

Accuracy of the Papanicolaou Test in Screening for and Follow-up of Cervical Cytologic Abnormalities: A Systematic Review

Kavita Nanda, MD, MHS; Douglas C. McCrory, MD, MHSc; Evan R. Myers, MD, MPH; Lori A. Bastian, MD, MPH; Vic Hasselblad, PhD; Jason D. Hickey; and David B. Matchar, MD

Purpose: To evaluate the accuracy of conventional and new methods of Papanicolaou (Pap) testing when used to detect cervical cancer and its precursors.

Data Sources: Systematic search of English-language literature through October 1999 using MEDLINE, EMBASE, other computerized databases, and hand searching.

Study Selection: All studies that compared Pap testing (conventional methods, computer screening or rescreening, or monolayer cytology) with a concurrent reference standard (histologic examination, colposcopy, or cytology).

Data Extraction: Two reviewers independently reviewed selection criteria and completed 2×2 tables for each study.

Data Synthesis: 94 studies of the conventional Pap test and three studies of monolayer cytology met inclusion criteria. No studies of computerized screening were included. Data were organized by cytologic and histologic thresholds used to define disease. For conventional Pap tests, estimates of sensitivity and specificity varied greatly in individual studies. Methodologic quality and frequency of histologic abnormalities also varied greatly between studies. In the 12 studies with the least biased estimates, sensitivity ranged from 30% to 87% and specificity ranged from 86% to 100%.

Conclusions: Insufficient high-quality data exist to estimate test operating characteristics of new cytologic methods for cervical screening. Future studies of these technologies should apply adequate reference standards. Most studies of the conventional Pap test are severely biased: The best estimates suggest that it is only moderately accurate and does not achieve concurrently high sensitivity and specificity. Cost-effectiveness models of cervical cancer screening should use more conservative estimates of Pap test sensitivity.

Since the implementation of widespread screening with the Papanicolaou (Pap) test, rates of cervical cancer in the United States have decreased from 14.2 per 100 000 in 1973 to 7.8 per 100 000 in 1994. Nevertheless, cervical cancer is still the ninth-leading cause of cancer deaths among U.S. women (1). Most of these deaths occur in women who have never had a Pap test, but some occur in women who recently received negative test results. Approximately two thirds of false-negative results are caused by sampling error, and the rest are caused by detection error.

Sampling error occurs when abnormal cells are not collected or are not transferred to the Pap slide, and detection error occurs when abnormal cells on the Pap slide are missed or misinterpreted. The most common sampling error is lack of cells from the cervical transformation zone. To reduce sampling error, an endocervical cytobrush and a spatula can be used instead of a cotton swab. However, a recent meta-analysis found that the Pap test did not differ in sensitivity or specificity when different sampling devices were used (2). The Food and Drug Administration (FDA) has approved another potential solution: liquid-based monolayer preparation (ThinPrep, Cytoc Corp., Boxborough, Massachusetts). With this technique, the sample is collected as in the conventional Pap test, but cells are then placed in a fixative solution. The cells are dispersed, collected onto a filter, and transferred to a microscopic slide for interpretation. Because samples are fixed immediately after collection, fewer cellular morphologic artifacts occur. Fewer cells on the slide are obscured because the process reduces the amounts of other sampled material, such as blood and mucus, and deposits cells on the slide in a monolayer.

To reduce detection error, some researchers advocate rescreening slides initially reported to be normal. The Clinical Laboratory Improvement Amendments of 1988 mandate rescreening of a 10% random sample of normal slides as a quality assurance measure. Rescreening can also be performed on a higher proportion of slides by using computerized technologies. The FDA has approved

Ann Intern Med. 2000;132:810-819.

For author affiliations and current addresses, see end of text.

two such systems, one that is algorithm-based (AutoPap QC System, TriPath Imaging, Inc., Redmond, Washington), and one that uses neural networks (PAPNET, Neuromedical Systems, Inc., Suffern, New York). PAPNET uses neural network computerized imaging of Papanicolaou smear slides to identify cells or clusters of cells that require review; it then displays up to 128 images per slide that are likely to contain abnormalities. A cytotechnologist reviews these images and decides whether to review the actual slide using light microscopy. AutoPap uses its algorithm-based decision-making technology to identify slides that exceed a certain threshold for the likelihood of abnormal cells. The laboratory can select different thresholds, corresponding to 10%, 15%, or 20% review rates. In contrast to random rescreening, AutoPap selects a sample of slides that is enriched with abnormalities, thereby including most of the slides that contain abnormalities missed by manual screening.

Another approach to reducing detection error is improving the sensitivity of the initial screening step. The FDA has recently approved a new method (AutoPap Primary Screening System, TriPath Imaging) for this indication. AutoPap Primary Screening System uses proprietary computerized algorithms to identify slides that exceed a certain threshold for the likelihood of abnormal cells. A cytotechnologist then reviews these slides. The system allows laboratories to concentrate on the 75% of slides that most likely contain abnormal cells while immediately archiving the remainder.

Sampling and detection errors are reduced when Pap test screening is repeated frequently. However, cost-effectiveness analyses have concluded that if persons are screened more than every 3 years, cost-effectiveness ratios exceed \$50 000 per life-year saved (3, 4).

Precise estimates of cytologic test sensitivity and specificity are important because they may be used to determine policy decisions, such as recommendations for optimal frequency of screening, management of mild abnormalities, and use of newer methods. Our primary objective was to systematically review the operating characteristics of conventional and new methods (computer screening and monolayer slide preparation) of Pap testing in the detection of cervical cancer and its precursors. We also evaluated test performance among women with previous cytologic abnormalities. The Agency for Healthcare Research and Quality (AHRQ), under contract to Duke University (Durham, North Carolina), funded the study. An AHRQ-approved advisory panel assisted in the design, conduct, and reporting of this work, and the evidence report on which this manuscript is based was reviewed by an external peer review panel (5).

Table 1. Search Strategy

1. Vaginal smears/
2. ((Pap or Papan\$) and (smear\$ or test\$)).tw.
3. (PAPNET or autoPap or ThinPrep).tw.
4. 1 or 2 or 3
5. exp Cervix neoplasms/
6. Cervix dysplasia/
7. Cervical Intraepithelial Neoplasia/
8. dyskaryo\$.tw.
9. 5 or 6 or 7 or 8
10. exp "Sensitivity and Specificity"/
11. (sensitivity or specificity).tw.
12. exp Diagnostic errors/
13. 4 and (10 or 11 or 12)
14. 13 and 9
15. limit 14 to (human and English language)
16. Papillomavirus, Human/
17. 15 not 16

Methods

Data Sources

Data sources, including MEDLINE (from 1966), EMBASE (from 1980), HealthStar (from 1975), CancerLit (from 1983), and CINAHL (from 1983) were searched through October 1999 by using a strategy developed with a medical librarian (Table 1). Searches were limited to English-language studies in humans. We manually searched newly published relevant journal issues, bibliographies of included studies, and recent systematic reviews (6–9). To locate unpublished studies, we also contacted relevant professional societies and manufacturers of cytologic devices.

Study Selection

We identified studies of conventional Pap testing (with or without manual rescreening), Pap testing using monolayer slide preparation (ThinPrep), Pap testing with primary computer screening (AutoPap or PAPNET), and Pap testing with computer rescreening (AutoPap or PAPNET). Other recently developed methods, the AutoCyte PREP System and the AutoCyte SCREEN system (TriPath Imaging, Inc., Burlington, North Carolina), had not been approved by the FDA at the time of our review and were not evaluated in this study.

