive Rate for Validated Conventional Slides, %*	False-Positive Rate for Educational Conventional Slides, %*	False-Positive Rate for Educational ThinPrep Slides, %*	*P* Value of Difference for Educational Slides
Laboratory	4.9 (n = 120)	14.2 (n = 191)	2.4 (n = 24)	.002
Individual	6.6 (n = 142)	15.7 (n = 252)	7.1 (n = 49)	<.001

* A false positive is any response in the 200 series.

Table 4. Summary of Exact Match Error Rates for Both Conventional and ThinPrep Slides With a Reference Diagnosis of Reparative Changes

Type of Response	Exact Match Error Rate for Validated Conventional Slides, %*	Exact Match Error Rate for Educational Conventional Slides, %*	Exact Match Error Rate for Educational ThinPrep Slides, %*	*P* Value of Difference for Educational Slides
Laboratory	27.4 (n = 120)	40.5 (n = 191)	58.9 (n = 25)	<.001
Individual	32.2 (n = 142)	41.4 (n = 252)	57.5 (n = 49)	<.001

* An exact match error is any response other than 120 (reparative changes).

Table 5. Distribution of Responses for Educational Slides With a Reference Diagnosis of Reparative Changes Showing Nonexact Matches for 2000–2003 College of American Pathologists Program Years, Conventional and ThinPrep Preparations

	Individual Responses		Laboratory Responses	
Diagnostic Category	Conventional, %	ThinPrep, %	Conventional, %	ThinPrep, %
001 Unsatisfactory	1.2	1.3	1.3	<1
101 NILM-NOS*	18.7	43.1	20.1	49.3
111 Fungal	<1	2.5	<1	5.3
113 *Trichomonas*	2.8	1.6	2.5	<1
115 Herpes	<1	<1	<1	<1
121 Atrophic vaginitis	1.2	1.3	1.4	1.9
127 Follicular cervicitis	<1	<1	<1	<1
120 Repair (exact match)	58.6	42.5	59.5	41.1
All 200 level	15.7	7.1	14.2	2.4

* *P* values of differences between conventional and ThinPrep rates for individual and laboratory responses are both <.001. Differences in other categories were not formally evaluated. NILM-NOS indicates negative for intraepithelial lesions and malignancy, not otherwise specified.

value under the alternative hypothesis and used the estimated pooled standard error estimate as the variance. With a type I error rate of 5%, the power of the tests ranged from 74% for the laboratory responses for exact match error rates to 98% for the individual responses for exact match error rates. Results for laboratory and individual false-positive discordancy rates were 96% and 86%, respectively. The power for the laboratory exact match error rates is relatively low, both because of the smaller sample size of ThinPrep slides and the larger slide-to-slide variability for the exact match error rates.

RESULTS

A total of 11 200 individual participant responses for conventional slides and 1155 individual participant responses for ThinPrep slides with a reference diagnosis of reparative change were available for analysis in this study from more than 400 slides. The data included an additional 2951 laboratory responses for conventional cases and 207 laboratory responses for ThinPrep slides. Slides with a reference diagnosis of reparative changes made up 1.2% of all ThinPrep slides and 3.7% of all conventional smears in circulation during the study period. A summary of the response data for slides with a reference diagnosis of reparative changes for the 2000–2003 PAP years is given in Table 2.

Table 3 shows the discordancy rates (ie, response of a 200 series diagnosis) for both ThinPrep and conventional slides. For individual responses on educational slides, conventional preparations were associated with a significantly higher false-positive rate (15.7%) than ThinPrep smears (7.1% false-positive rate, $P < .001$). Laboratory responses showed the same trend. Responses on conventional preparations had a significantly higher false-positive rate (14.2%) than ThinPrep smears (2.4% false-positive rate, $P = .002$).

The proportion of exact matches by participants on conventional and ThinPrep slides is summarized in Table 4. Using the exact match criteria for educational slides, conventional preparations had significantly lower error rates for both individual and laboratory responses. For individuals, the exact match error rate was 41.4% for conventional preparations and 57.5% for ThinPrep preparations ($P < .001$). For laboratories, the exact match error rate was 40.5% for educational conventional preparations and 58.9% for educational ThinPrep preparations ($P < .001$).

