Int J Clin Exp Pathol (2008) 1, 285-290
www.ijcep.com/IJCEP708013

Original Article

Papanicolaou Test in the Detection of High-Grade Cervical Lesions: A Re-evaluation Based on Cytohistologic Non-correlation Rates in 356 Concurrently Obtained Samples

Bhavini Carns[1, 2, 3] and Oluwole Fadare[1, 4]

[1]Department of Pathology, Wilford Hall Medical Center, Lackland AFB, TX 78236, USA; [2]Department of Pathology, Brooke Army Medical Center, Fort Sam Houston, TX 78234, USA; [3]Pathology Program, San Antonio Uniformed Services Health Education Consortium, San Antonio, TX 78236 and [4]Department of Pathology, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA

Received 25 Aug 2007; accepted and available online 20 Sept 2007

Abstract: Studies evaluating the routine Papanicolaou (Pap) test have traditionally used as the reference gold standard, the diagnoses on the follow-up histologic samples. Since the latter are typically obtained days to weeks after the Pap test, the accuracy of the resultant comparison may be affected by interim factors, such as regression of human papillomavirus, new lesion acquisitions or colposcopy-associated variability. A subset of our clinicians have routinely obtained cervical cytology samples immediately prior to their colposcopic procedures, which presented a unique opportunity to re-evaluate the test performance of liquid-based cervical cytology in detecting the most clinically significant lesions (i.e. cervical intraepithelial neoplasia 2 or worse: CIN2+), using as gold standard, diagnoses on cervical biopsies that were essentially obtained simultaneously. For each patient, cytohistologic non-correlation between the Pap test and biopsy was considered to be present when either modality displayed a high-grade squamous intraepithelial lesion (HGSIL)/CIN2+ while the other displayed a less severe lesion. Therefore, HGSIL/CIN2+ was present in both the Pap test and biopsy in true positives, and absent in both modalities in true negatives. In false positives, the Pap test showed HGSIL while the biopsy showed less than a CIN2+. In false negatives, Pap tests displaying less than a HGSIL were associated with biopsies displaying CIN2+. Combinations associated with "atypical" interpretations were excluded. A cytohistologic non-correlation was present in 17 (4.8%) of the 356 combinations reviewed. The non-correlation was attributed, by virtue of having the less severe interpretation, to the Pap test in all 17 cases. There were 17, 322, 0, and 17 true positives, true negatives, false positives and false negatives respectively. The sensitivity, specificity, positive predictive value and negative predictive value of the Pap test, at a diagnostic threshold of HGSIL, in identifying a CIN2+ lesion were 50%, 100%, 100% and 95% respectively. Even in Pap test/biopsy combinations obtained on the same day by the same colposcopist and evaluated by the same pathologist, there is a 4.8% (17/356) false negative rate associated with the Pap test. Our findings suggest that there may be an intrinsic error rate associated with this test modality.

Key Words: Cytohistologic correlation, Pap test, accuracy, sensitivity, specificity, performance, colposcopy

Introduction

As a screening modality that is ultimately geared towards the reduction of cervical cancer-related mortality by earlier detection, the Papanicolaou (Pap) test has been remarkably successful [1-6]. Nonetheless, worldwide, cervical cancer remains the 2nd most commonly diagnosed non-cutaneous malignancy and is associated with the 5th highest cancer-related mortality [7]. The vast majority of the latter is generally attributed to the absence of well-developed cancer screening programs in many developing countries [7, 8, 9]. However, even in countries where these programs are robust, thousands of cervical cancer-related preventable deaths

The views expressed in this article are those of the authors and do not reflect the official policy of the Department of Defense or other Departments of the United States Government.

occur annually [10]. Failures of a Pap test-based cervical cancer screening program in preventing otherwise preventable deaths may be attributed to several factors, including limited access to healthcare by a subset of the population, failure of eligible women to present for screening and/or follow-up, suboptimal sampling, lesional and processing-related factors, interpretation errors, and clinical management deficiencies [11]. Since the presence of any of the aforementioned factors may theoretically preclude effective screening for an individual patient, coupled with the concept of a widely performed screening test for a disease whose prevalence is relatively low, it may be anticipated that Pap test-based cervical cancer screening programs would be associated with an inherent error rate.

