Papanicolaou Smear Quality Assurance: Providing Feedback to Physicians

Peter Curtis, MD; Bron Skinner, PhD; John J. Varenholt, MD; Lois Addison, MA, MLT (ASCP); Jacqueline Resnick, MA; and Mizanu Kebede, MS, MT (ASCP)

Chapel Hill, North Carolina, Pensacola, Florida, and Dunrobin, Ontario

Background. The effective management of Papanicolaou (Pap) smears depends on the reliability and accuracy of obtaining and interpreting the specimen. Provider sampling error is one of the important factors contributing to inadequate specimens. Feedback on provider performance may be an effective way to improve the quality of Pap smears.

Methods. A pilot study in a university-based residency program involving resident and faculty physicians was initiated to assess the impact of feedback on performance of Pap smears. After establishing adequacy and inadequacy criteria and recording adequacy rates for 3 months, individual and group feedback was implemented. No formal educational intervention on Pap smear technique was undertaken.

Results. The quality of 836 Pap smears performed by 9 faculty and 13 resident physicians showed continued improvement in both sampling and slide preparation to 90% adequacy over a 9-month period. This improvement, though clinically useful, was not statistically significant owing to the relatively small numbers of smears performed by each physician. This form of feedback may be useful in both practice and educational settings.

Conclusions. Feedback without any formal educational intervention led to a clinically useful trend of improvement in the quality of Pap smears, which has been sustained since the study began. This type of simple feedback may be useful in practice settings and particularly valuable in pinpointing areas for improvement for learners in residency programs.

Key words. Papanicolaou smear; quality assurance, health care. *J Fam Pract 1993; 36:309-312.*

The effectiveness of cytologic screening for cervical cancer depends on reaching at-risk women for testing as well as the reliability and accuracy of obtaining, handling, and interpreting the smear. Potential errors leading to false-negative tests may occur when obtaining the specimen, in the laboratory, and in reporting results to the health care provider.[1,2]

The presence of endocervical cells indicates that the transformation zone has been sampled. Studies have suggested that both an endocervical and ectocervical specimen must be collected to ensure an adequate smear.[3,4] Elias et al noted that the rate of detection of dysplasia increased by at least 60% in smears containing endocervical cells compared to those without endocervical cells.[3,4] Some research indicates, however, that endocervical harvesting contributes little to improving false-negative testing rates.[5,6] Physician sampling error is probably one of the most important factors contributing to the inadequate cervical smear,[7] but laboratories do not uniformly give individual profiles of collection adequacy rates unless specifically requested by the provider (personal communication, P. Ashton, Roche Biomedical Laboratories, March 1988).

There is evidence that regular performance feedback to physicians alters behavior and is most successful when the feedback is individualized, includes comparisons with peers, is in the ambulatory setting, and is delivered by a respected source of information.[8,9] A number of studies have shown that acceptance or adoption of improved techniques occurs when individuals perceive a discrepancy between their current level of performance and the expected standard.[10,11]

A pilot study undertaken in the family practice center in 1985 at the University of North Carolina showed a cervical smear adequacy rate of 74% (presence of endocervical cells) using the swab and spatula method, as well as numerous problems with tracking and follow-up

Submitted, revised, December 1, 1992.

From the Departments of Family Medicine, School of Medicine (P.C., B.S., M.K.), and Maternal and Child Health, School of Public Health (J.R.), University of North Carolina, Chapel Hill; the Medical Center Clinic, Pensacola, Florida (J.J.V.), and LAA Associates, Dunrobin, Ontario, Canada (L.A.). Requests for reprints should be addressed to Peter Curtis, MD, Department of Family Medicine, Campus Box 7595, University of North Carolina, Chapel Hill, NC 27599.

