Unsatisfactory Reporting Rates

2006 Practices of Participants in the College of American Pathologists Interlaboratory Comparison Program in Gynecologic Cytology

Ann T. Moriarty, MD; Amy C. Clayton, MD; Sue Zaleski, SCT(ASCP); Michael R. Henry, MD; Mary R. Schwartz, MD; Galen M. Eversole, MD; William D. Tench, MD; Lisa A. Fatheree, SCT(ASCP); Rhona J. Souers, MS; David C. Wilbur, MD

• ***Context.*—Minimum cellular criteria for satisfactory Papanicolaou tests were established with the Bethesda System in 2001, and unsatisfactory rates are used as a quality-reporting measure.**

***Objective.*—To evaluate practices and unsatisfactory rates from laboratories responding to the 2007 College of American Pathologists supplemental questionnaire survey.**

***Design.*—In 2007, a supplemental questionnaire was mailed to 1621 laboratories enrolled in the 2006 College of American Pathologists Interlaboratory Comparison Program in Gynecologic Cytology (PAP Education), requesting data from the 2006 calendar year. Unsatisfactory rates, reasons for unsatisfactory specimens, laboratory size, and specimen preparation type were analyzed.**

***Results.*—A total of 42% of the laboratories responded to the survey. Most of those laboratories (637 of 674; 94.5%) used the Bethesda System minimum cellularity criteria. Of those laboratories responding, 79% (527 of 667) used the Bethesda System criteria for atrophic or postirradiation specimens. Unsatisfactory rates have increased since 1996. SurePath preparations were associated with the lowest unsatisfactory rate (50th percentile, 0.30; 95th percentile, 1.3), conventional Papanicolaou tests had the highest 95th percentile rates (50th percentile, 1.0; 95th percentile, 5.90), and ThinPrep specimens had the highest median percentile (50th percentile, 1.1; 95th percentile, 3.4). The most-common reason for unsatisfactory Papanicolaou tests was too few squamous cells. Air-drying artifact was the least-common reason for unsatisfactory reporting for liquid-based preparations.**

***Conclusions.*—Use of the Bethesda System criteria for unsatisfactory specimens is widespread. Unsatisfactory rates have increased since 1996; however, the median rates are 1.1% or less for all preparations. Results from the College of American Pathologists PAP Education supplemental questionnaire continue to provide valuable benchmarking data for cytologic quality-improvement programs in laboratories.**

(*Arch Pathol Lab Med.* 2009;133:1912–1916)

The College of American Pathologists (CAP) assesses practice patterns using survey questionnaires sent to laboratories participating in CAP programs. Cervicovaginal cytology surveys have been used since 1994 and have evaluated practices related to implementation of the Bethesda System (TBS) terminology and reporting rates for interpretive categories used in Papanicolaou (Pap) testing.

Benchmarks for unsatisfactory rates were last reported in 2004,[1] soon after implementation of the third version of TBS, which defined cellular criteria for adequacy for both liquid-based preparations and conventional smears.[2] Since the publication of the 2003 practice patterns, the use of liquid-based Pap specimens has expanded, whereas use of conventional preparations has decreased.[3] Although there was no increase in median unsatisfactory rates and no difference in unsatisfactory rates reported between liquid-based or conventional preparation types in the prior 2003 CAP Interlaboratory Comparison Program in Gynecologic Cytology (PAP Education) supplemental questionnaire (SQ), some authors reported changes in adequacy rates after implementing TBS criteria for adequacy.[4] Liquid-based preparations are marketed as decreasing unsatisfactory rates by eliminating or diminishing the effects of obscuring inflammation, blood, and air-drying. Therefore, unsatisfactory rates could be expected to change with increased implementation of liquid-based preparations. Because of the expansion of liquid-based preparations in the United States since the 2003 PAP SQ and because the survey includes results from a larger segment of laboratories using liquid-based preparations, the 2007 SQ may more accurately represent current cytology practices.

Accepted for publication March 31, 2009.