Previous studies evaluating the accuracy of the Pap test have traditionally used as the reference gold standard, the diagnoses on the follow-up histologic samples [12, 13, 14]. Indeed, cytohistologic correlation studies have been an integral component of quality improvement and quality assurance programs in most cytopathology laboratories in the United States for many years [15]. However, since the cervical biopsies are typically obtained days to weeks after the Pap test, the accuracy of the resultant comparison may be affected by interim factors, such as regression of human papillomavirus infection, new lesion acquisitions or colposcopy-associated variability. A subset of our clinicians have routinely obtained liquid-based cytologic and histologic samples concurrently, which presented a unique opportunity to re-evaluate the test performance of the Pap test in detecting the most clinically significant lesions (i.e. cervical intraepithelial neoplasia 2 or worse: CIN2+), using as gold standard, diagnoses on cervical biopsies that were essentially obtained simultaneously.

Material and Methods

Following approval from our Institutional Review Board, the computerized database of the department of pathology at Wilford Hall Medical Center (WHMC, Lackland Air Force Base, San Antonio, TX) was searched for all patients with a Pap test and cervical biopsy that were accessioned within 24 hours of each other for the period 1/3/06-7/27/07. The cytopathology division at WHMC is a large reference laboratory that receives cytology samples from variably sized military facilities throughout the United States. The screened population is comprised predominantly of young military women. In 2006, over 150,000 Pap tests were processed.

Based on telephonic interviews of a representative cross section of submitting clinicians, the Pap test is typically obtained before the start of the colposcopic procedure, after which the latter proceeds routinely. The clinicians included residents, attending physicians and nurse practitioners, and were selected in a largely random fashion based on contactability, diversity of practice setting and geographic location. They represented approximately 37% of all clinicians that submitted cytohistologic combinations for the study period. In all instances, the same individual performed both the Pap test and the biopsy. The stated objective for obtaining concurrent samples in all individuals interviewed was the possibility of increasing diagnostic yield.

For each patient, diagnostic discordance (cytohistologic non-correlation) between the Pap test and biopsy was considered to be present when either modality displayed a high grade squamous intraepithelial lesion (HGSIL)/CIN2+ while the other modality displayed a less severe interpretation [i.e. no dysplasia or CIN 1 in biopsies or negative for intraepithelial lesion or malignancy (NILM) or low grade squamous intraepithelial lesion (LSIL) in Pap tests]. Therefore, HGSIL/CIN2+ was present in both the Pap test and cervical biopsy in true positives and absent in both modalities in true negatives. In false positives, the Pap showed HGSIL while the cervical biopsy showed less than a CIN2+ (i.e. CIN 1 or negative). In false negatives, Pap tests displaying less than a HGSIL were associated with concurrent cervical biopsies displaying CIN2+. For the purpose of this study, CIN2+ included CIN2, CIN3, adenocarcinoma-in-situ, invasive squamous cell carcinoma, and adenocarcinoma. Combinations associated with ASC-US (atypical squamous cells of undetermined significance) or ASC-H (Atypical squamous cells, cannot exclude HGSIL interpretations) were excluded, as were cases in which either the Pap test and/or the cervical biopsy were deemed unsatisfactory for pathologic evaluation. All slides were reviewed for the remaining patients to confirm the

reference interpretations. In our laboratory, any Pap test/cervical biopsy combinations for the same patient are flagged and segregated to ensure that the same pathologist signs out the diagnostic reports for both samples.

