

Head-to-Head Evaluation of Five Chlamydia Tests Relative to a Quality-Assured Culture Standard

WILBERT J. NEWHALL,^{1,2} ROBERT E. JOHNSON,^{1*} SUSAN DELISLE,^{1,3} DAVID FINE,³
ALULA HADGU,¹ BESSIE MATSUDA,^{4†} DONNA OSMOND,⁵ JOYCE CAMPBELL,⁵
AND WALTER E. STAMM⁶

*Centers for Disease Control and Prevention, Atlanta, Georgia*¹; *JSI Research and Training Institute, Denver, Colorado*²; *The Center for Health Training, Washington State Health Department Laboratory,*³ *and University of Washington,*⁴ *Seattle, Washington;* *and Oregon State Health Department Laboratory, Portland, Oregon*⁴

Received 23 July 1998/Returned for modification 15 October 1998/Accepted 15 December 1998

Nucleic acid amplification tests offer superior sensitivity for the detection of *Chlamydia trachomatis* infection, but many laboratories still use nonamplification methods because of the lower cost and ease of use. In spite of their availability for more than a decade, few studies have directly compared the nonamplification tests. Such comparisons are still needed in addition to studies that directly compare individual nonamplification and amplification tests. The purpose of this study was to evaluate and compare the performance characteristics relative to culture of five different tests for the detection of *C. trachomatis* with and without confirmation of positive results. The tests were applied to endocervical specimens from 4,980 women attending family planning clinics in the northwestern United States. The five nonculture tests included Chlamydiazyme (Abbott), MicroTrak direct fluorescent antibody (DFA) (Syva), MicroTrak enzyme immunoassay (EIA) (Syva), Pace 2 (Gen-Probe), and Pathfinder EIA (Sanofi/Kallestad). All positive results obtained with a nonculture test (except MicroTrak DFA) were confirmed by testing the original specimens with a blocking antibody test (Chlamydiazyme), a cytospin DFA (MicroTrak EIA and Pathfinder EIA), and a probe competition assay (Pace 2). The prevalence of culture-proven chlamydia was 3.9%. The sensitivities of the nonculture tests were in a range from 62 to 75%, and significant differences between tests in terms of sensitivity were observed. The positive predictive value for each test was 0.85 or higher. The specificities of the nonculture tests without performance of confirmations were greater than 99%. Performing confirmatory tests eliminated nearly all of the false positives.

One of the key decisions that must be made when implementing a control program for chlamydia is which laboratory test to use. The standard for over a decade has been the isolation of the organism in cell culture. Cell culture is the standard because it is nearly 100% specific and has been traditionally also more sensitive than most nonculture methods. The high specificity of culture results from the visualization of inclusions in tissue culture cells by a fluorescent antibody that is specific for *Chlamydia trachomatis*. The sensitivity of culture, however, is not ideal and ranges widely from laboratory to laboratory. Estimates of culture sensitivity vary from as low as 50 to 90% (24). Nevertheless, with the exception of nucleic acid amplification methods such as ligase chain reaction, PCR, and transcription-mediated amplification, commercial nonculture methods have proven to be less sensitive than high-quality cultures in head-to-head comparisons (2). They also may give false-positive results that, in many clinical settings, may cause psychosocial harm to the patient. Unfortunately, cell culture is beyond the capabilities of most public and private laboratories due to its technical demands, labor intensity, and high cost. Amplification tests potentially have superior sensitivity but are new, somewhat more demanding technically, and more expensive. As a result, nonculture, non-nucleic acid amplification

methods are routinely used by many clinicians and health departments to detect chlamydia in genital specimens as a substitute for culture.

The reported sensitivities of nonculture tests for which an extensive evaluation literature exists (Chlamydiazyme and MicroTrak direct fluorescent antibody [DFA] tests) vary greatly (approximately 60 to 90%) (2). This between-study variation is much greater than what can be accounted for by sampling error. Important sources of this variation include variability in the sensitivity of the cell culture methods used (the reference standard in most published studies), variation in the specimen collection techniques used by clinicians, variation in transport time and methodology, and the small number of culture-positive cases analyzed in most studies (1, 11, 17, 19, 20, 24, 25, 28). Whatever the source, this large variation makes it virtually impossible to compare performances of different tests if they have not been evaluated in the same study. Therefore, evaluations which compare the candidate tests of interest simultaneously with culture to control for the large variation in sensitivity due to factors other than the performance of candidate tests are needed. Only a few such comparative evaluations of nonamplification methods have been reported in the last 5 years (3, 7, 11, 30). Expanding the number of studies that compare multiple tests simultaneously and that include a more substantial number of reference test-positive specimens than were studied in most earlier evaluations will be a principal recommendation of new guidelines for chlamydia laboratory tests now in development at the Centers for Disease Control and Prevention (CDC).