Table 5 shows response rates for educational slides with a reference diagnosis of reparative changes by diagnostic category for 100-level responses for individual and laboratory responses. As noted, a higher proportion of ThinPrep responses are correctly matched to series, but a lower proportion are exact matches to the reference di-

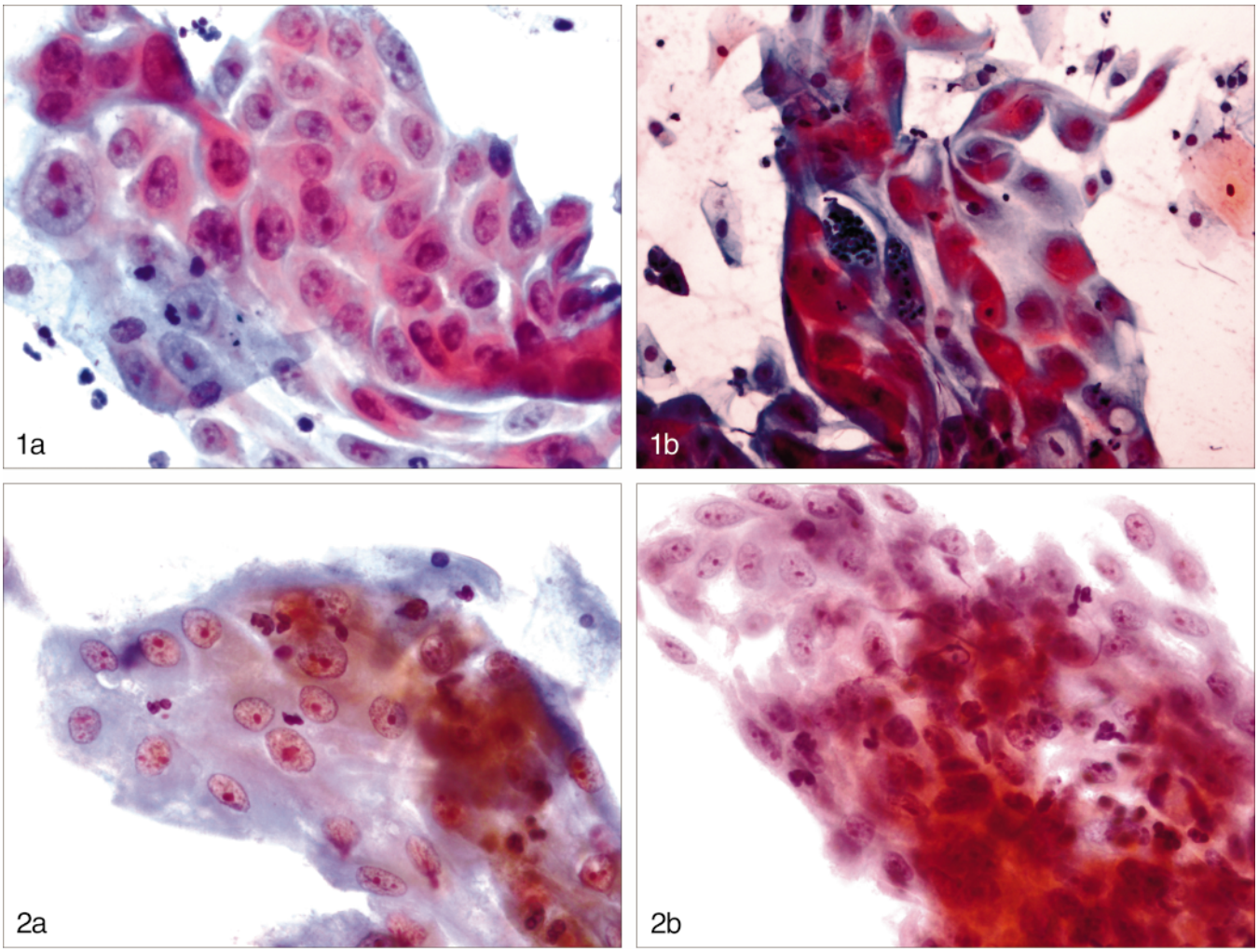

Figure 1. *a and b, Reparative changes in a conventional smear (Papanicolaou, original magnification ×400). Both figures show flat sheets of squamous cells with cellular streaming, intercellular windows, and prominent nucleoli. b, Intracytoplasmic neutrophils are commonly seen.*

