Comparative Analysis of Conventional Papanicolaou Tests and a Fluid-Based Thin-Layer Method

Anjali Limaye, MD; Amsy J. Connor, CT(ASCP); Xiaohua Huang, MS; Ronald Luff, MD, MPH

● Context.—A fluid-based, direct-to-vial method of thin-layer gynecologic cytology (ThinPrep Pap Test) is reported to be more effective than the conventional Papanicolaou test in the detection of squamous intraepithelial lesions.

Objective.—This retrospective analysis evaluated the validity of the findings on the thin-layer method using case material at a large independent laboratory and represented a comparison of performance of both methods over an identical period.

Methods.—Data for conventional and ThinPrep tests were compared for 2 periods. Period 1 included 1 421 080 conventional and 56 835 ThinPrep specimens, and period 2 included 564 270 conventional and 109 784 ThinPrep specimens. Squamous intraepithelial lesions were used to determine detection of disease. These 2 sets of data were also analyzed to eliminate effects of any selection bias toward ThinPrep for high-risk patients.

Gynecologic cytology as represented by the Papanicolaou (Pap) test has been an important part of preventive medicine. In the United States, use of the Pap test as a screening test has led to a decrease in the incidence of cervical cancer from 14.2 per 100 000 in 1973 to 7.8 per 100 000 in 1994. The ThinPrep Pap Test (Cytyc Corporation, Boxborough, Mass) is a liquid-based cell-collection method. The US Food and Drug Administration (FDA) approved ThinPrep as a replacement for the conventional Pap test in 1996. The clinical trial leading to FDA approval showed a statistically significant increase in the detection of squamous intraepithelial lesions (SILs; both low and high grade) with the ThinPrep method, utilizing a split-sample methodology. Since then, multiple studies have been undertaken to validate the original findings. The ThinPrep sample sizes in these studies have ranged from 1 500 to 56 000. This study was undertaken to analyze case material in direct-to-vial use at the Quest Diagnostics Incorporated (Teterboro, NJ) laboratory and represented a comparison of performance for an identical period of time totaling 27 months for both conventional Pap and ThinPrep tests. Overall, data from approximately 2 million conventional Pap tests and 166 000 ThinPrep tests were analyzed.

Results.—Use of ThinPrep showed a greater than 100% increase in the detection rate of squamous intraepithelial lesions (1.3%–3.4% in period 1 and 1.3%–2.9% in period 2), which was statistically significant after correcting for selection bias. We also found a significant decrease in the false-negative proportion (57% in period 1 and 35% in period 2). There was a marked improvement (233%) in the detection of high-grade squamous intraepithelial lesions in high-risk cases and a decrease in the atypical squamous cells of undetermined significance to squamous intraepithelial lesion ratio from 3.1 to 1.5 in period 2.

Conclusion.—ThinPrep is better than the conventional Papanicolaou test in detecting squamous intraepithelial lesions and is a superior screening test in detection of precancerous changes of the cervix.

For editorial comment, see p 143.

MATERIALS AND METHODS

The cases consisted of all the gynecologic cytology specimens sent to the laboratory and included all ThinPrep and conventional Pap tests submitted to the laboratory from July 1997 to October 1999. No cases were excluded. All specimens were collected independently, either as conventional Pap tests or as ThinPrep specimens. The specimens were obtained mainly from outpatient medical practices in the greater New York–New Jersey area, which included principally gynecologists and family practitioners. Fewer than 2% of the specimens were submitted from hospital-based physicians. The study was done over 2 periods. Period 1 represented the pilot study and extended from July 1997 through December 1998. Period 2 included January through October 1999. More detailed data were available for period 2, allowing analysis of additional variables, including SILs in the high-risk group. Between the 2 periods, many physicians converted their practices from conventional Pap tests to ThinPrep, leading to a greater percentage of ThinPrep tests during period 2. In period 1, ThinPrep represented 4% of all gynecologic cytology cases, and in period 2, ThinPrep constituted 16% of the cases. The cases represented a typical screening population, as defined in earlier studies, with a less than 5% prevalence of SILs. The patients included a wide range of ages and reproductive histories, as indicated by the ordering physicians on the laboratory requisition form. The history included information that the lab-

Reprints: Anjali Limaye, MD, Department of Anatomic Pathology, Quest Diagnostics Incorporated, Teterboro, NJ.