The sensitivity, specificity, positive predictive value and negative predictive value for the Pap test (at a HGSIL threshold) for detecting CIN2+, using the diagnoses on the concurrently obtained cervical biopsies as the reference gold standard, was then determined. Sensitivity was calculated by dividing true positives by the sum of true positives and false negatives (X100); Specificity was calculated by dividing true negatives by the sum of true negatives and false positives (X100); positive predictive value was calculated by dividing true positives by the sum of true positives and false positives (X100); negative predictive value was calculated by dividing true negatives by the sum of true negatives and false negatives (X100).

Results

There were a total of 606 patients with Pap test/cervical biopsy combinations for the study period. Nine combinations were excluded due to insufficient material for pathologic evaluation in either the Pap test (n=4) or biopsy (n=5). 241 of the remaining 597 combinations were excluded because the Pap test involved an ASC-H (n=30) or ASC-US (n=211) interpretation. The remaining 356 combinations formed the basis for this study. All cytologic preparations were liquid-based (ThinPrep®, Cytyc Corporation, Marlborough, MA).

A cytohistologic non-correlation was present in 17 (4.8%) of the 356 combinations. The cytohistologic non-correlation was attributed, by virtue of having the less severe interpretation, to the Pap test in all 17 cases. Upon review, all 17 cases had cells diagnostic of LGSIL but no HGSIL cells. There were 17, 322, 0, and 17 true positives, true negatives, false positives and false negatives respectively. Using the cervical biopsy diagnoses as the gold standard, the sensitivity, specificity, positive predictive value and negative predictive value of a HGSIL interpretation on a Pap test in identifying a CIN2+ lesion was 50%, 100%, 100% and 95% respectively.

The present study design was primarily geared towards the detection of CIN2+ by the Pap test i.e. performance of the Pap test, at a diagnostic threshold of HGSIL, in detecting CIN2+. However, since the Pap test is a screening modality, abnormalities on which should theoretically trigger some further action, we performed additional analyses on all 597 cases with differing thresholds and definitions. The distribution of interpretations in all 597 combinations is outlined in **Table 1**.

In the second analysis, we investigated the ability of any cytologic abnormality (ASC-US and above, ASCUS+) to detect CIN2+. In this 2nd analysis, true positives were defined as an ASC-US+ interpretation on the Pap test and a CIN2+ on the biopsy, true negatives as NILM on the Pap test and no CIN2+ on the biopsy, false positives as ASC-US+ on the Pap test and no CIN2+ on the biopsy, and false negatives as NILM on the Pap test and CIN2+ on the biopsy. Using these definitions, there were 68, 124, 405, and 0 true positives, true negatives, false positives and false negatives respectively. The calculated negative predictive value, positive predictive value, sensitivity and specificity of a ASC-US+ Pap test interpretation for detecting CIN2+ was 100%, 14.4%, 100%, 23.4%. Therefore, the large amount of false positives attributed to ASC-US interpretations significantly improved sensitivity and negative predictive value while unacceptably lowering positive predictive value and specificity.

In the third analysis, we investigated the ability of a Pap test, at an interpretation threshold of LGSIL and above (LGSIL+), to detect CIN2+. ASC-H was considered to be above LGSIL in severity for the purposes of this analysis. True positives were defined as a LGSIL+ interpretation on the Pap test and a CIN2+ on the biopsy, true negatives as NILM or ASC-US on the Pap test and no CIN2+ on the biopsy, false positives as LGSIL+ on the pap test and no CIN2+ on the biopsy, and false negatives as NILM or ASC-US on the Pap test and CIN2+ on the biopsy. Using these definitions, there were 50, 317, 212, and 18 true positives, true negatives, false positives and false negatives respectively. The calculated negative predictive value, positive predictive value, sensitivity and specificity of a LGSIL+ Pap test interpretation for detecting CIN2+ was 94.6%, 19.1%, 73.5% and 60% respectively.