From AmeriPath Indiana, Indianapolis (Dr Moriarty); the Department of Anatomic Pathology, Mayo Clinic, Rochester, Minnesota (Drs Clayton and Henry); the Department of Pathology, University of Iowa Medical Center, Iowa City (Ms Zaleski); the Department of Pathology, The Methodist Hospital, Houston, Texas (Dr Schwartz); Quest Diagnostics Inc, Las Vegas, Nevada (Dr Eversole); Palomar Medical Center, Escondido, California (Dr Tench); the Departments of Cytology Surveys (Ms Fatheree) and Biostatistics (Ms Souers), College of American Pathologists, Northfield, Illinois; and the Department of Pathology, Massachusetts General Hospital, Boston (Dr Wilbur).

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

Reprints: Ann T. Moriarty, MD, AmeriPath Indiana, 2560 N Shadeland Ave, Ste A, Indianapolis, IN 46219-1739 (e-mail: amoriarty@ameripath.com)

MATERIALS AND METHODS

The 2007 CAP SQ was mailed to 1621 laboratories enrolled in the 2006 CAP PAP Education program, requesting data from the 2006 calendar year. Not every laboratory responded to every question. The number of laboratories responding to each question is indicated in the tables. Some laboratories only provided data for all preparation types combined.

Laboratories were asked whether they used TBS 2001 reporting terminology and whether they used the minimum squamous cellularity criteria for cervical cytology, that is, whether they estimated a minimum of 8000 to 12 000 squamous cells on conventional smears or estimated a minimum of 5000 squamous cells on liquid-based cytology. Laboratories were also asked whether they applied the minimum squamous cell adequacy criteria for atrophic and/or postirradiation specimens. They were asked to rank the reasons for unsatisfactory results for each of 3 preparation types (conventional, ThinPrep, and SurePath) as (1) too few squamous cells, (2) obscuring inflammation, (3) obscuring blood, (4) obscuring foreign material/lubricant, (5) air-drying, or (6) other specified reason. Ranking ranged from 1 (most frequent) to 6 (least frequent). Laboratories could also select a *not available* or *not applicable* response for all ranks. Each reason was given a mean rank score, based on the reported frequency and the number of participants reporting that reason as a cause for unsatisfactory results. Laboratories were also asked to estimate the annual volume and unsatisfactory rates of each preparation type for 2006.

Reporting rate distribution tables were constructed for unsatisfactory rates for each preparation type. Unsatisfactory distribution rates were not Gaussian in distribution; therefore, the results are presented in percentile reporting rates instead of means and standard deviations. The 50th percentile reporting rate represents the median. Reporting rates that were unclear or appeared to represent impossible values were deleted. Some laboratories only provided data for all preparation methods combined.

Differences between reporting rates for each specimen type were evaluated using the Wilcoxon signed rank test, a nonparametric analysis of paired data. A Pearson χ^2 test was used to determine whether there was a significant association between the adequacy criteria for minimum squamous cellularity in TBS 2001 and in the dominant preparation type, conventional or liquid based.[5] When applicable, results of the survey were compared with results from the previous 1996 and 2003 CAP PAP Education questionnaires.[1]

RESULTS

Of the 1621 questionnaires mailed, 679 laboratories (42%) returned the 2007 questionnaire reporting data from 2006. This response rate is similar to the 43% (759 of 1751) response rate for the 2003 SQ. Not all laboratories answered every question. Most laboratories (662 of 677; 97.8%) used TBS for reporting Pap tests; this is a 14.4% increase in TBS use compared with what was reported in the 2003 survey (648 of 758; 85.5%). Although the participants reported using TBS, only 94.5% (637 of 674) of the responding laboratories reported using criteria for minimum squamous cellularity (8000 to 12 000 estimated squamous cells on conventional smears or 5000 estimated squamous cells for liquid-based cytology.)