Study samples included women undergoing Pap testing for primary screening and those undergoing evaluation for previous cytologic abnormalities. The main outcome measures were the sensitivity and specificity of the cytologic test for detecting "cases." Cytologic abnormality was defined by one of three thresholds: atypical squamous cells of undetermined significance (ASCUS), low-grade squamous intraepithelial lesions (LSIL), and high-grade squamous intraepithelial lesions (HSIL). "Cases" were defined as histologic diagnosis of cervical intraepithelial neoplasia (CIN), grades I to III, or carcinoma.

| Classification System | Cytology Classification | | | | | | |
|-----------------------|-------------------------|---------------------------------|----------------|--|-------------------|-------------------|--------------------|
| The Bethesda System | | Infection Reactive Repair | ASCUS | Squamous Intraepithelial Lesion (SIL) | | | |
| | | | | Low Grade (LSIL) | High Grade (HSIL) | | |
| Richart | | | Condyloma | Cervical Intraepithelial Neoplasia (CIN) | | | |
| | | | | Grade I | Grade II | Grade III | |
| Reagan (WHO) | Normal | Atypia | Mild Dysplasia | Moderate Dysplasia | Severe Dysplasia | In situ Carcinoma | Invasive Carcinoma |
| Papanicolaou | I | II | III | | IV | | V |

Figure. Map of classification schemes for cervical cytology. ASCUS = atypical squamous cells of undetermined significance; WHO = World Health Organization.

Equivalent categories in other classification schemes (10–14) were also used (Figure).

We included studies of the conventional Pap test if a reference standard of histologic examination or colposcopy was reasonably concurrent to the cytologic screening test (within 3 months) and if sufficient data were reported to complete all four cells of a 2×2 table. Comparison with such a reference standard provides a more relevant outcome for clinical decision makers because colposcopic or histologic diagnoses form the basis of most clinical management decisions.

Only one study of ThinPrep (15) provided enough information to allow us to extract data on sensitivity and specificity compared with a gold standard of histologic examination or colposcopy. We therefore used a separate set of screening criteria for studies of the new methods, based on cytology society guidelines (16, 17) and FDA documents (18): 1) The study must prospectively compare screening tests or test and reference standard on the same set of patients or slides; 2) if cytologic examination is the reference standard, discordant results from the two study tests must be adjudicated by an independent panel of experienced cytology professionals; 3) at least 50% of patients testing positive for HSIL must be verified by histologic examination or colposcopy; and 4) the study design must allow for separate analyses of sensitivity (or relative true-positive rate) and specificity (or relative false-positive rate).

Data based on a cytologic reference standard cannot be integrated with data based on a histologic reference standard (19–23). However, when negative test results are not verified with the reference standard, information about incremental characteristics of test performance may be obtained by directly comparing independently applied conventional and new tests (21). In this case, both tests must be applied independently to all patients, and all positive results on either test must be verified

with the reference standard. A relative true-positive rate and a relative false-positive rate, which can be used to determine relative estimates of the performance of the new test, can then be calculated.

Two investigators independently screened each study. Differences of opinion were reconciled by consensus. The title and abstract of each citation were screened first, and the full report was screened second. Of the 1193 bibliographic references we reviewed, 761 (approximately 64%) were excluded on the basis of title or abstract. We reviewed the full reports of 346 studies of the conventional Pap test and 86 studies of the new methods (18 on AutoPap, 42 on PAPNET, and 26 on ThinPrep).

We developed a numeric quality score to evaluate included citations. Nine members of the study's working group (6 clinicians, 2 economists, and 1 health policy analyst) initially identified more than 12 evaluation criteria on the basis of previously reported criteria (6, 24, 25). We used a consensus process to narrow this list to 7. Blinded to the rest of the group, each participant then independently assigned numerical weights to the criteria. The means of these "votes" were calculated. Each participant received a copy of his or her responses, depicted graphically in relation to the mean for each criterion, and was requested to confirm or reconsider his or her responses. The means of the revised responses for each criterion were then calculated and used as the assigned weights.

Only 94 studies of conventional Pap testing permitted estimation of both sensitivity and specificity. The most common reasons for excluding an article on the conventional Pap test were lack of histologic examination or colposcopy as a reference standard and lack of sufficient data to complete all four cells of a 2×2 table.

Only three studies of ThinPrep (15, 26, 27) permitted estimation of both sensitivity and specificity. No studies of PAPNET, the AutoPap 300 QC, or the AutoPap Primary Screening System permitted

Table 2. Quality Evaluation of 94 Studies of the Conventional Papanicolaou Test

| Quality Criterion (Points) | Studies, n (%) |
|---|----------------|
| Reference standard | |
| Histologic examination (2) | 67 (71) |
| Histologic examination or negative colposcopy or colposcopy (1) | 27 (29) |
| Independence of assessments | |
| Blinded (2) | 23 (24) |
| Not blinded (0) | 71 (76) |
| Verification | |
| All positive and negative test results verified (2) | 48 (51) |
| Positive test results and random fraction of negative test results verified (1) | 1 (1) |
| Positive test results and selected sample of negative test results verified (0) | 45 (48) |
| Study sample | |
| Consecutive or random (2) | 77 (82) |
| Other (0) | 17 (18) |
| Spectrum of disease/nondisease | |
| Defined (1) | 80 (85) |
| Not defined (0) | 14 (15) |
| Publication type | |
| Paper (1) | 94 (100) |
| Abstract (0) | 0 (0) |
| Industry relation | |
| Not performed or supported by a manufacturer (1) | 91 (97) |
| Supported by a manufacturer (0.5) | 1 (1) |
| Performed by a manufacturer (0) | 2 (2) |

estimates of sensitivity or specificity. Studies that applied manual screening followed by computerized rescreeing could not be evaluated because rescreeing is conditional on a negative initial screen; thus, the two tests are not applied independently. The other studies of the new methods were excluded because they were not two-armed prospective studies or because they failed to verify at least 50% of all patients who tested positive for HSIL on histologic examination or colposcopy

Data Extraction

Two reviewers independently completed 2×2 tables for each study. Where available, we abstracted data at four different combinations of cytologic and histologic thresholds: ASCUS/CIN-I, LSIL/CIN-I, LSIL/CIN-II-III, and HSIL/CIN-II-III. Cases that were indeterminate because of uninterpretable cytologic results or lack of reference standard confirmation were documented but were excluded from the calculation of sensitivity and specificity. If more than one study sample was included in a single report and data for each sample were provided separately, we treated the samples as if they had been presented in individual studies.

Data Synthesis

Three studies described the accuracy of thin-layer cytology relative to histologic examination or conventional Pap testing. Bolick and Hellman (15) compared ThinPrep Pap smear diagnoses of LSIL or higher to a histologic reference standard of CIN-II-III or higher, permitting direct estimation of 94.2% sensitivity and 57.7% specificity. Conventionally prepared Pap smears achieved a sensitivity of 84.6% and a specificity of 37.0% according to the same thresholds. Most negative test results in this study were not verified with histologic examination (15). Both Roberts and colleagues (26) and Hutchinson and coworkers (27) compared conventional and ThinPrep slides prepared with a split sample technique and used a combination of cytologic and histologic examination as the reference