Presented as a poster at the 48th Annual Scientific Meeting of the American Society of Cytology, Philadelphia, Pa, November 2000.

Accepted for publication August 7, 2002.

From the Department of Anatomic Pathology, Quest Diagnostics Incorporated, Teterboro, NJ.

For editorial comment, see p 143.
was defined as SILs discovered on random rescreening and con-
both conventional and ThinPrep smears. Estimated false negative
(high-grade included all cases of moderate and severe dysplasia
mild dysplasia (cervical intraepithelial neoplasia [CIN] 1), and
false-negative cases. In our laboratory, all high-risk cases, when
normal limits, including both low- and high-risk cases, to detect
cases initially interpreted by a single cytotech nologist as within
Amendments of 1988 [CLIA '88]) to reevaluate at least 10% of
viewed by 1 cytotechnologist. The cytology laboratory was re-
verted to an estimate based on 80% of normal cases being re-
viewed by at least 2 observers (cytotechnologist and/or pathologist); thus, 80% of normal cases were reviewed by 1 cytotechnologist. The false-negative proportion was defined as follows: False-Negative Cases/(True Positive + False-Negative Cases). This proportion had been consid-
ered the best measure of cytotechnologist and laboratory perfor-
ance in detecting cytologic evaluation errors. To determine the false-negative proportion, it is necessary to know the total num-
ber of false-negative cases. An estimated total number of false-
negative cases could be determined by extrapolating the number of false-negative cases detected on random quality control re-
screening to 80%, which represents the percentage of normal cas-
es reviewed by a single cytotechnologist. Using the estimated false negative in the false-negative proportion formula, it be-
comes an estimated false-negative proportion. A potential pitfall in using the estimated false-negative proportion is in not know-
ing the error rate of the reevaluation process. However, that prob-
lem was generally controlled for by having the same cytotechn-
ologists follow the same procedure for interpreting quality control
cases for both conventional and ThinPrep smears.

Four separate computer simulations were done to test the fol-
lowing null hypotheses: equal SIL detection rates for conventional
and ThinPrep smears for periods 1 and 2, equal SIL detection rates for conventional and ThinPrep smears in the high-risk group in period 2, and equal high-grade SIL detection rates for conventional and ThinPrep specimens in the high-risk groups of period 2. The simulations were based on the assumption of a common sensitivity of 0.95 and a common prevalence for SILs in the high-risk groups for both test types and a different common prevalence for SILs in the non-risk groups. In our laboratory with longitudinal studies for more than 20 years, the estimated false-negative proportion has been consistently around 5%. Since sensitivity equals 1 minus the false-negative proportion, the sen-
sitivity for detection of abnormal cells is 95% (likewise, estimat-
ed). The prevalence was based on the expected percentage of true
diseased patients and should be the same for conventional and
ThinPrep smears. The prevalence values were varied from 1% to 6% in the simulations; the conclusions of the simulations re-
mained unchanged over this range. Using random binominal num-
bers and the normal approximation, computer simulations were
run 10000 times for each hypothesis to model the detection ratios
for periods 1 and 2. The ratio of SIL detection rates for the 2
methods was computed from each simulation and compared with
the observed ratio. The P value was computed as the propor-
tion of the number of times the simulated ratios were greater
than the observed ones and was used to determine whether the
observed ratio was statistically significant. The design of the sim-
ulation accounted for any selection bias toward ThinPrep for
high-risk patients. The S-Plus statistical software (MathSoft Inc,
Seattle, Wash) was used for the simulations. P values were de-
rivered using either the simulation or the chi² tests. The significance
level was set to .05.