Most (397 of 656; 60.4%) of the laboratories report using a combination of conventional and liquid-based preparations. The percentage of laboratories that used only conventional smears dropped 10.7 percentage points (a 43.9% decrease) from 24.4% (181 of 742) in the 2003 survey to 13.7% (90 of 656) in the present survey. Concomitantly, the percentage of laboratories offering only liquid-based preparations increased 173% from 9.3% (69 of 742) in 2003 to 25.4% (167 of 656) in the present survey.

Table 1. Percentile Reporting Rates for Unsatisfactory Papanicolaou Tests by Method

Method	Response, No.[a]	Percentile						
		5th	10th	25th	50th	75th	90th	95th
Conventional	109	0	0.1	0.5	1.0	2.4	4.7	5.9
ThinPrep	197	0.1	0.3	0.6	1.1	1.7	2.9	3.4
SurePath	67	0.1	0.1	0.2	0.3	0.6	1.1	1.3
All[b]	354	0.1	0.2	0.4	0.9	1.3	2.1	2.9

[a] Number of laboratories responding.
[b] This refers to compiled data reported by laboratories and represents overall rated for all methods. Not all laboratories reported combined method rate.

Of the 520 respondents that reported using TBS criteria for adequacy, more laboratories (437 of 520; 84%) used predominantly liquid-based preparations versus those (83 of 520; 16%) that used predominantly conventional smears ($P < .001$). Almost all (436 of 450; 97%) of the participants that reported using a predominance of liquid-based preparations used TBS criteria for adequate squamous cellularity; whereas only 86% (84 of 98) of the laboratories that reported using a predominance of conventional preparations used TBS criteria for minimum cellularity ($P < .001$). It was not possible to identify whether laboratories applied criteria differently for conventional or liquid-based preparations within laboratories. Of the 32 laboratories that did not use the Bethesda 2001 criteria for adequacy, 12 (38%) were international: 4 were from Canada, 4 from Japan, and the remaining laboratories (n = 24) were from the Arab Emirates, India, Israel, and Peru. Of the 667 laboratories that responded to this question, 527 (527 of 667; 79%) also extended use of minimum squamous cell adequacy criteria to atrophic and/or postirradiation specimens.

Percentile reporting rates for individual Pap methods as well as a combined rate for all methods are shown in Table 1. SurePath preparations were associated with the lowest median and 95th percentile unsatisfactory rates (0.3% and 1.3% respectively). ThinPrep Pap tests were reported with the highest median percentile ranking (1.1%), and conventional preparations had the highest 95th percentile ranking (5.9%). Using the Wilcoxon rank sum test of the predominant preparation type at each laboratory, SurePath preparations had a significantly lower unsatisfactory rate ($P < .001$).

Table 2 compares the median and 90th percentile unsatisfactory rates reported for 1996, 2002, and 2006. Median and 90th percentile unsatisfactory rates have increased since 2002. Although it appears there were decreasing median unsatisfactory rates with increasing laboratory volume (Table 3), the overall distribution is not statistically different between volume groups ($P = .69$). Likewise, in 2002, there was no significant laboratory volume effect on the unsatisfactory rates for the overall group or for any specific slide types.

The mean rank score of each reason for an unsatisfactory specimen can be seen in Table 4. A low mean rank score indicates that the indicator is more significant as the cause of unsatisfactory results; the lower the mean rank score, the more often this is reported as a reason for an unsatisfactory specimen. All methods rank ''too few squamous cells'' as the leading cause of unsatisfactory Pap tests. Obscuring inflammation was the second most common reason for an unsatisfactory conventional smear ''ob-

Table 2. Comparison of Unsatisfactory Reporting Rate for All Slides, Conventional, ThinPrep, and SurePath

	1996[a]		2003			2006		
Method	Median	90th Percentile	Median (95% CI)	90th Percentile	Response, No.[b]	Median (95% CI)	90th Percentile	Response, No.[b]
All	0.5	2.0	0.5 (0.4, 0.6)	1.4	482	0.9 (0.8, 0.9)	2.1	354
Conventional	0.5	2.0	0.5 (0.4, 0.6)	2.0	221	1.0 (0.7, 1.4)	4.7	109
ThinPrep			0.4 (0.4, 0.5)	1.2	182	1.1 (1.0, 1.1)	2.9	197
SurePath			0.2 (0.2, 0.5)	0.9	41	0.3 (0.3, 0.4)	1.1	67

Abbreviation: CI, confidence interval.