Table 3. Summary Statistics for Studies of the Conventional Papanicolaou Test*

| Threshold | Studies, n | Sensitivity | | | Specificity | | | Positive Likelihood Ratio | | | Negative Likelihood Ratio | | | Prevalence | | |
|-----------------|------------|-------------|------|---------|-------------|------|---------|---------------------------|------|---------|---------------------------|------|---------|------------|------|---------|
| | | Minimum | 50% | Maximum | Minimum | 50% | Maximum | Minimum | 50% | Maximum | Minimum | 50% | Maximum | Minimum | 50% | Maximum |
| ASCUS/CIN-I | | | | | | | | | | | | | | | | |
| Overall | 37 | 0.18 | 0.74 | 0.98 | 0.17 | 0.68 | 0.99 | 0.96 | 1.93 | 81.9 | 0.08 | 0.47 | 1.18 | 0.02 | 0.51 | 0.94 |
| Verification | | | | | | | | | | | | | | | | |
| All or random | 21 | 0.31 | 0.68 | 0.92 | 0.17 | 0.75 | 0.99 | 0.96 | 2.38 | 81.9 | 0.11 | 0.50 | 1.18 | 0.02 | 0.36 | 0.91 |
| Some or unclear | 16 | 0.18 | 0.78 | 0.98 | 0.20 | 0.60 | 0.92 | 1.11 | 1.70 | 3.20 | 0.08 | 0.40 | 0.90 | 0.20 | 0.73 | 0.94 |
| LSIL/CIN-I | | | | | | | | | | | | | | | | |
| Overall | 71 | 0.17 | 0.69 | 0.99 | 0.09 | 0.81 | 1.0 | 1.0 | 2.90 | 1301.8 | 0.03 | 0.44 | 0.99 | 0.02 | 0.64 | 0.95 |
| Verification | | | | | | | | | | | | | | | | |
| All or random | 38 | 0.18 | 0.62 | 0.98 | 0.09 | 0.90 | 1.0 | 1.08 | 4.63 | 1301.8 | 0.13 | 0.49 | 0.86 | 0.02 | 0.43 | 0.94 |
| Some or unclear | 33 | 0.17 | 0.75 | 0.99 | 0.18 | 0.71 | 0.95 | 1.0 | 1.80 | 8.60 | 0.03 | 0.39 | 0.99 | 0.33 | 0.72 | 0.95 |
| LSIL/CIN-II-III | | | | | | | | | | | | | | | | |
| Overall | 54 | 0.23 | 0.83 | 1.0 | 0.06 | 0.66 | 0.99 | 0.78 | 1.92 | 52.6 | 0.01 | 0.32 | 2.76 | 0.01 | 0.32 | 0.91 |
| Verification | | | | | | | | | | | | | | | | |
| All or random | 31 | 0.23 | 0.81 | 0.99 | 0.06 | 0.77 | 0.99 | 0.78 | 2.27 | 52.6 | 0.01 | 0.37 | 2.76 | 0.01 | 0.24 | 0.91 |
| Some or unclear | 23 | 0.44 | 0.87 | 1.0 | 0.08 | 0.46 | 0.97 | 1.07 | 1.64 | 18.9 | 0.01 | 0.30 | 0.79 | 0.02 | 0.47 | 0.91 |
| HSIL/CIN-II-III | | | | | | | | | | | | | | | | |
| Overall | 43 | 0.06 | 0.58 | 1.0 | 0.21 | 0.92 | 1.0 | 1.27 | 4.46 | 289.0 | 0.01 | 0.52 | 0.95 | 0.02 | 0.41 | 0.91 |
| Verification | | | | | | | | | | | | | | | | |
| All or random | 25 | 0.18 | 0.53 | 0.92 | 0.64 | 0.96 | 1.0 | 2.10 | 9.95 | 289.0 | 0.12 | 0.53 | 0.84 | 0.02 | 0.28 | 0.91 |
| Some or unclear | 18 | 0.06 | 0.62 | 1.0 | 0.21 | 0.78 | 0.99 | 1.27 | 2.85 | 15.4 | 0.01 | 0.50 | 0.95 | 0.17 | 0.54 | 0.91 |

* ASCUS = atypical squamous cells of undetermined significance; CIN-I-III = cervical intraepithelial neoplasia, grades I-III. LSIL = low-grade squamous intraepithelial lesion; HSIL = high-grade squamous intraepithelial lesion.

Table 4. Studies of the Conventional Papanicolaou Test in Screening Samples without Verification Bias*

| Study (Reference) | Year | Patients and Methods | Study Design and Characteristics | Location and Time Period | Outcomes |
|---|------|--|--|---|--|
| University of Zimbabwe/JHPIEGO Cervical Cancer Project (115)† | 1999 | Consecutive women undergoing primary screening. Phase I included 8731 women (verification: some or unclear); phase II included 2203 women, but only 2147 had complete data (all verified). Mean age, 33 y. | Diagnostic test evaluation of visual inspection with acetic acid and Pap smear | Zimbabwe, 10/95–8/97 | Pap smear reported as \geq LSIL; reference standard: histology or negative colposcopy, reported as \geq LSIL, HSIL |
| Davison and Marty (59) | 1994 | 200 consecutive nonpregnant premenopausal women undergoing screening (196 had complete data). Women with previous abnormal smear were excluded. All included patients were verified. | Diagnostic test evaluation of screening colposcopy and Pap smear | United States, years not specified | Pap smear reported as \geq mild dysplasia; reference standard: histology or negative colposcopy, reported as \geq CIN-I |
| Giles et al. (79) | 1988 | 200 predominantly middle-class women presenting for screening (24 smears were unsatisfactory). Women with previous cervical abnormalities were excluded. Mean age, 39 y. All included patients were verified. | Diagnostic test evaluation of screening colposcopy and Pap smear | United Kingdom, years not specified | Pap smear reported as \geq mild dyskaryosis; reference standard: histology or negative colposcopy, reported as \geq CIN-I (koilocytosis regarded as negative) |
| Baldauf et al. (70)‡ | 1995 | 1539 consecutive women undergoing routine prenatal or gynecologic examination. Mean age, 36 y. 10% random verification of test-negative patients. | Diagnostic test evaluation of cervicography and Pap smear | France, 1/91–12/92 | Pap smear reported as \geq atypical cells; reference standard: histology or negative colposcopy, reported as \geq CIN-I |
| Guerra et al. (113) | 1998 | 3658 consecutive pregnant women presenting for prenatal care. Mean age, 29 y. All patients were verified. | Diagnostic test evaluation of screening colposcopy and Pap smear | Italy, 2/92–3/93 | Pap smear reported as \geq ASCUS, LSIL, HSIL, or microinvasion; reference standard: colposcopy, reported as \geq abnormal transformation zone with minor and major changes |
| Hockstad (47) | 1992 | 73 consecutive women presenting without previous abnormal Pap smear, pelvic symptoms, or hysterectomy (2 declined and 1 did not follow-up). Age range, 15–39 y. All patients were verified. | Diagnostic test evaluation of Pap smear, HPV test, and colposcopy | United States, 1988–1989 | Pap smear reported as \geq atypical or condylomatous changes; reference standard: histology or negative colposcopy, reported as \geq mild dysplasia. If women had negative screening colposcopy but positive Pap smear or HPV test, second colposcopy with biopsy was performed. |
| Garutti et al. (103) | 1994 | 200 nonpregnant women referred to clinic for screening. Mean age, 41 y. All included patients were verified. | Diagnostic test evaluation of Pap smear and cervicography | Italy, 1/90–9/90 | Pap smear reported as \geq HPV; reference standard: histology or negative colposcopy reported as \geq HPV |
| Loiudice et al. (108) | 1998 | 3342 consecutive nonpregnant women undergoing primary screening (42 protocol violations excluded). Women who were menstruating; had previous abnormal Pap smears, HPV, or HIV; or were immunocompromised were excluded. Mean age, 33 y. All included patients were verified. | Diagnostic test evaluation of Pap smear and speculoscopy | Italy, 2/95–11/95 | Pap smear reported as \geq LSIL; reference standard: histology or negative colposcopy, reported as \geq LSIL, high-grade |
| Kesic et al. (117) | 1993 | 418 asymptomatic women referred for screening (23 had defective cervicograms and were excluded). All patients were verified. | Diagnostic test evaluation of Papanicolaou smear and cervicography | Yugoslavia, 1/88–8/89 | Pap smear reported as \geq class III; reference standard: histology or negative colposcopy, reported as \geq CIN-I, CIN-II–III |
| Londhe et al. (110) | 1997 | 500 consecutive sexually active nonpregnant women presenting to gynecology clinic (only 372 were included in the study). All included patients were verified. | Diagnostic test evaluation of Pap smear and visual inspection with acetic acid | India, years not specified | Pap smear reported as \geq positive; reference standard: histology or negative colposcopy, reported as \geq LSIL, HSIL |
| Mannino (114) | 1998 | 3049 consecutive women presenting for screening. Women with previous abnormal Pap smears, vaginitis, or menopause were excluded. All patients were verified. | Diagnostic test evaluation of Pap smear and screening colposcopy | United States, unspecified 10-year period | Pap smear reported as \geq LSIL; reference standard: histology or negative colposcopy, reported as \geq koilocytosis/CIN-I, CIN-II–III |
| Mann et al. (71) | 1993 | 243 women presenting for screening. Women with previous cervical or vaginal pathologic characteristics were excluded. Mean age, 29 y. All patients were verified. | Diagnostic test evaluation of Pap smear and speculoscopy | United States, 1989–1992 | Pap smear reported as \geq atypia with condylomatous features; reference standard: histology or negative colposcopy, reported as \geq HPV/CIN-I, CIN-II–III |

* ASCUS = atypical squamous cells of undetermined significance; CIN-I–III = cervical intraepithelial neoplasia, grades I–III; HPV = human papillomavirus; HSIL = high-grade squamous intraepithelial lesion; JHPIEGO = Johns Hopkins Program for International Education in Gynecology and Obstetrics; LSIL = low-grade squamous intraepithelial lesion; NS = not specified; Pap = Papanicolaou.