ISSN 0094-3509

Table 1. Changes in Pap Smear Adequacy Rates of Physicians and Second- and Third-Year Residents Over 12 Months of Quality Assessment and Feedback

Physician Group	Baseline Adequacy Rates, % January–March	Pap Smear Adequacy Reports, % April–June*	July–September†	October–December
Faculty	88.3	94.1	90.8	92
3rd-year residents	77.4	84.5	Graduated	Graduated
2nd-year residents	64.3	92.8	94.2	91

**Feedback to physicians began in April.*
†Second-year residents began their 3rd year of training in July.

of abnormal test results. Consequently, an automated quality assurance, prompting, and feedback system for cervical smears was developed. The purpose of this was to develop routine systems that would help physicians in the management of normal and abnormal results and to improve the performance of cervical smears, particularly in a training environment. In this paper we report on the effects of setting criteria for determining the technical adequacy of Papanicolaou (Pap) smears and providing feedback to physicians.

Methods

The development of the system took place in the family practice center at the University of North Carolina, a primary care facility that was serving 10,000 patients (approximately 5000 of whom were women) with 22,000 visits annually. At the time of the study, 11 faculty physicians and 19 resident physicians worked in the center. On average, 116 Pap smears were performed each month, mainly by faculty and third-year residents. All cervical smears were taken using the Cytobrush and Ayre spatula method except in pregnant women, on whom a swab and spatula was used.[12] Slide preparation was performed by nurses assisting the physician, although some women physicians prepared their own slides when a nurse chaperone was not present. Two slides were used for each Pap smear; one for the Cytobrush sample (rolled over the slide) and one for the spatula sample. The specimens were spray-fixed immediately after the sample was placed on the slide, usually by the nurse. Instruction in technique for residents occurred through one-on-one precepting and for nurses by demonstration from the nursing or laboratory coordinator.

In 1988, the office laboratory at the family practice center developed and installed an automated Pap smear reporting and recall system. Algorithms were produced for decision management of normal, abnormal, inflammatory, adequate, and inadequate Pap smears. The algorithms were developed after discussion with consultant gynecologists and review of current literature. Reports of normal and abnormal results of individual patients were automatically generated for each physician. These could be displayed to include previous Pap smear results.

In consultation with the university hospital cytology laboratory, the adequacy of cervical smears was defined by (1) *specimen quality*: presence of endocervical cells, and (2) *process quality*: adequate slide preparation (not too thick, obscured, or dried out). Inadequacy for each cervical smear was defined by the following *specimen quality* factors: (1) absent endocervical cells, (2) too few cells for interpretation; and by *process quality* factors: (3) cervical smear totally dried before fixation, and (4) smear too thick for interpretation, or obscured. Using these criteria, data on adequacy of smears performed by individuals and groups (faculty and residents) of physicians were recorded for 3 months (Table 1). To assess the impact of feedback, physicians were given no orientation or warning that feedback would occur, and no educational intervention was attempted until 6 months later. At this stage, 3 months after data collection was started, feedback to the physicians began on a monthly basis, a comparison of their cervical smear adequacy rates (including cervical sampling and slide preparation errors) was made with those of colleagues.

Adequacy rates before and after feedback were compared using a paired *t* test.

Results

Over a 1-year period, faculty and third-year residents performed an average of 62 Pap smears each; second-year residents performed an average of 34 tests, and first-year residents an average of 18 tests. Of all tests performed, 11.2% were recorded as inadequate, of which 70% had a previous history of an abnormal test result, such as inflammatory changes or low-grade or high-grade squamous intraepithelial lesions.