RESULTS
Table 1 shows that in period 1 the detection of SILs increased by 160% (from 1.3% to 3.4%) in ThinPrep

| Table 1. The Number of High-Risk Patients and the Percentage of Squamous Intraepithelial Lesions in the Total Population in Period 1* |
|-----------------|-------------|----------------|-------------|----------------|
|                 | CPAP %  | TPPT % | Ratio of %  | P |
| High-risk patients | 50.762 | 4865 | 1.24 | . . . |
| Total population  | 1421 080 | 56835 | 3.4 | 1.26 | <.001 |
| SIL in total population | 17921 | 1930 |  |

* CPAP indicates conventional Papanicolaou test; TPPT, ThinPrep Pap test; and SIL, squamous intraepithelial lesion.
smears compared with conventional smears. There was a selection bias for ThinPrep in the high-risk population (1:2.4). The simulation showed that the increase in detection was statistically significant ($P < .001$), even in the presence of this selection bias.

Table 2 shows the results from period 2. These data are similar to those of period 1, with the selection bias of high-risk cases having decreased (1:1.6). The detection of SILs increased by 130%, from 1.3% to 2.9%. Period 2 data allowed for a direct comparison of detection in the high-risk population for both conventional Pap smears and ThinPrep. For the high-risk population, the detection of SILs increased by 160%, from 2.7% to 6.9%. Simulation showed that both these increases were statistically significant ($P < .001$), even in the presence of selection bias.

Table 3 shows that follow-up histologic (biopsy) diagnoses for SILs had essentially similar rates for conventional Pap tests and for ThinPrep. The difference was not statistically significant ($P = .001$). The detection of low-grade SILs increased by 59%, from 2.7% to 4.3%. The detection of high-grade SILs increased by 60%, from 0.18% to 0.57%. Simulation showed that both these increases were statistically significant ($P < .001$).

Table 4 shows that for both periods 1 and 2 there was a statistically significant decrease in false-negative proportion with use of ThinPrep. A 57% reduction, from 6.9% to 3.0%, was observed for period 1. A 35% reduction, from 9.1% to 5.9%, was observed for period 2. These decreases in the estimated false-negative proportion for both periods are statistically significant using the $\chi^2$ test ($P < .001$).

Table 6 compares the detection of low- and high-grade SILs in the high-risk group in period 2, a group for which the selection bias had been already eliminated, allowing direct comparison within this group. The detection of high-grade SILs increased by 233%, from 0.3% to 1.0%, with ThinPrep. This increase was statistically significant ($P < .001$ using the simulation). The detection of low-grade SILs increased by 59%, from 2.7% to 4.3%. The detection rate of cancer was 0.004% in ThinPrep specimens and 0.003% in conventional Pap tests.

Additional analysis of period 2 data showed a decrease in the ASCUS-SIL ratio in ThinPrep from 3.1 to 1.5. Although there was an increase in ASCUS to 4.4% in ThinPrep from 3.8% in conventional Pap tests, the diagnosis of SIL also increased to 2.9% in ThinPrep from 1.2% in conventional Pap tests, leading to an overall decrease in the ASCUS-SIL ratio.

Table 7 shows that in period 2, the number of satisfactory-but-limited cases was 22.1% in ThinPrep versus 27.7% in conventional Pap tests. The number of satisfactory-but-limited cases due to absent endocervical component was 12.2% of the total population in ThinPrep versus 14.8% in conventional Pap tests. Both these findings are statistically significant ($P < .001$) using the $\chi^2$ test. Other causes of satisfactory-but-limited designations included lack of age or history of last menstrual period, which was approximately 7.5% in both types of specimens; partially obscuring inflammation, which was 1.7% using ThinPrep versus 3.8% using conventional Pap tests; partially obscuring blood, 0.2% using ThinPrep versus 1.0% using conventional Pap tests; and scant cellularity, 0.4% with ThinPrep and 0.2% with conventional Pap tests.