Figure 2. *a and b, Reparative changes in a ThinPrep preparation (Papanicolaou, original magnification ×400). Although these squamous cells show features similar to those seen in conventional smears, these typical cytomorphologic features were less prevalent in this preparation.*

agnosis of reparative changes. The bulk of these discordant matches are negative for intraepithelial lesions (101) (NILM-NOS). More than 40% of the responses for ThinPrep slides are NILM-NOS (101) for both individuals and laboratories, while about 20% of the responses for conventional slides are in this category ($P < .001$ for both individuals and laboratories).

A cytomorphologic review of individual ThinPrep specimens with a reference diagnosis of reparative changes revealed both similarities and dissimilarities with respect to the appearances of reparative changes observed in conventional slides. Both types of preparations showed flat sheets of cells with a uniform arrangement. The nuclei tended to line up and showed a streaming nuclear polarity with prominent nucleoli (Figure 1, a and b). In ThinPrep smears, however, the cells were more rounded, and hence, the streaming was less apparent (Figure 2, a and b).

COMMENT

On review of the data in the CAP-PAP, it was suspected that cases with a diagnosis of reparative change in ThinPrep specimens were performing differently from conventional smears with the same interpretation, and our analysis from the data for a 4-year period confirms this finding. ThinPrep slides with a reference diagnosis of reparative changes have a lower false-positive discordancy rate than conventional slides with the same reference diagnosis (Table 3), but they are less likely to be labeled exactly as reparative changes (Table 4), which is reflected in a higher exact match error rate for these ThinPrep slides. To our knowledge, these results are the first to suggest that the interpretation of reparative change in ThinPrep specimens is different from conventional smears. An earlier study from the CAP-PAP described some differences in response rates between ThinPrep and conventional slides with respect to invasive squamous carcinoma.[12]

This finding has been identified using only educational conventional and ThinPrep slides. Validating ThinPrep slides showing reparative changes may be a particular challenge. Although the false-positive rate of educational ThinPrep slides is low (Table 3), these slides have a high

exact match error rate, which may preclude them from obtaining validation status (Table 4).

The finding of a lower false-positive discordancy rate in cases of ThinPrep reparative changes might suggest that this diagnosis is more easily made in this type of preparation than in the conventional smear. However, the finding of a higher exact match error rate in cases of ThinPrep slides with reparative changes (Table 4) and the high prevalence of NILM responses (Table 5) is not consistent with this explanation, since this finding suggests that participants are failing to recognize reparative changes in many cases. The combination of these results suggests that the ThinPrep reparative changes in the CAP-PAP represent only a subset of all cases that are diagnosed as reparative change in conventional smears and are composed of the least atypical and alarming of all possible cases. Hence, the false-positive discordancy rate (Table 2) is lower and, in many cases, unrecognized, leading to high exact match error rates (Table 4) and responses of NILM-NOS (Table 5). Both original and duplicate ThinPrep slides were used in this study, but it is unlikely that the use of duplicate slides would fully explain this finding. Although duplicate ThinPrep slides may contain fewer (or more) diagnostic reparative sheets, any duplicate slide was reviewed both by the original, submitting laboratory and the Committee to confirm the finding of reparative changes prior to slide circulation.

Slides with a reference diagnosis of reparative changes constitute a smaller proportion of ThinPrep slides than conventional slides in the CAP-PAP (1.2% vs 3.7%, Table 2). While the CAP-PAP program data are not adequate to address this question formally, it is likely that this is a reflection of the rarity of these cases being submitted to the CAP-PAP by donor laboratories. It is possible that cases with more florid features of repair are no longer categorized as reparative changes in ThinPrep specimens and are routinely categorized as something else, most probably atypical glandular cells. The reduction of blood and inflammation within a ThinPrep preparation may allow the recognition of cells with a higher diagnostic significance that would otherwise pass unrecognized in a conventional smear with these obscuring factors. This interpretative tendency would also be consistent with the preceding explanation. Moreover, on review of the cases of repair in ThinPrep specimens in the CAP-PAP, it appears that the amount of atypia is relatively mild.