Table 1 Distribution of interpretations in 597 concurrently obtained cervical biopsies and Pap tests

Pap test interpretation	Cervical biopsy diagnosis	Number
LSIL	Negative for Dysplasia	11
LSIL /ASC-H	Negative for Dysplasia	2
NILM	CIN, grade 1	83
LSIL	CIN, grade 2-3	17
NILM	Negative for Dysplasia	41
ASC-US	Negative for Dysplasia	48
ASC-H	Negative for Dysplasia	1
ASC-US	CIN, grade 1	145
ASC-H	CIN, grade 1	4
LSIL	CIN, grade 1	187
LSIL / ASC-H	CIN, grade 1	7
ASC-US	CIN, grade 2-3	17
ASC-H	CIN, grade 2-3	7
HGSIL	CIN, grade 2-3	16
LSIL / ASC-H	CIN, grade 2-3	9
ASC-US	CIN, grade 1 adenocarcinoma in-situ	1
HGSIL	Microinvasive squamous cell carcinoma	1
Total		597

LSIL, low grade squamous intraepithelial lesion; HGSIL, high grade squamous intraepithelial lesion; NILM, negative for Intraepithelial lesion or malignancy; ASC-H, atypical squamous cells, cannot exclude HGSIL; CIN, cervical intraepithelial neoplasia; ASC-US, atypical squamous cells of undetermined significance

The performance of the Pap test in detecting CIN2+, stratified by the various definitional thresholds outlined above, is shown in **Table 2**.

The average age of the 356 women in our study dataset was 27.8 years (95% CI ±8.64; range 16-65, median 25). The average age for the entire data set of 597 women was 28.39 years (95% CI ±8.84; range 16-65, median 26).

Discussion

The accuracy of the Pap test has traditionally been determined using the diagnoses on the follow-up histologic samples as the gold standard [12, 13, 14]. Although undoubtedly practical, this determination is potentially fraught with many errors related to factors arising in the interim, such as regression of human papillomavirus infection, new lesion acquisitions or colposcopy-associated variability. In the active medicolegal environment in which gynecologic cytopathology exists [16, 17], it is imperative that there be continuous research efforts aimed at reaffirming the error rate of the Pap test in routine practice settings. The central objective of the present study was to re-evaluate the performance of the liquid-based Pap test, at a diagnostic threshold of HGSIL, in detecting CIN2+. Notable distinctions between the present study and many others include the entire composition of our dataset of liquid-based samples and the nearly concurrent retrieval of the cytologic and histologic samples. Our study was designed to limit potentially confounding factors. The Pap tests and cervical biopsies were obtained by the same individual in most instances; both were obtained on the same day, and both were evaluated by the same pathologist.

Using the cervical biopsy as the gold standard, we found the sensitivity, specificity, positive predictive value and negative predictive value of a HGSIL interpretation on a Pap test in identifying a CIN2+ lesion was 50%, 100%, 100% and 95% respectively. We identified 17 (4.8%) false negatives out of 356 Pap test/cervical biopsy combinations, i.e. 17 cases in which the biopsy showed CIN2+ while

Table 2 Performance of Pap test in detecting CIN2+ with different definitional thresholds

Parameter	Definitional thresholds on the Papanicolaou test		
	High-grade squamous intraepithelial lesion	Low-grade squamous intraepithelial lesion+	Any cytologic abnormalities
NPV	95%	94.6%	100%
PPV	100%	19.1%	14.4%
Sensitivity	50%	73.5%	100%
Specificity	100%	60%	23.4%