† Results of this study are given for phase II only.

‡ Results of this study were estimated from 10% of the patients whose negative test results were verified.

Table 4—Continued

| Prevalence | Study Results | | Quality Evaluation | | | | | | | |
|---|---|--|--------------------|--------------------|---------|--------------|-------------|----------|-------------|----------|
| | Sensitivity | Specificity | Quality Score | Reference Standard | Blinded | Verification | Consecutive | Spectrum | Publication | Industry |
| CIN-I: 487/2092 (0.23) CIN-II-III: 201/2092 (0.10) | LSIL/CIN-I: 0.30 LSIL/CIN-II-III: 0.44 | LSIL/CIN-I: 0.92 LSIL/CIN-I: 0.91 | 10 | 1 | 2 | 2 | 2 | 1 | 1 | 1 |
| CIN-I: 30/196 (0.15) | LSIL/CIN-I: 0.53 | LSIL/CIN-I: 1.0 | 10 | 1 | 2 | 2 | 2 | 1 | 1 | 1 |
| CIN-I: 17/176 (0.10) | LSIL/CIN-I: 0.58 | LSIL/CIN-I: 0.95 | 10 | 1 | 2 | 2 | 2 | 1 | 1 | 1 |
| CIN-I: 62/1343 (0.05) | ASCUS/CIN-I: 0.56 | ASCUS/CIN-I: 0.98 | 9 | 1 | 2 | 1 | 2 | 1 | 1 | 1 |
| CIN-I: 72/3658 (0.02) CIN-II-III: 55/3658 (0.02) | ASCUS/CIN-I: 0.90 LSIL/CIN-I: 0.87 LSIL/CIN-II-III: 0.96 HSIL/CIN-II-III: 0.88 | ASCUS/CIN-I: 0.97 LSIL/CIN-I: 0.98 LSIL/CIN-II-III: 0.98 HSIL/CIN-II-III: 1.0 | 8 | 1 | 0 | 2 | 2 | 1 | 1 | 1 |
| CIN-I: 7/70 (0.10) | ASCUS/CIN-I: 0.31 | ASCUS/CIN-I: 0.96 | 8 | 1 | 0 | 2 | 2 | 1 | 1 | 1 |
| CIN-I: 72/200 (0.36) | LSIL/CIN-I: 0.42 | LSIL/CIN-I: 0.86 | 8 | 2 | 2 | 2 | 0 | 0 | 1 | 1 |
| CIN-I: 267/3300 (0.08) CIN-II-III: 25/3300 (0.01) | LSIL/CIN-I: 0.40 LSIL/CIN-II-III: 0.75 | LSIL/CIN-I: 0.96 LSIL/CIN-II-III: 0.93 | 8 | 1 | 0 | 2 | 2 | 1 | 1 | 1 |
| CIN-I: 27/395 (0.07) CIN-II-III: 19/395 (0.05) | LSIL/CIN-I: 0.52 LSIL/CIN-II-III: 0.53 | LSIL/CIN-I: 0.94 LSIL/CIN-II-III: 0.93 | 7 | 1 | 2 | 2 | 0 | 0 | 1 | 1 |
| CIN-I: 98/372 (0.26) CIN-II-III: 23/372 (0.06) | NS/CIN-I: 0.14 NS/CIN-II-III: 0.23 | NS/CIN-I: 0.96 NS/CIN-II-III: 0.95 | 7 | 1 | 0 | 2 | 2 | 0 | 1 | 1 |
| CIN-I: 904/3049 (0.30) CIN-II-III: 60/3049 (0.02) | LSIL/CIN-I: 0.30 LSIL/CIN-II-III: 0.99 | LSIL/CIN-I: 1.0 LSIL/CIN-II-III: 0.93 | 7 | 1 | 0 | 2 | 2 | 0 | 1 | 1 |
| CIN-I: 29/243 (0.12) CIN-II-III: 6/243 (0.02) | LSIL/CIN-I: 0.32 LSIL/CIN-II-III: 0.64 | LSIL/CIN-I: 0.99 LSIL/CIN-II-III: 0.97 | 6 | 1 | 0 | 2 | 0 | 1 | 1 | 1 |

standard. Although sensitivity and specificity could not be calculated directly, the performance of ThinPrep and conventional Pap smears could be estimated and compared. In the study by Roberts and colleagues (26), any positive result on either test was verified cytologically or histologically; histologic verification was obtained on a majority of HSIL samples. The relative true-positive rate was 1.13, indicating that ThinPrep had higher sensitivity, and the relative false-positive rate was 1.12, indicating that ThinPrep had slightly lower specificity (26). In the study by Hutchinson and coworkers (27), final reference diagnoses were made by using a combination of cytologic and histologic examination, but histologic verification was obtained for more than 90% of Pap smears that showed HSIL or cancer. The relative true-positive rate was 1.19, indicating higher sensitivity for ThinPrep, and the relative false-positive rate was 2.05, indicating lower specificity for ThinPrep (27).

Only 94 studies of the conventional Pap test met the inclusion criteria (Tables 2 and 3) (28–121). Sample sizes ranged from 9 to 22 412 (median, 202). Most of these studies were conducted in samples of women who were referred for previous cytologic abnormalities, had visible cervical lesions, or were at high risk for cervical cancer (for example, immunocompromised patients). Few studies evaluated low-prevalence screening samples. Most studies used histologic examination as a reference standard, but only 51% obtained verification of all or a random fraction of patients whose test results were negative. Few studies independently assessed the test and reference standard. Although most studies used adequate sample selection procedures, 15% did not provide adequate information on the spectrum of disease in their sample. All included studies were published in full-length reports; no abstracts identified in the screening process provided enough data to meet the inclusion criteria.

Data on sensitivity and specificity were available at four different combinations of test and reference standard thresholds: ASCUS/CIN-I (37 studies), LSIL/CIN-I (71 studies), LSIL/CIN-II–III (54 studies), and HSIL/CIN-II–III (43 studies). Most studies allowed construction of 2×2 tables using more than one combination of test and reference standard threshold. For ASCUS/CIN-I, sensitivity estimates ranged from 18% to 98% and specificity estimates ranged from 17% to 99% (Table 3). For studies with data at the LSIL/CIN-I threshold, sensitivity ranged from 17% to 99% and specificity ranged from 9% to 100%. As expected, when a higher disease threshold of CIN-II–III was used with the same test threshold, sensitivity was higher and specificity was lower.

Our primary objective was to obtain the best

estimates of Pap test performance that were applicable to a low-prevalence screening sample. However, only 12 studies identified low-risk patients undergoing screening Pap smears and also verified all or a random fraction of patients with negative test results (Table 4). For the 9 studies that provided data at the LSIL/CIN-I threshold, sensitivity ranged from 30% to 87% (mean, 47%) and specificity ranged from 86% to 100% (mean, 95%). In 8 of these 9 studies, sensitivity was less than 60%. For the LSIL/CIN-II–III threshold, sensitivity was higher (range, 44% to 99%) and specificity was lower (range, 91% to 98%).