### Table 2. The Number of High-Risk Patients and the Percentage of Squamous Intraepithelial Lesions in the High-Risk and Total Population in Period 2*

<table>
<thead>
<tr>
<th></th>
<th>CPAP</th>
<th>TPPT</th>
<th>Ratio of</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-risk patients</td>
<td>15 341</td>
<td>2.7</td>
<td>4759</td>
<td>4.3</td>
</tr>
<tr>
<td>Total population</td>
<td>564 270</td>
<td>109 784</td>
<td>60%</td>
<td>.001</td>
</tr>
<tr>
<td>SIL in total population</td>
<td>7099</td>
<td>1.3</td>
<td>3232</td>
<td>2.9</td>
</tr>
<tr>
<td>SIL in high-risk population</td>
<td>410</td>
<td>2.7</td>
<td>326</td>
<td>6.9</td>
</tr>
</tbody>
</table>

* CPAP indicates conventional Papanicolaou test; TPPT, ThinPrep Pap test; and SIL, squamous intraepithelial lesion.

### Table 3. Validation of Cytologic Diagnosis of Squamous Intraepithelial Lesion by Histologic Examination*

<table>
<thead>
<tr>
<th></th>
<th>CPAP</th>
<th>TPPT</th>
<th>Percentage</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSIL</td>
<td>259</td>
<td>159</td>
<td>59%</td>
<td>.001</td>
</tr>
<tr>
<td>HSIL</td>
<td>123</td>
<td>74</td>
<td>60%</td>
<td>.001</td>
</tr>
<tr>
<td>$P$</td>
<td>...</td>
<td>...</td>
<td>0.18</td>
<td>...</td>
</tr>
</tbody>
</table>

* CPAP indicates conventional Papanicolaou test; TPPT, ThinPrep Pap test; LSIL, low-grade squamous intraepithelial lesion, and HSIL, high-grade squamous intraepithelial lesion.

### Table 4. False-Negative Proportion for Period 1*

<table>
<thead>
<tr>
<th></th>
<th>CPAP</th>
<th>TPPT</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total No. of cases</td>
<td>1 421 080</td>
<td>56 835</td>
<td>.001</td>
</tr>
<tr>
<td>No. of SILs on initial screen</td>
<td>17 795</td>
<td>1925</td>
<td>.76</td>
</tr>
<tr>
<td>Random rescreening, % (QC)</td>
<td>7.6</td>
<td>6.8</td>
<td>.18</td>
</tr>
<tr>
<td>No. of SILs in QC review</td>
<td>126</td>
<td>5</td>
<td>.57</td>
</tr>
<tr>
<td>Total No. of SILs in QC review (eFN) estimated on 80% single screen</td>
<td>1321</td>
<td>59</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>False-negative proportion</td>
<td>6.9%</td>
<td>3.0%</td>
<td>.001</td>
</tr>
</tbody>
</table>

* CPAP indicates conventional Papanicolaou test; TPPT, ThinPrep Pap test; SIL, squamous intraepithelial lesion; QC, quality control; and eFN, estimated false negative.
were 0.4% unsatisfactory specimens using ThinPrep versus 0.27% using conventional Pap tests.