CIN, cervical intraepithelial neoplasia; PPV, positive predictive value; NPV, negative predictive value

the concurrently obtained Pap test showed less than a HGSIL. The fact that all of our CIN2+-related cytohistologic non-correlating cases were attributable to the Pap test suggests that sampling error is a significant contributor to screening errors. However, the possibility that a subset of our false negatives is actually attributable to the concurrence of the sample collections is worthy of note. Some authors have found no relationship between the Pap test-to-cervical biopsy interval and the false negative rate of the Pap test [18]. However, several others have reported that Pap tests obtained within a short interval after a previous one tend to have a low sensitivity and a high false negative rate [15, 19, 20, 21, 22]. Cells on the most superficial layers of the ectocervix may require a certain "regeneration" interval following a Pap test to achieve full exfoliability for the next test. Since almost all the patients in our dataset were referred for colposcopy following previous abnormal Pap tests, this may have contributed to the false negative rate. Secondly, the colposcopist's efforts to preserve the surface of the cervix, with the knowledge of an impending colposcopic examination, may have resulted in suboptimally aggressive Pap tests [23]. Finally, there is some published data that suggests that the false negative rate of the Pap test is higher in younger women [24]. As previously noted, the average age of the 356 women in our study dataset was 27.8 years.

The sensitivity of the Pap test, at a diagnostic threshold of HGSIL, in detecting CIN2+ in the concurrently obtained sample, was found to be 50%. This is considerably lesser than the 76.1% and 100% that have been reported in the 2 largest studies of concurrently obtained samples that have specifically examined the question [15, 27]. Both were based on conventional Pap smears. A closer examination of their data indicated that the observed differences in results are largely attributable to threshold definition. In the study of DiBonito et al [27], the cytologic threshold used was "any cytologic cervical abnormality". In the College of American Pathologists study, the cytologic threshold was essentially identical, as sensitivity was defined as the ability of "cervical cytology to identify correctly the presence of a lesion or malignancy in the biopsy" [15]. At a cytologic threshold of "any cytologic abnormality", our sensitivity is also 100% but specificity drops to 14.4%. Our study confirmed the well-established negative correlation between the specificity and sensitivity of the Pap test [14]. Using the CIN2+ detection endpoint, sensitivity can be substantially increased by lowering the definitional threshold for abnormality in the Pap test, at a significant cost to specificity, and vice versa (**Table 2**). In the study of Mayeaux et al [21], a repeat conventional Pap smear prior to colposcopy had a sensitivity of 48% for CIN and only 25% for high-grade lesions.

The present study was designed to answer the most clinically relevant question, i.e. whether the Pap test will capture a CIN2+ if it is present. Our data indicates that the Pap test will indeed capture CIN2+ in 95.2% of cases. However, even in Pap test/biopsy combinations obtained on the same day by the same colposcopist and evaluated by the same pathologist, there is a 4.8% (17/356) false negative rate associated with the Pap test. Our findings suggest that there is probably an intrinsic error rate associated with this screening test modality. Practitioners should continuously educate participants in Pap test-based screening programs that it is associated with a certain false negative rate.

Please address all correspondences to Oluwole Fadare, MD, Director of Surgical Pathology, Department of Pathology, Wilford Hall Medical

Center, 2200 Bergquist Dr., Ste 1, Lackland AFB, TX 78236, USA. E-mail address: oluwolefadare@yahoo.com

References

[1] Koss LG. The Papanicolaou test for cervical cancer detection: a triumph and tragedy. *JAMA* 1989;261:737-743.

[2] Wingo PA, Cardinez CJ, Landis SH, Greenlee RT, Ries LA, Anderson RN and Thun MJ. Long-term trends in cancer mortality in the United States, 1930-1998. *Cancer* 2003;97:3133-3275.

[3] Adami HO, Ponten J, Sparen P, Bergstrom R, Gustafsson L and Friberg LG. Survival trend after invasive cervical cancer diagnosis in Sweden before and after cytologic screening, 1960-1984. *Cancer* 1994;73:140-147.

[4] Benedet JL, Anderson GH and Matistic JP. A comprehensive program for cervical cancer detection and management. *Am J Obstet Gynecol* 1992;166:1254-1259.

[5] van de Graaf Y, Vooijs GP and Zielhuis GA. Cervical screening revisited. *Acta Cytol* 1990; 34:366-372.

[6] Devesa SS, Young JL Jr, Brinton LA and Fraumeni JF Jr. Recent trends in cervix uteri cancer. *Cancer* 1989;64:2184-2190.