For the purpose of evaluating Pap smear adequacy, data from 836 cervical smears performed by nine faculty

Table 2. Pap Smear Sampling and Preparation Errors Before and After Performance Feedback to Physicians

Group/Error Category	Baseline Period (3 mo), %	Feedback Phase 1st Period (3 mo), %	Feedback Phase 2nd Period (3 mo), %	Feedback Phase 3rd Period (3 mo), %
Faculty				
Sampling error	4.1	3.9	9.2	8.0
Preparation error	7.6	2.0	0.0	0.0
3rd-year residents				
Sampling error	4.8	10.0		
Preparation error	17.8	5.5	Graduated	Graduated
2nd-year residents*				
Sampling error	19.0	3.6	6.8	8.0
Preparation error	16.7	3.6	0.0	1.0

**Second-year residents began their 3rd year of training during the second 3-month period of the feedback phase.*

physicians (507 smears), seven third-year (230 smears), and six second-year residents (99 smears) are reported in this paper. Three faculty physicians who were absent for a significant part of the study and first-year residents who performed only a few cervical smears were excluded from the analysis.

The adequacy rates through 1990 of second- and third-year residents as well as faculty are shown in Table 1. These rates improved for all groups after feedback, but only the improvement for second-year residents was close to being statistically significant ($P = .06$). During the first 3-month observation period, sampling errors occurred in 4.8% and slide preparation errors in 17.8% of all smears by third-year residents; for second-year residents these errors were 19% and 16.7%, respectively. As shown in Table 2, second-year resident preparation errors decreased progressively by 13% after the first 3 months of feedback to zero when they entered the third year. Third-year residents and faculty all showed progressive declines in preparation error rates, but sampling errors were not consistently decreased. In the following year (1991) the overall adequacy rate for the faculty was 93%, for third-year residents it was 94%, and for second-year residents, 90%, which approximate the optimal levels reported in the literature.[13,14] The feedback system has continued as an integral part of clinical practice in the family practice center since its inception, and serves as both a quality assurance program and a way of monitoring resident performance.

Discussion

The quality assurance system for monitoring Pap smears was started because of the need to track atypical or abnormal test results in a complex clinical setting with many clinicians and to address a low cervical smear adequacy rate of 74%. However, by the time the quality assurance system was functioning, the sampling technique in the practice had changed from using a swab to using the Cytobrush, an improved method.[14] It was clear that sampling and slide preparation were distinctly different sources of error, particularly for Pap smears performed by residents who did not have a nurse present to assist them with the procedure. These residents fixed their own slides, which were frequently dried out because of delays in fixation. This type of error is important to recognize, since it can lead to false-positive interpretation by laboratory technologists.[15]

After the feedback began, laboratory staff reported significantly increased numbers of inquiries from clinicians regarding their personal feedback data and issues of technique. Informal conversations on the subject between clinicians were also noted, suggesting that feedback stimulated more interest in individual performance.

Although physicians received regular Pap smear results on individual patients, they were given no prior indication that feedback was to be initiated. No education regarding cervical smear sampling and fixation techniques was given until 6 months after feedback started, when aggregated data were reported at a departmental conference. It would appear that feedback alone stimulated clinicians and nurses to seek ways to improve their performance. No extra guidance or training in performing Pap smears was initiated. The effect of feedback on physician performance of Pap smears showed a trend of clinically useful improvement in adequacy rates for both trainees and faculty physicians. To show statistically significant improvement would require greater power and numbers of tests than were reported in this study. Similarly, given the already high adequacy rates achieved by the physicians, larger numbers of Pap smears would be

needed to show significant maintenance of this performance over a longer period.

For all groups of clinicians, improvement was most consistent in the preparation of slides. Variability in sampling performance could be explained either by uneven clinical skills or possibly by the proportion of each clinician's population of postmenopausal women with atrophic cervices. In the latter it can be more difficult to successfully sample the transformation zone. Second-year residents showed the most consistent improvement in both sampling and preparation performance, though they performed fewer Pap smears than more senior residents and faculty. These second-year residents had a smaller percentage of patients who were over 45 years of age than did the faculty; therefore, the initial sampling errors could probably be attributed to technique rather than to atrophic changes in the cervix. Another explanation for improvement in performance, particularly for second-year residents, could be a very steep learning curve that coincided with the feedback phase.