**COMMENT**

The ThinPrep test was initially approved by the FDA based on split-sample analysis. Subsequently, other studies have validated the tests in the premarket approval process, including a large study by Hutchinson et al. in 8000 high-risk patients. Additionally, since that approval, several studies favorably comparing ThinPrep to conventional Pap tests in direct-to-vial use have been reported. These include the study by Carpenter and Davey, which was based on 2727 ThinPrep specimens from a high-risk university hospital practice, and that by Guidos and Selvaggi, which included 9583 ThinPrep specimens from a mixture of both a screening population and high-risk patients. The study by Bolick and Hellman focused on a screening population and included 10694 ThinPrep specimens. The largest previously published study compared 56339 ThinPrep specimens with 74756 conventional Pap tests in a screening population. However, none of those studies corrected for a possible clinician selection bias for high risk patients to have a ThinPrep test instead of a conventional Pap test. The focus of this study was to assess the presence of such a selection bias and to ascertain whether it eliminated previously reported improved detection of disease by use of ThinPrep. Indeed, our study showed a substantial selection bias favoring high-risk patients as identified by ordering physicians on the requisition to have ThinPrep tests in both periods of study, although the bias was less evident in the second period (1:2.4 vs 1:1.6). We suspect that initially physicians used ThinPrep more often for high-risk patients and that this selection bias waned as use of ThinPrep increased. Despite this selection bias, disease detection defined as SIL (both high and low grade) showed a statistically significant improvement ($P < .001$). Squamous intraepithelial lesion was chosen to define disease detection, based on its higher intraobserver and interobserver reproducibility compared with inclusion of borderline epithelial abnormalities (ASCUS plus SIL cases). However, since SIL cases are substantially less frequent than ASCUS cases, a larger population must be compared to allow for a statistically valid comparison. Fortunately, the high volume of both conventional and ThinPrep tests in our laboratory allowed for such comparison.

The cytologic results were validated by a subset analysis of histologic diagnoses of subsequent biopsies. The validation for ThinPrep tests and the conventional Pap test was statistically similar for both low- and high-grade SILs, confirming a lack of selection bias in the reporting of the cytologic results for either technique. This is consistent with previous studies and the findings corroborate ongoing intralaboratory CLIA ‘88–required cytohistologic correlation. Additional studies showed a decreased ratio of ASCUS to SIL for ThinPrep compared with conventional Pap tests due to an absolute increase in the detection of SILs (Table 6), consistent with decreases observed in other studies. Impressively, our study showed that high-grade SIL detection using ThinPrep in high-risk women was higher (233%) than that with conventional Pap tests. The detection rate of cancer was 0.004% in ThinPrep specimens and 0.003% in conventional Pap tests. The number of satisfactory-but-limited cases was 22.1% in ThinPrep versus 27.7% in conventional Pap tests, with fewer cases lacking endocervical component in ThinPrep tests (12%) compared with conventional Pap tests (14%). There were 0.4% unsatisfactory specimens using ThinPrep versus 0.27% using conventional Pap tests. During the period of this study, preprocessing methods, as suggested by Bentz et al., were not used to improve specimen adequacy. Overall, our findings are similar to those observed by Bernstein et al. in their meta-analysis.

The conventional Pap test is an imperfect screening test that has an overall estimated sensitivity of 51%. One component of the sensitivity level is the laboratory's ability to detect abnormal cells. The inability to detect those abnormal cells is best defined as the false-negative propor-
tion based on a rescreening of once-screened negative cases selected on a random basis. Although the determination of the false-negative proportion has been criticized, when a consistent method of determination of false-negative proportion is applied within the same laboratory, regardless of the screening technique, it should allow for a meaningful comparison. Intuitively, if a new gynecologic cytology screening test detects more disease than its predecessor, there should also be a reduction in the number of false-negative cases. Our study showed such a reduction, which was statistically significant (P < .001), corroborating the findings of Linder and Zahniser. The false-negative proportion of 5.9% for ThinPrep in period 2 is similar to the observations of Belinson et al, although in their study the definition of true positive was based on the presence of high-grade SIL in a biopsy specimen, rather than in the cytology preparation.

In summary, our findings show that the ThinPrep methodology is better than the conventional Pap test in detection of cervical epithelial abnormalities, even accounting for a selection bias toward the use of ThinPrep in high-risk patients. ThinPrep had a greater decrease in false-negative cases compared with conventional Pap tests. These findings strongly support ThinPrep as a superior screening test as compared with the conventional Pap test in detection of precancerous changes of the cervix.

References