[7] Stewart BW and Kleihues P. World Cancer Report. IARC Press, Lyon, 2003.

[8] Suba EJ, Raab SS: Viet/American Cervical Cancer Prevention Project. Papanicolaou screening in developing countries: an idea whose time has come. *Am J Clin Pathol* 2004; 121:315-320.

[9] Suba EJ, Murphy SK, Donnelly AD, Furia LM, Huynh ML and Raab SS. Systems analysis of real-world obstacles to successful cervical cancer prevention in developing countries. *Am J Public Health* 2006;96:480-487.

[10] Jemal A, Siegel R, Ward E, Murray T, Xu J and Thun MJ. Cancer statistics, 2007. *CA Cancer J Clin* 2007;57:43-66.

[11] DeMay RM. The Pap test. ASCP Press, Chicago, 2005.

[12] Raab SS. Diagnostic accuracy in cytopathology. *Diagn Cytopathol* 1994;10:68-75.

[13] Nanda K, McCrory DC, Myers ER, Bastian LA, Hasselblad V, Hickey JD and Matchar DB. Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review. *Ann Intern Med* 2000;132:810-819.

[14] Fahey MT, Irwig L and Macaskill P. Meta-analysis of Pap test accuracy. *Am J Epidemiol* 1995;141:680-689.

[15] Jones BA and Novis DA. Cervical biopsy-cytology correlation. A College of American Pathologists Q-Probes study of 22 439 correlations in 348 laboratories. *Arch Pathol Lab Med* 1996;120:523-531.

[16] Slater DN. False-negative cervical smears: medico-legal fallacies and suggested remedies. *Cytopathology* 1998;9:145-154.

[17] Frable WJ. Error reduction and risk management in cytopathology. *Semin Diagn Pathol* 2007;24:77-88.

[18] Beeby AR, Wadehra V, Keating PJ and Wagstaff TI. A retrospective analysis of 94 patients with CIN and false negative cervical smears taken at colposcopy. *Cytopathology* 1993;4:331-337.

[19] Wheelock JB and Kaminski PF. Value of repeat cytology at the time of colposcopy for the evaluation of cervical intraepithelial neoplasia on Papanicolaou smears. *J Reprod Med* 1989; 34:815-817.

[20] Davis GL, Hernandez E, Davis JL and Miyazawa K. Atypical squamous cells in Papanicolaou smears. *Obstet Gynecol* 1987;69:43-46.

[21] Mayeaux EJ Jr, Harper MB, Abreo F, Pope JB and Phillips GS. A comparison of the reliability of repeat cervical smears and colposcopy in patients with abnormal cervical cytology. *J Fam Pract* 1995;40:57-62.

[22] Jones DE, Creasman WT, Dombroski RA, Lentz SS and Waeltz JL. Evaluation of the atypical Pap smear. *Am J Obstet Gynecol* 1987; 157:544-549.

[23] Zardawi IM and Rode JW. Clinical value of repeat Pap smear at the time of colposcopy. *Acta Cytol* 2002;46:495-498.

[24] Paterson ME, Peel KR and Joslin CA. Cervical smear histories of 500 women with invasive cervical cancer in Yorkshire. *Br Med J (Clin Res Ed)* 1984;289:896-898.

[25] Spitzer M, Ryskin M, Chernys AE and Shifrin A. The value of repeat Pap smear at the time of initial colposcopy. *Gynecol Oncol* 1997;67:3-7.

[26] Simsir A, Ioffe OB, Bourquin P, Brooks SE and Henry M. Repeat cervical cytology at the time of colposcopy. Is there an added benefit? *Acta Cytol* 2001;45:23-27.

[27] DiBonito L, Falconieri G, Tomasic G, Colautti I, Bonifacio D and Dudine S. Cervical cytopathology. An evaluation of its accuracy based on cytohistologic comparison. *Cancer* 1993;72:3002-3006.