In summary, these 3 findings of reparative changes in ThinPrep specimens (lower false-positive discordancy rate, lower exact match diagnosis rate, and frequent responses of NILM), coupled with the possible reduction in the number of ThinPrep slides showing reparative changes, suggest that cases that previously would have been recognized as florid reparative changes are not being labeled as such in ThinPrep preparations and are being classified using another cytodiagnostic category.

As previously noted, the diagnosis of repair in conventional smears has traditionally been associated with an elevated risk of squamous intraepithelial lesion. Thus, the fact that the recognition of reparative change in ThinPrep specimens is different from that of conventional smears is important. Since reparative changes in ThinPrep preparations may represent only a subpopulation of the entire group, it is possible that reparative change in these specimens is associated with a significantly lower risk of squamous intraepithelial lesions than the same interpretation in conventional smears.

In conclusion, a diagnosis of reparative changes on ThinPrep specimens in the CAP-PAP is significantly more specific than is such a diagnosis on conventional smears. Reparative changes on a ThinPrep slide are less often exactly labeled or recognized when compared to conventional smears. Cases with a reference interpretation of reparative changes make up a smaller percentage of ThinPrep cases than conventional smears. In combination, these findings suggest that reparative changes on ThinPrep slides in the CAP-PAP are only a proportion of all possible reparative changes, consisting of the less atypical and noticeable cases. A diagnosis of reparative changes in ThinPrep may not share the same clinical significance as that of one on a conventional smear.

We thank Ms Jennifer Haja, BA, and Ms Barbara Blond, MBA.

References

1. Ueki M, Veda M, Kurokawa A, et al. Cytologic study of the tissue repair cells of the uterine cervix with special reference to their origin. *Acta Cytol.* 1992; 36:310–318.

2. Bibbo M, Keebler CM, Weid GL. The cytologic diagnosis of tissue repair in the female genital tract. *Acta Cytol.* 1971;15:133–140.

3. Patten SF. Diagnostic cytopathology of the uterine cervix. In: Weid GL, ed. *Monographs on Clinical Cytology.* New York, NY: S Karger; 1978:93–114.

4. Geirsson G, Woodworth FE, Patten SF, Bonfiglio TA. Epithelial repair and regeneration in the uterine cervix: an analysis of cells. *Acta Cytol.* 1977;21:371–378.

5. Flenker H. Tissue repair in the cervix. In: Wied GL, Keebler CM, Koss LG, Patten SF, Rosenthal DL, eds. *Compendium on Diagnostic Cytology.* 7th ed. Chicago, Ill: Tutorials of Cytology; 1992:75–76.

6. Rimm DL, Gmitro S, Frable WJ. Atypical reparative change on cervical/vaginal smears may be associated with dysplasia. *Diagn Cytopathol.* 1996;14: 374–379.

7. Ng WK, Li AS, Cheung LK. Significance of atypical repair in liquid based gynecologic cytology: a follow-up study with molecular analysis for human papillomavirus. *Cancer Cytopathol.* 2003;99:141–148.

8. Chang JH, Rutkowski MA, Phillips MA, Wilbur DC, Mesonero CE. The potential biological and clinical significance of reactive/reparative cellular changes in cervical smears: a five year follow-up study [abstract]. *Acta Cytol.* 1996;40: 1038.

9. Soofer SB, Sidawy MK. Reactive cellular change: is there an increased risk for squamous intraepithelial lesions? *Cancer Cytopathol.* 1997;81:144–147.

10. Colgan TJ, Woodhouse SL, Styer PE, Kennedy M, Davey DD. Reparative changes and the false-positive/false-negative Papanicolaou test: a study from the College of American Pathologists Interlaboratory Comparison Program in Cervicovaginal Cytology. *Arch Pathol Lab Med.* 2001;125:134–140.

11. Renshaw AA, Davey DD, Birdsong GG, et al. Precision in gynecologic cytologic interpretation: a study from the College of American Pathologists Interlaboratory Comparison Program in Cervicovaginal Cytology. *Arch Pathol Lab Med.* 2003;127:1413–1420.

12. Renshaw AA, Young NA, Birdsong GG, et al. Comparison of performance of conventional and ThinPrep gynecologic preparations in the College of American Pathologists Gynecologic Cytology Program. *Arch Pathol Lab Med.* 2004; 128:17–22.