Discussion and Conclusions

Thin-layer cytology (ThinPrep), the computerized rescreening device (PAPNET), and the algorithmic classifier (AutoPap) have all received regulatory approval from the FDA. However, because of three deficiencies in methods, most studies of these technologies were excluded from our review. First, many studies did not apply the new technology and conventional Pap testing prospectively to the same sample of women. Although this allows comparison of detection rates in separate samples, it does not directly compare results in individual patients; therefore, even relative sensitivity and specificity cannot be calculated. Second, almost all studies of thin-layer cytology or computer screening or rescreening failed to verify the disease status of women who had negative test results on cytologic screening tests. Most studies applied the reference standard (adjudicated cytologic or histologic examination) only to cases in which diagnoses differed between conventional and new methods. Concordant positive and concordant negative test results are assumed to be true-positives and true-negatives but may actually be concordant false results. This study design has a consistent underlying bias that can be expected to overestimate the sensitivity and specificity of the new test (122). When the conventional tests and the new tests are conditionally dependent—that is, when tests may have similar problems with sample collection or interpretation of mild disease—this bias can be substantial. Third, little evidence is available with which to assess the effects of thin-layer cytology or computer screening or rescreening on specificity.

Conventional Pap testing is less efficient at discriminating between women who have disease and those who do not than is generally believed. However, the conventional Pap test is still the only screening test that has definitively been shown to reduce the incidence and mortality rates of cervical cancer. Because cervical cancer is usually a slow-growing disease and many low-grade lesions regress

spontaneously, serial testing with Pap smears is effective. Decision analyses have shown that Pap testing every 3 to 5 years is valuable because abnormalities missed during one screening interval will probably be detected during the next (3, 5). Most women who develop cervical cancer do so because of lack of screening rather than errors in cytodiagnosis.

Many studies of the conventional Pap test are biased. Some studies were conducted in patients from colposcopy clinics who were referred for previous cytologic abnormalities. In many of these studies, colposcopy or histologic examination was compared with the original cytologic results, all of which were abnormal. Therefore, only two cells of the 2×2 table could be filled in because no negative test results were verified. Such studies were excluded from our analysis. Other investigators repeated the Pap smear at the time of colposcopy and compared colposcopy or histologic examination with these repeated smears, some of which were negative. Although this design allowed completion of a 2×2 table, it probably biased the spectrum of disease because the study sample was taken from a population that was referred because of abnormalities on an initial Pap test (123). Women with subsequent normal Pap smears must have had an initial false-positive smear or a false-negative smear on repeated testing. We assumed the former, but if our assumption was untrue, we could have underestimated the accuracy of the test.

Our data included studies with several gold standards: histologic examination with cone biopsy, hysterectomy, or punch biopsy, and colposcopy. These reference standards are themselves subject to inaccuracy, and sensitivity and specificity may be underestimated when a diagnostic test is compared with an imperfect reference standard (124). In this case, the sensitivity and specificity of the test may vary with frequency of disease. In the included studies, the proportion of patients with disease ranged from 0.02 to 0.94.

When disease is preferentially verified among women with abnormalities but not among all women, the study sample is further biased. Many studies performed colposcopy only on women with abnormalities (on the Pap test or another test, such as human papillomavirus testing or cervicography) or those with condyloma or abnormal-appearing cervixes. By selecting such patients for verification of results, a high frequency of histologic abnormalities are included in the study sample. This is an example of verification (workup) bias, which can lead to elevated estimates of sensitivity and lowered estimates of specificity (125). In our analysis, studies that verified some or an unclear proportion of patients who tested negative had higher sensitivity and lower specificity than those that verified all patients

or a random sample (**Table 3**). This finding was consistent across all cytologic and histologic thresholds.

In a previous systematic review of the accuracy of the Pap test, Fahey and colleagues (6) identified 59 studies, many of which did not meet our strict inclusion criteria. We included 58 additional studies, and many were published after the earlier review. However, Fahey and colleagues (6) found that cytologic methods had a mean sensitivity of 58% and a mean specificity of 69% in screening samples, results that are generally consistent with those of our study.

The few studies of Pap screening that were conducted in low-prevalence samples and avoided verification bias provided the best estimates of sensitivity and specificity. Although specificity was high, the sensitivity estimates are much lower than generally believed. Future decision models, cost-effectiveness studies, and health policy decisions should consider these lower sensitivity estimates in their analyses.

From Duke University Medical Center and Durham Veterans Affairs Medical Center, Durham, North Carolina.

Disclaimer: This article is based on an evidence report prepared by Duke University under contract with the Agency for Healthcare Research and Quality (AHRQ contract 290-97-0014). The authors are responsible for its content, including any clinical or treatment recommendations. No statement in this article should be construed as an official position of the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

Acknowledgments: The authors thank Jean Slutsky of the Agency for Healthcare Research and Quality for overall project guidance, Jane Kolimaga for organization and management of the project, Ruth Goslin and Rebecca Gray for assistance with manuscript preparation, and the Advisory Panel and peer reviewers of the Evidence Report for many helpful suggestions throughout the entire process.

Grant Support: Dr. Nanda was a Veterans Affairs Women's Health Fellow at the Durham Veterans Affairs Medical Center while conducting this study.

Requests for Single Reprints: Kavita Nanda, MD, MHS, Clinical Research Department, Family Health International, Box 13950, Research Triangle Park, NC 27709; e-mail, knanda@fhi.org.

Requests To Purchase Bulk Reprints (minimum, 100 copies): Barbara Hudson, Reprints Coordinator; phone, 215-351-2657; e-mail, bhudson@mail.acponline.org.

Current Author Addresses: Dr. Nanda: Clinical Research Department, Family Health International, Box 13950, Research Triangle Park, NC 27709.

Drs. McCrory and Matchar: Center for Clinical Health Policy Research, Duke University Medical Center, First Union Tower, 2200 West Main Street, Suite 230, Durham, NC 27705.

Dr. Myers: Department of Obstetrics and Gynecology, Box 3279 Medical Center, Duke University Medical Center, Durham, NC 27710.

Dr. Bastian: Health Services Research and Development (152), Durham Veterans Affairs Medical Center, 508 Fulton Street, Durham, NC 27705.

Dr. Hasselblad: Department of Community and Family Medicine, Box 3850 Medical Center, Duke University Medical Center, Durham, NC 27710.

Mr. Hickey: School of Medicine, Duke University, Durham, NC 27710.