[a] Rates represent conventional Papanicolaou smears only.
[b] Number of laboratories responding with data for unsatisfactory rates.

Table 3. Annual Laboratory Volumes Versus Median Unsatisfactory Percentiles

	All		Conventional		ThinPrep		SurePath	
Laboratory Volumes	50th Percentile	Response, No.[a]	50th Percentile	Response, No.[a]	50th Percentile	Response, No.[a]	50th Percentile	Response, No.[a]
<5000	1.0	52	1.0	30	0.9	47	0.3	13
5000–9999	1.0	50	0.4	15	0.7	29	0.2	12
10 000–19 999	0.9	60	1.1	23	1.0	37	0.4	16
20 000–49 999	0.9	50	1.0	23	1.2	29	0.3	12
50 000–99 999	0.8	14	2.0	9	1.5	7	0.3	3
100 000–199 999	0.6	7	0.8	4	0.7	5	0.3	1
>200 000	0.7	11	0.7	5	0.8	4	0.2	4
Total		**244**		**105**		**158**		**61**

[a] Number of laboratories responding.

Table 4. Comparison of Mean Rank Score of Reasons for Unsatisfactory Papanicolaou Tests by Method

	Conventional		ThinPrep		SurePath	
Reason	Cited, No.[a]	Score[b]	Cited, No.[a]	Score[b]	Cited, No.[a]	Score[b]
Squamous cellularity	430	1.2	433	1.2	163	1.2
Obscuring inflammation	412	2.0	341	2.8	72	2.7
Obscuring blood	375	2.8	324	2.7	59	3.2
Foreign material/lubricant	251	4.5	250	3.3	32	4.1
Air-drying	335	3.3	110	4.8	21	4.9
Other	71	3.8	41	3.5	21	2.1

[a] Total number of times the reason is cited for an unsatisfactory preparation.
[b] The lower the rank score, the more significant the weighted reason for an unsatisfactory specimen.

scuring blood'' for a ThinPrep, and ''other'' for SurePath preparations. Foreign material or lubricant was ranked higher as a reason for unsatisfactory specimens for ThinPrep as compared with either conventional or SurePath preparations. Air-drying was the least common reason for an unsatisfactory Pap test for liquid-based preparations. For those 98 laboratories that provided a response in the ''other'' category, 15 of 49 laboratories (30.6%) using conventional smears, 3 of 20 laboratories (15%) using SurePath, and 3 of 29 laboratories (10%) using ThinPrep reported ''too thick'' as a reason for unsatisfactory specimens. In addition, 6 of 49 laboratories (12%) indicating ''other'' in conventional preparations reported receiving broken or unlabeled slides as a reason for unsatisfactory specimens.

COMMENT

The CAP SQ has been a valuable tool for investigating the self-reported practices of laboratories in the United States. The results of these questionnaires have led to benchmarking data for quality assurance purposes and are used in the CAP Laboratory Accreditation Program checklist.[6] Despite the previous lack of standardized criteria for assessing adequacy, unsatisfactory rates have been published since 1992.[7] In 1992, TBS was novel, and conventional smears were the predominant Pap test method. The median unsatisfactory rate at that time was 0.5%. One year later, a 1993 Q-Probe from the CAP cited an unsatisfactory rate of 0.28%.[8] The 2003 SQ was circulated shortly after the newly defined TBS semiquantitative criteria for adequate cellularity were published in 2002, and the use of the new criteria was likely not yet widespread. The 2003 survey did not find a difference in unsatisfactory rates compared with 1992, despite expectations that the new, well-defined TBS criteria might result in an increased rate of unsatisfactory specimens. At the same time, liquid-based cytology was beginning to flourish in the United States, which theoretically would reduce the number of unsatisfactory specimens because these preparatory methods optimized cellularity and reduced potentially obscuring inflammation and blood. Unsatisfactory rates were most likely affected by the countering effects of new criteria for unsatisfactory specimens, which increase the number of unsatisfactory specimens, and the concurrent increase of liquid-based Pap tests, which would theoretically decrease the unsatisfactory rate. Therefore, it was not