The potential effect of regular feedback on individual performance of tests in comparison with peers to produce significant and sustained improvement needs fuller study in primary care.[16] For instance, studies from Holland, where centralized laboratories are used in an organized health care system, indicate that performance feedback can be a powerful tool in changing physician behavior.[17]

Our clinical setting, a family practice training program, was different from the typical private practice, and the applicability and usefulness of this type of feedback would need to be assessed for different types of practice settings and laboratory services, whether serving the community or the hospital. Usually, cytology laboratories provide the clinician only with individual Pap smear reports. It would be useful, if clinicians can be tracked by the laboratory, to provide a mean adequacy rate for all clinicians, against which the individual's average rate could be compared.

Acknowledgment

Funded in part by a grant from the National Cancer Institute, grant No. JR18CA52748-01, National Institutes of Health.

References

1. Hakama M, Miller AB, Day NE, eds. International Agency for Research on Cancer Working Group on Cervical Cancer Screening—summary chapter in screening for cancer of the uterine cervix. Lyons, France: International Agency for Research on Cancer, 1986:133–41.
2. Bethesda System for reporting cervical/vaginal cytological diagnoses. A National Cancer Institute workshop. JAMA 1989; 262: 931–4.
3. Elias A, Linthorst G, Bekker B, Vooijs PG. The significance of endocervical cells in the diagnosis of cervical epithelial changes. Acta Cytol 1983; 27:225–9.
4. Vooijs PG, Elias A, Van der Graaf Y, Veling S. Relationship between the diagnosis of epithelial abnormalities and the composition of cervical smears. Acta Cytol 1985; 29:323–8.
5. Mitchell H, Medley G, Giles G. Cervical cancer diagnosed after negative results on cervical cytology. BMJ 1990; 300:1622–6.
6. Sidaway MK. Should we report cervical smears lacking endocervical components as unsatisfactory. Proceedings of the Bethesda System Second Conference. Bethesda, Md: National Cancer Institute, 1991.
7. Gay JD, Donaldson LD, Goellner JR. False negative results in cervical cytologic studies. Acta Cytol 1985; 29:1043–6.
8. Batstone GF. Educational aspects of medical audit. BMJ 1990; 301:326–8.
9. Fineberg HV, Funkhouser AR, Marks H. Variation in medical practice: a review of the literature. Presented at the Conference on Cost-Effective Medical Care: Implications of Variation in Medical Practice, sponsored by the Institute of Medicine, National Academy of Sciences, Washington, DC, February 1983.
10. Berwick DM, Coltin KL. Feedback reduces test use in a health maintenance organization. JAMA 1986; 255:1450–4.
11. Eisenberg J. Doctor's decisions and the cost of medical care. In: Do education and feedback change doctors' decisions? Ann Arbor, Mich: Health Administration Press, 1986:100–24.
12. Hamblin JE, Brock CD, Litchfield L, Dias J. Papanicolaou smear adequacy: effect of different techniques in specific fertility states. J Fam Pract 1985; 20:257–60.
13. The 1988 Bethesda System for reporting cervical/vaginal cytological diagnosis. National Cancer Institute workshop. JAMA 1988; 262:65–8.
14. Kivlahan C, Ingram E. Improved yield of endocervical cells on Papanicolaou smears in a residency setting, J Fam Pract 1985; 20:381–5.
15. Buntinx F, Knottnerus JA, Crebolder HFJM, Essed GGM, Schouten H. Relation between quality of cervical smears and probability of abnormal results. BMJ 1992; 304:1224–6.
16. Mugford M, Barfield P, O'Hanlon M. Effects of feedback of information on clinical practice: a review. BMJ 1991; 303:398–402.
17. Winkins RAG, Pop P, Grol RPTM, Kester ADM, Knotterus JA. Effect of feedback on test ordering behavior of general practitioners. BMJ 1992; 304:1092–5.