References

- Landis SH, Murray T, Bolden S, Wingo PA. Cancer statistics, 1999. *CA Cancer J Clin*. 1999;49:8-31.
- Buntinx F, Brouwers M. Relation between sampling device and detection of abnormality in cervical smears: a meta-analysis of randomised and quasi-randomised studies. *BMJ*. 1996;313:1285-90.
- Eddy DM. Screening for cervical cancer. *Ann Intern Med*. 1990;113:214-26.
- Fahs MC, Mandelblatt J, Schechter C, Muller C. Cost effectiveness of cervical cancer screening for the elderly. *Ann Intern Med*. 1992;117:520-7.
- Duke University. Evaluation of Cervical Cytology. Evidence Report/Technology Assessment No. 5. Prepared by Duke University. Rockville, MD: Agency for Health Care Policy and Research; 1999.
- Fahey MT, Irwig L, Macaskill P. Meta-analysis of Pap test accuracy. *Am J Epidemiol*. 1995;141:680-9.
- Brown AD, Garber AM. The cost-effectiveness of three new technologies to enhance Pap testing. Chicago: Blue Cross Blue Shield Association Technical Evaluation Center Assessment Program; 1998.
- Abulafia O, Sherer DM. Automated cervical cytology: meta-analyses of the performance of the AutoPap 300 QC System. *Obstet Gynecol Surv*. 1999;54:469-76.
- Abulafia O, Sherer DM. Automated cervical cytology: meta-analyses of the performance of the PAPNET system. *Obstet Gynecol Surv*. 1999;54:253-64.
- The Bethesda system for reporting cervical/vaginal cytologic diagnoses: revised after the second National Cancer Institute Workshop, April 29-30, 1991. *Acta Cytol*. 1993;37:115-24.
- Richart R. Cervical intraepithelial neoplasia. *Pathol Annu*. 1973;8:301-28.
- Reagan J, Fu Y. Histologic types and prognosis of cancers of the uterine cervix. *Radiation Oncology Biology Physics*. 1979;5:1015-20.
- Nyirjesy I. Atypical or suspicious cervical smears. An aggressive diagnostic approach. *JAMA*. 1972;222:691-3.
- Kurman RJ, Henson DE, Herbst AL, Noller KN, Schiffman MH. Interim guidelines for management of abnormal cervical cytology. The 1992 National Cancer Institute Workshop. *JAMA*. 1994;271:1866-9.
- Bolick DR, Hellman DJ. Laboratory implementation and efficacy assessment of the ThinPrep cervical cancer screening system. *Acta Cytol*. 1998;42:209-13.
- Proposed guidelines for primary screening instruments for gynecologic cytology. Intersociety Working Group for Cytology Technologies. *Acta Cytol*. 1997;41:924-33.
- Proposed guidelines for secondary screening (rescreening) instruments for gynecologic cytology. Intersociety Working Group for Cytology Technologies. *Acta Cytol*. 1998;42:273-6.
- Food and Drug Administration. Points to consider for cervical cytology devices, version 7/25/94. Washington, DC: Food and Drug Administration Center for Devices and Radiological Health; 1994.
- Giard RW, Hermans J. The evaluation and interpretation of cervical cytology: application of the likelihood ratio concept. *Cytopathology*. 1993;4:131-7.
- Schulzer M, Anderson DR, Drance SM. Sensitivity and specificity of a diagnostic test determined by repeated observations in the absence of an external standard. *J Clin Epidemiol*. 1991;44:1167-79.
- Chock C, Irwig L, Berry G, Glasziou P. Comparing dichotomous screening tests when individuals negative on both tests are not verified. *J Clin Epidemiol*. 1997;50:1211-7.
- Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *Am J Epidemiol*. 1995;141:263-72.
- Phelps CE, Hutson A. Estimating diagnostic test accuracy using a "fuzzy gold standard." *Med Decis Making*. 1995;15:44-57.
- Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. *JAMA*. 1995;274:645-51.
- Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC, et al. Guidelines for meta-analyses evaluating diagnostic tests. *Ann Intern Med*. 1994;120:667-76.
- Roberts JM, Gurley AM, Thurloe JK, Bowditch R, Laverty CR. Evaluation of the ThinPrep Pap test as an adjunct to the conventional Pap smear. *Med J Aust*. 1997;167:466-9.
- Hutchinson ML, Zahniser DJ, Sherman ME, Herrero R, Alfaro M, Bratti MC, et al. Utility of liquid-based cytology for cervical carcinoma screening: results of a population-based study conducted in a region of Costa Rica with a high incidence of cervical carcinoma. *Cancer*. 1999;87:48-55.
- Ferenczy A, Robitaille J, Franco E, Arseneau J, Richart RM, Wright TC. Conventional cervical cytologic smears vs. ThinPrep smears. A paired comparison study on cervical cytology. *Acta Cytol*. 1996;40:1136-42.
- Rasbridge SA, Nayagam M. Discordance between cytologic and histologic reports in cervical intraepithelial neoplasia. Results of a one-year audit. *Acta Cytol*. 1995;39:648-53.
- Soost HJ, Lange HJ, Lehmacher W, Ruffing-Kullmann B. The validation of cervical cytology. Sensitivity, specificity and predictive values. *Acta Cytol*. 1991;35:8-14.
- Okagaki T, Zelterman D. Information, discrimination and divergence in cytology. II. Total discrimination as a measure of performance. *Acta Cytol*. 1991;35:25-9.
- Lozowski MS, Mishriki Y, Talebian F, Solitare G. The combined use of cytology and colposcopy in enhancing diagnostic accuracy in preclinical lesions of the uterine cervix. *Acta Cytol*. 1982;26:285-91.
- Davis JR, Hindman WM, Paplanus SH, Trego DC, Wiens JL, Suci TN. Value of duplicate smears in cervical cytology. *Acta Cytol*. 1981;25:533-8.
- Johansen P, Arffmann E, Pallesen G. Evaluation of smears obtained by cervical scraping and an endocervical swab in the diagnosis of neoplastic disease of the uterine cervix. *Acta Obstet Gynecol Scand*. 1979;58:265-70.
- Naslund I, Auer G, Pettersson F, Sjövall K. Evaluation of the pulse wash sampling technique for screening of uterine cervical carcinoma. *Acta Radiol Oncol*. 1986;25:131-6.
- Ritter DB, Kadish AS, Vermund SH, Romney SL, Villari D, Burk RD. Detection of human papillomavirus deoxyribonucleic acid in exfoliated cervicovaginal cells as a predictor of cervical neoplasia in a high-risk population. *Am J Obstet Gynecol*. 1988;159:1517-25.
- Wetrich DW. An analysis of the factors involved in the colposcopic evaluation of 2194 patients with abnormal Papanicolaou smears. *Am J Obstet Gynecol*. 1986;154:1339-49.
- Jones BA, 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;120:523-31.
- Coibion M, Autier P, Vandam P, Delobelle A, Huet F, Hertens D, et al. Is there a role for cervicography in the detection of premalignant lesions of the cervix uteri? *Br J Cancer*. 1994;70:125-8.
- Skehan M, Soutter WP, Lim K, Krausz T, Pryse-Davies J. Reliability of colposcopy and directed punch biopsy. *Br J Obstet Gynaecol*. 1990;97:811-6.
- Giles JA, Deery A, Crow J, Walker P. The accuracy of repeat cytology in women with mildly dyskaryotic smears. *Br J Obstet Gynaecol*. 1989;96:1067-70.
- DiBonito L, Falconieri G, Tomasic G, Colautti I, Bonifacio D, Dudine S. Cervical cytopathology. An evaluation of its accuracy based on cytopathologic comparison. *Cancer*. 1993;72:3002-6.
- Ronco G, Montanari G, Aimone V, Parisio F, Segnan N, Valle A, et al. Estimating the sensitivity of cervical cytology: errors of interpretation and test limitations. *Cytopathology*. 1996;7:151-8.
- Parham DM, Wiredu EK, Hussein KA. The cytological prediction of cervical intraepithelial neoplasia in colposcopically directed biopsies. *Cytopathology*. 1991;2:285-90.
- Baldauf JJ, Dreyfus M, Ritter J, Philippe E. Colposcopy and directed biopsy reliability during pregnancy: a cohort study. *Eur J Obstet Gynecol Reprod Biol*. 1995;62:31-6.
- Kwikkel HJ, Quaak MJ, de With C. Predictive value of the abnormal PAP smear: a retrospective analysis of error rates. *Eur J Obstet Gynecol Reprod Biol*. 1986;21:101-12.
- Hockstad RL. A comparison of simultaneous cervical cytology, HPV testing, and colposcopy. *Fam Pract Res J*. 1992;12:53-60.
- Fung Kee Fung M, Senterman M, Eid P, Faught W, Mikhael NZ, Wong PT. Comparison of Fourier-transform infrared spectroscopic screening of exfoliated cervical cells with standard Papanicolaou screening. *Gynecol Oncol*. 1997;66:10-5.
- Del Priore G, Maag T, Bhattacharya M, Garcia PM, Till M, Lurain JR. The value of cervical cytology in HIV-infected women. *Gynecol Oncol*. 1995;56:395-8.
- Smith JJ, Cowan L, Hemstreet G 3d, Smith ML, Pippitt CH Jr, Doggett R, et al. Automated quantitative fluorescent image analysis of cervical cytology. *Gynecol Oncol*. 1987;28:241-54.
- Pang JC, Hsu C, Kwan-Chau KM. Cytology and colposcopy in the diagnosis of cervical neoplasia. *Gynecol Oncol*. 1977;5:134-41.
- Kealy WF. Correlation of cervical cytodiagnosis and histopathology—an exercise in quality control. *Ir J Med Sci*. 1986;155:381-8.
- Korach J, Menczer J, Lusky A. A comparison between referral and hospital cervical cytology laboratory reports. *Isr J Med Sci*. 1995;31:352-5.
- Bolick DR, Hellman DJ. Laboratory implementation and efficacy assessment of the ThinPrep cervical cancer screening system. *Acta Cytol*. 1998;42:209-13.
- Mayeaux EJ Jr, Harper MB, Abreo F, Pope JB, 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.
- Frisch LE, Milner FH, Ferris DG. Naked-eye inspection of the cervix after acetic acid application may improve the predictive value of negative cytologic screening. *J Fam Pract*. 1994;39:457-60.
- Lederer H, Lambourne A. The results of screening by cervical cytology and of histological examination of gynaecological operation specimens. *J Obstet Gynaecol Br Commonw*. 1973;80:67-71.
- Gonzalez D, Hernandez E, Anderson L, Heller P, Atkinson BF. Clinical significance of a cervical cytologic diagnosis of atypical squamous cells of undetermined significance. Favoring a reactive process or low grade squamous intraepithelial lesion. *J Reprod Med*. 1996;41:719-23.
- Davison JM, Marty JJ. Detecting premalignant cervical lesions. Contribution of screening colposcopy to cytology. *J Reprod Med*. 1994;39:388-92.
- Ramirez EJ, Hernandez E, Miyazawa K. Cervical conization findings in women with dysplastic cervical cytology and normal colposcopy. *J Reprod Med*. 1990;35:359-61.
- Whelock JB, 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;34:815-7.
- Morrison BW, Erickson ER, Doshi N, Russo JF. The significance of atypical cervical smears. *J Reprod Med*. 1988;33:809-12.
- Fahim HI, Faris R, Arab SE, Sammour MB. An epidemiologic study of Papanicolaou smear data at Ain-Shams University hospitals. *J Egypt Public Health Assoc*. 1991;66:97-111.
- Korn AP, Autry M, DeRemer PA, Tan W. Sensitivity of the Papanicolaou smear in human immunodeficiency virus-infected women. *Obstet Gynecol*. 1994;83:401-4.
- German M, Heaton R, Erickson D, Henry M, Nash J, O'Connor D. A comparison of the three most common Papanicolaou smear collection techniques. *Obstet Gynecol*. 1994;84:168-73.