surprising that the 2003 SQ did not demonstrate a statistically significant difference in unsatisfactory rates compared with the 1992 data. The 2007 SQ was circulated 5 years after publication of the criteria for cellularity, and by that time, liquid-based preparation was the predominant preparation type reported by participants responding to the survey.[9] The 2006 benchmarking data demonstrate an increase in the median and 90th percentile unsatisfactory rates. However, at the median level, the 95% confidence intervals for 2002 SurePath preparations overlap with those of 2006, indicating that the trend may be due to sampling variability. ThinPrep Pap tests have a higher median unsatisfactory rate than either SurePath or conventional preparations; SurePath specimens have the lowest median and 95th percentile unsatisfactory rate of the 2 liquid-based methodologies. Median rates for all preparation types were 1.1% or less.

Low squamous cellularity was the most common cause of unsatisfactory Pap tests regardless of the preparation used; this is probably because of the acceptance of the semiquantitative TBS criteria for cellularity embraced by 94.5% (637 of 674) of the responding laboratories. The 5.5% of laboratories that did not use the Bethesda adequacy criteria were more likely to report using predominantly conventional smears. Laboratories using TBS minimum squamous cellularity criteria were more likely to report using liquid-based preparations as a predominant method.

Air-drying was predictably of little consequence for liquid-based preparations. Obscuring blood and inflammation were the second and third leading causes, respectively, of unsatisfactory specimens for conventional and ThinPrep specimens but were ranked third and fourth for SurePath. This most likely reflects the differential gradient sedimentation technique of the SurePath methodology that selectively eliminates approximately 50% of the inflammation and greater than 90% of the blood in the material used to create the slide. All methods identified preparations that were too thick as an ''other'' reason for unsatisfactory specimens. Additionally, unlabeled or broken conventional smears were cited as a cause for unsatisfactory specimens. Broken slides are a technical reason for unsatisfactory specimens and are usually rejected before accessioning or reporting.

There is no significant laboratory volume effect on the unsatisfactory rates for the overall group or any specific slide types, similar to the 2002 data. This may be skewed because most laboratories did not report high volumes of Pap tests. Most of the laboratories responding to this survey were laboratories that reported less than 50 000 gynecologic specimens per year, which is the typical demographic profile of laboratories participating in CAP PAP Education in 2006. Of the laboratories participating in CAP PAP Education, 83.8% reported annual volumes of less than 50 000 gynecologic specimens. Laboratories participating in the CAP PAP Education for 2006 were from a variety of facilities, including 58 international laboratories; most respondents were hospital laboratories with profiles similar to those responding to the 2007 SQ.[9]

Most respondents report using adequacy criteria for postirradiation and atrophic preparations. The Bethesda System explicitly states that the minimum squamous criteria are for screening cervical cytology preparations only. In patients with special circumstances (hysterectomy, atrophy, or irradiation), it may be necessary to modify the criteria for minimum squamous cell adequacy; surprisingly, 79% (527 of 667) of laboratories report using the minimum cellularity criteria for patient specimens from this population with special circumstances. However, it cannot be ascertained from the responses whether the laboratory modified in any way the minimal squamous cellularity criteria for women falling into these clinical categories.