66. Koonings PP, Dickinson K, d'Ablaing G 3d, Schlaerth JB. A randomized clinical trial comparing the Cytobrush and cotton swab for Papanicolaou smears. *Obstet Gynecol.* 1992;80:241-5.
67. Andrews S, Hernandez E, Miyazawa K. Paired Papanicolaou smears in the evaluation of atypical squamous cells. *Obstet Gynecol.* 1989;73:747-50.
68. Wright TC, Sun XW, Koulos J. Comparison of management algorithms for the evaluation of women with low-grade cytologic abnormalities. *Obstet Gynecol.* 1995;85:202-10.
69. Spitzer M, Krumholz BA, Chernys AE, Seltzer V, Lightman AR. Comparative utility of repeat Papanicolaou smears, cervicography, and colposcopy in the evaluation of atypical Papanicolaou smears. *Obstet Gynecol.* 1987;69:731-5.
70. Baldauf JJ, Dreyfus M, Lehmann M, Ritter J, Philippe E. Cervical cancer screening with cervicography and cytology. *Eur J Obstet Gynecol Reprod Biol.* 1995;58:33-9.
71. Mann W, Lonky N, Massad S, Scotti R, Blanco J, Vasilev S. Papanicolaou smear screening augmented by a magnified chemiluminescent exam. *Int J Gynaecol Obstet.* 1993;43:289-96.
72. Wright TC Jr, Ellerbrock TV, Chiasson MA, Van Devanter N, Sun XW. Cervical intraepithelial neoplasia in women infected with human immunodeficiency virus: prevalence, risk factors, and validity of Papanicolaou smears. *New York Cervical Disease Study.* *Obstet Gynecol.* 1994;84:591-7.
73. Cox JT, Schiffman MH, Winzelberg AJ, Patterson JM. An evaluation of human papillomavirus testing as part of referral to colposcopy clinics. *Obstet Gynecol.* 1992;80:389-95.
74. Tawa K, Forsythe A, Cove JK, Saltz A, Peters HW, Watring WG. A comparison of the Papanicolaou smear and the cervigram: sensitivity, specificity, and cost analysis. *Obstet Gynecol.* 1988;71:229-35.
75. Cecchini S, Bonardi R, Mazzotta A, Grazzini G, Iossa A, Ciatto S. Testing cervicography and cervicoscopy as screening tests for cervical cancer. *Tumori.* 1993;79:22-5.
76. Hellberg D, Axelsson O, Gad A, Nilsson S. Conservative management of the abnormal smear during pregnancy. A long-term follow-up. *Acta Obstet Gynecol Scand.* 1987;66:195-9.
77. Cox JT, Lorincz AT, Schiffman MH, Sherman ME, Cullen A, Kurman RJ. Human papillomavirus testing by hybrid capture appears to be useful in triaging women with a cytologic diagnosis of atypical squamous cells of undetermined significance. *Am J Obstet Gynecol.* 1995;172:946-54.
78. Haddad NG, Hussein I, Livingstone JR, Smart GE. Colposcopy in teenagers [Letter]. *BMJ.* 1988;297:1128.
79. Giles JA, Hudson E, Crow J, Williams D, Walker P. Colposcopic assessment of the accuracy of cervical cytology screening. *Br Med J (Clin Res Ed).* 1988;296:1099-102.
80. Chomet J. Screening for cervical cancer: a new scope for general practitioners? Results of the first year of colposcopy in general practice. *Br Med J (Clin Res Ed).* 1987;294:1326-8.
81. Gundersen JH, Schauberg CW, Rowe NR. The Papanicolaou smear and the cervigram. A preliminary report. *J Reprod Med.* 1988;33:46-8.
82. Jones MH, Jenkins D, Cuzick J, Wolfendale MR, Jones JJ, Balogun-Lynch C, et al. Mild cervical dyskeratosis: safety of cytological surveillance. *Lancet.* 1992;339:1440-3.
83. Bigrigg MA, Codling BW, Pearson P, Read MD, Swingler GR. Colposcopic diagnosis and treatment of cervical dysplasia at a single clinic visit. Experience of low-voltage diathermy loop in 1000 patients. *Lancet.* 1990;336:229-31.
84. Singh P, Pang M, Ratnam SS. Selective cervical conization following colposcopy: a critical evaluation of 107 cases. *Singapore Med J.* 1985;26:289-94.
85. Syrjanen KJ, Mantjarvi R, Vayrynen M, Yliskoski M, Syrjanen SM, Saarikoski S, et al. Cervical smears in assessment of the natural history of human papillomavirus infections in prospectively followed women. *Acta Cytol.* 1987;31:855-65.
86. Morrison EA, Goldberg GL, Hagan RJ, Kadish AS, Burk RD. Self-administered home cervicovaginal lavage: a novel tool for the clinical-epidemiologic investigation of genital human papillomavirus infections. *Am J Obstet Gynecol.* 1992;167:104-7.
87. Jones DE, Creasman WT, Dombroski RA, Lentz SS, Waeltz JL. Evaluation of the atypical Pap smear. *Am J Obstet Gynecol.* 1987;157:544-9.
88. Stafil A. Cervicography: a new method for cervical cancer detection. *Am J Obstet Gynecol.* 1981;139:815-25.
89. Tay SK, Jenkins D, Singer A. Management of squamous atypia (borderline nuclear abnormalities): repeat cytology or colposcopy? *Aust N Z J Obstet Gynaecol.* 1987;27:140-1.
90. Soutter WP, Wisdom S, Brough AK, Monaghan JM. Should patients with mild atypia in a cervical smear be referred for colposcopy? *Br J Obstet Gynaecol.* 1986;93:70-4.
91. Tait IA, Alawattagama AB, Rees E. Screening for cervical dysplasia in department of genitourinary medicine. *Genitourin Med.* 1988;64:255-8.
92. Maggi R, Zannoni E, Giorda G, Biraghi P, Sideri M. Comparison of repeat smear, colposcopy, and colposcopically directed biopsy in the evaluation of the mildly abnormal smear. *Gynecol Oncol.* 1989;35:294-6.
93. Oyer R, Hanjani P. Endocervical curettage: does it contribute to the management of patients with abnormal cervical cytology? *Gynecol Oncol.* 1986;25:204-11.
94. Upadhyay SN, Jha RS, Sinha TK, Mishra NK. Accuracy of cytology in screening for cervical cancer. *Indian J Med Res.* 1984;80:457-62.
95. Walker EM, Dodgson J, Duncan ID. Does mild atypia on a cervical smear warrant further investigation? *Lancet.* 1986;2:672-3.
96. Parker A, Ueki M. A comparison of preoperative exfoliative cervical cytology with subsequent histology. *N Z Med J.* 1986;99:414-6.