Although the survey response rate was high and represents the population of participants in the 2006 CAP PAP Education, possible weaknesses of the survey include that not all questions were answered by all respondents and that results were skewed toward laboratories with smaller volumes. The emphasis on smaller laboratories may not reflect the unsatisfactory rates seen in the population at large. Of the laboratories reporting separate preparation methods, there were fewer laboratories reporting SurePath data (82 laboratories; 17%) than conventional (167 laboratories; 36%) or ThinPrep data (221 laboratories; 47%). Another weakness of these data is that they are self-reported. Self-reported data may be suspect if laboratories misunderstand the criteria and report that they use TBS when their local practice does not reflect the actual recommendations. There may be inconsistency in reporting causes of unsatisfactory specimens as well. For example, blood may be trapped in filtered preparations, such as ThinPrep, precluding squamous cell representation on the slide. In this case, laboratories may cite either scant squamous cellularity or obscuring blood as a cause of the unsatisfactory preparation, depending upon laboratory policy. Likewise, lubricant may interfere with squamous cell representation on the preparation, and laboratories may choose to cite scant squamous cellularity as a cause of the unsatisfactory preparation instead of interfering foreign material. Rank distributions represent sources of confusion as well. For example, air-drying was cited as the reason for unsatisfactory liquid-based preparations by 7% (131 of 1867) reasons cited for unsatisfactory liquid-based preparations. However, the rank number was high. The conclusion that air-drying is not a significant cause of unsatisfactory specimens in liquid-based preparations is based on rank *score.* A generated score does not imply that the method was a significant source of unsatisfactory specimens for any laboratory using liquid-based preparations.

In summary, TBS criteria for minimal cellular criteria are widely used in most laboratories responding to the 2006 SQ in gynecologic cytology. The 2007 SQ provides valuable benchmarking data for laboratory use in quality-assurance practices. The latest survey indicates that liquid-based preparations make up an increasing proportion of Pap test volumes. Unsatisfactory rates have also increased since the implementation of TBS, but median rates are still 1.1% or less. Low squamous cellularity remains the leading cause of an unsatisfactory Pap test, with blood and inflammation ranking relatively high as confounding, obscuring elements. Air-drying is predictably the lowest-ranked cause for unsatisfactory specimens that are processed using liquid-based methods.

References

1. Davey DD, Neal MH, Wilbur DC, Colgan TJ, Styer PE, Mody DR. Bethesda 2001 implementation and reporting rates: 2003 practices of participants in the College of American Pathologists Interlaboratory Comparison Program in Cervicovaginal Cytology. *Arch Pathol Lab Med.* 2004;128(11):1224–1229.

2. Solomon D, Davey D, Kurman R, et al. The 2001 Bethesda System: terminology for reporting cervical cytology. *JAMA.* 2002;287(16):2114–2119.

3. Eversole G, Moriarty AT, Schwartz MR, et al. Practices of participants in the

College of American Pathologists Interlaboratory Comparison Program in Cervicovaginal Cytology: 2006. *Arch Pathol Lab Med.* In press.
4. Fidda N, Miron J, Rodgers W, Rader A. Impact of the new Bethesda System 2001 on specimen adequacy of conventional cervicovaginal smears. *Diagn Cytopathol.* 2004;30(4):235–239.
5. Snedecor GW, Cochran WG. *Statistical Methods.* 8th ed. Ames, IA: The Iowa State University Press; 1989.
6. Commission on Laboratory Accreditation. *Laboratory Accreditation Program: Cytopathology Checklist.* Northfield, IL: College of American Pathologists; 2007.
7. Davey DD, Nielsen ML, Rosenstock W, Kline TS. Terminology and specimen adequacy in cervicovaginal cytology: the College of American Pathologists Interlaboratory Comparison Program experience. *Arch Pathol Lab Med.* 1992;116(9):903–907.
8. Jones BA. Rescreening in gynecologic cytology: rescreening of 3762 previous cases for current high-grade squamous intraepithelial lesions and carcinoma—a College of American Pathologists Q-Probes study of 312 institutions. *Arch Pathol Lab Med.* 1995;119(12):1097–1103.
9. 2006 Year End Summary Report: PAP Gynecologic Cytology PT Program (PAP PT). Northfield, IL: College of American Pathologists; 2007.