97. Sheets EE, Constantine NM, Dinisco S, Dean B, Cibas ES. Colposcopically directed biopsies provide a basis for comparing the accuracy of Thin-Prep and Papanicolaou smears. *J Gynecol Tech.* 1995;1:27-34.
98. Kaufman RH, Adam E, Icenogle J, Reeves WC. Human papillomavirus testing as triage for atypical squamous cells of undetermined significance and low-grade squamous intraepithelial lesions: sensitivity, specificity, and cost-effectiveness. *Am J Obstet Gynecol.* 1997;177:930-6.
99. Spinillo A, Capuzzo E, Tenti P, De Santolo A, Piazzi G, Iaschi A. Adequacy of screening cervical cytology among human immunodeficiency virus-seropositive women. *Gynecol Oncol.* 1998;69:109-13.
100. Herrington CS, Evans MF, Hallam NF, Charnock FM, Gray W, McGee JD. Human papillomavirus status in the prediction of high-grade cervical intraepithelial neoplasia in patients with persistent low-grade cervical cytological abnormalities. *Br J Cancer.* 1995;71:206-9.
101. Ferris DG, Wright TC, Litaker MS, Richart RM, Lorincz AT, Sun XW, et al. Triage of women with ASCUS and LSIL on Pap smear reports: management by repeat Pap smear, HPV DNA testing, or colposcopy? *J Fam Pract.* 1998;46:125-34.
102. Glenthøj A, Rank F, Peen U, Bostofte E. Diagnostic efficiency of brush cytology from the uterine cervix. In: Goerttler K, Feichter G, Witte S, eds. *New Frontiers in Cytology: Modern Aspects of Research and Practice.* Berlin, NY: Springer-Verlag; 1988:421-4.
103. Garutti P, Immovilli S, Chiari U, Segala V, Mollica G. Cervical screening for detection of cervical HPV. *The Cervix and the Lower Female Genital Tract.* 1994;12:31-4.
104. Slawson DC, Bennett JH, Simon LJ, Herman JM. Should all women with cervical atypia be referred for colposcopy: a HARNET study. *Harrisburg Area Research Network.* *J Fam Pract.* 1994;38:387-92.
105. Beeby AR, Keating PJ, Wagstaff TI, Wilson EMJ, Sims PF, Manning PJ, et al. The quality and accuracy of cervical cytology: a study of five sampling devices in a colposcopy clinic. *Journal of Obstetrics and Gynaecology.* 1993;13:276-81.
106. Young NA, Naryshkin S, Bowman RL. Value of repeat cervical smears at the time of colposcopic biopsy. *Diagn Cytopathol.* 1993;9:646-9.
107. Sankaranarayanan R, Wesley R, Somanathan T, Dhakad N, Shyamalakumary B, Amma NS, et al. Visual inspection of the uterine cervix after the application of acetic acid in the detection of cervical carcinoma and its precursors. *Cancer.* 1998;83:2150-6.
108. Louidice L, Abbiati R, Boselli F, Cecchini G, Costa S, Grossi E, et al. Improvement of Pap smear sensitivity using a visual adjunctive procedure: a co-operative Italian study on speculscopy (GISPE). *Eur J Cancer Prev.* 1998;7:295-304.
109. Recio FO, Sahai Srivastava BI, Wong C, Hempling RE, Eltabbakh GH, Piver MS. The clinical value of digene hybrid capture HPV DNA testing in a referral-based population with abnormal pap smears. *Eur J Gynaecol Oncol.* 1998;19:203-8.
110. Londhe M, George SS, Seshadri L. Detection of CIN by naked eye visualization after application of acetic acid. *Indian J Cancer.* 1997;34:88-91.
111. Manos MM, Kinney WK, Hurley LB, Sherman ME, Shieh-Ngai J, Kurman RJ, et al. Identifying women with cervical neoplasia: using human papillomavirus DNA testing for equivocal Papanicolaou results. *JAMA.* 1999;281:1605-10.
112. Shulman LP, Klett CC, Phillips OP. Preliminary comparison of papanicolaou smear and speculscopy for assessing cervical cancer precursor lesions in pregnant adolescents. *J Pediatr Adolesc Gynecol.* 1999;12:78-82.
113. Guerra B, De Simone P, Gabrielli S, Falco P, Montanari G, Bovicelli L. Combined cytology and colposcopy to screen for cervical cancer in pregnancy. *J Reprod Med.* 1998;43:647-53.
114. Mannino JR. Natural history of false-negative papanicolaou smears: a prospective study using screening colposcopy in addition to cytology. *J Am Osteopath Assoc.* 1998;98:542-6.
115. Visual inspection with acetic acid for cervical-cancer screening: test qualities in a primary-care setting. University of Zimbabwe/HPIEGO Cervical Cancer Project. *Lancet.* 1999;353:869-73.
116. Eskridge C, Begnaud WP, Landwehr C. Cervicography combined with repeat Papanicolaou test as triage for low-grade cytologic abnormalities. *Obstet Gynecol.* 1998;92:351-5.
117. Kesic VI, Soutter WP, Sulovic V, Juznic N, Aleksic M, Ljubic A. A comparison of cytology and cervicography in cervical screening. *International Journal of Gynecological Cancer.* 1993;3:395-8.
118. Cruickshank ME, Buchan S, Melvin WT, Kitchener HC. Human papillomavirus type 16 and 18 detection in the management of mild dyskaryosis. *Br J Obstet Gynaecol.* 1999;106:969-76.
119. Carta G, Di Stefano L, Catellani Perelli A, Toro G, Moscarini M. Colposcopy, cytology and histology in the diagnosis of squamous intraepithelial lesions of the cervix. *Clin Exp Obstet Gynecol.* 1999;26:60-6.
120. Giraud J, Coiffic J, Poulain P, Kerisit J. High prevalence of cervical intra-epithelial neoplasia in women treated for pelvic inflammatory disease. *Eur J Obstet Gynecol Reprod Biol.* 1998;81:51-4.
121. Maiman M, Fruchter RG, Sedlis A, Feldman J, Chen P, Burk RD, et al. Prevalence, risk factors, and accuracy of cytologic screening for cervical intraepithelial neoplasia in women with the human immunodeficiency virus. *Gynecol Oncol.* 1998;68:233-9.
122. Miller WC. Bias in discrepant analysis: when two wrongs don't make a right. *J Clin Epidemiol.* 1998;51:219-31.
123. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med.* 1978;299:926-30.
124. Boyko EJ. Re: "Meta-analysis of Pap test accuracy" [Letter]. *Am J Epidemiol.* 1996;143:406-7.
125. Feinstein AR. *Clinical Epidemiology: The Architecture of Clinical Research.* Philadelphia: WB Saunders; 1985:434.