* Corresponding author. Mailing address: Division of STD Prevention, Centers for Disease Control and Prevention, 1600 Clifton Rd., Mailstop E-02, Atlanta, GA 30333. Phone: (404) 639-1894. Fax: (404) 639-8610. E-mail: rej1@cdc.gov.

† Retired.

Additionally, in most evaluations of nonculture tests, there has been little attempt to evaluate procedures for confirming positive test results. In the practical application of these tests, a specificity less than 100% can have significant consequences, particularly in populations with a low prevalence of chlamydial infection. For example, a nonculture test with a specificity of 98% and a sensitivity of 80% has a positive predictive value (PPV) that decreases from 87.6 to 45.1% as the prevalence decreases from 15 to 2%. Specificity and the PPV can be increased by performing a confirmatory or supplemental test for those patients who have a positive screening test. The CDC recommends confirmatory testing of all specimens giving positive nonculture screening test results in populations with a chlamydia prevalence less than 5.0% (5). Although several approaches to confirmatory testing have been developed and are commercially available, evaluations have been published only for the Chlamydiazyme test (using a blocking antibody) (13, 19, 20, 26, 27) and for the Syva MicroTrak enzyme immunoassay (EIA) using DFA (6, 26) and PCR (21).

In summary, evaluations of nonculture tests for chlamydia which (i) compare multiple candidate tests simultaneously with a quality-assured culture standard, (ii) compare the effectiveness of methods for confirming positive nonculture test results, (iii) employ valid statistical analyses for comparing the performance of different tests, and (iv) include sufficient numbers of reference test-positive persons to allow adequate precision in estimating test sensitivity are needed. Such evaluations are also needed to test the validity of CDC chlamydia prevention recommendations for confirmatory testing and to enable health departments and health care providers to select the most accurate screening and confirmatory tests for their needs. Once the better-performing nonamplification tests are identified, more head-to-head studies comparing these tests with the new amplification tests should be conducted.

Following are the results of a clinical trial that was designed to provide information regarding the relative accuracy of nonculture tests for detecting chlamydia in cervical specimens from women. Of interest was both the performance of the initial screening test and the performance of confirmatory tests done when the screening test was positive. The tests that were selected for comparison with standard culture techniques included Syva's MicroTrak EIA, Abbott's Chlamydiazyme EIA, Sanofi's EIA, Gen-Probe's Pace 2 nucleic acid hybridization test, and Syva's MicroTrak DFA test. These tests were selected because they are (i) older tests that have been widely used and extensively evaluated and thereby serve as historical standards (Chlamydiazyme and MicroTrak DFA) or (ii) extensively marketed tests that have received relatively limited comparative evaluation (Sanofi EIA, Gen-Probe, and Syva EIA). Commercial nucleic acid amplification tests were not yet available at the time of this study.

MATERIALS AND METHODS

Patient population. Female patients attending participating family planning clinics in the states of Washington and Oregon during 1992 and 1993 were considered for enrollment in the study. The previously published screening criteria of the Region X Chlamydia Project were used to establish eligibility for enrollment (4). These criteria included any of the following: (i) mucopurulent cervicitis, pelvic inflammatory disease, friable cervix, or abnormal bleeding; (ii) a partner with signs and/or symptoms suggestive of urethritis; (iii) client request; (iv) rape within the previous 60 days; (v) candidacy for intrauterine device insertion; and (vi) a positive pregnancy test and a bimanual pelvic examination. Alternatively, the criteria included two or more of the following: (i) age under 24 years and being sexually active; (ii) new sex partner in the previous 60 days; (iii) sex partner with multiple partners in the previous 30 days; (iv) multiple sex partners in the previous 30 days; and (v) use of nonbarrier birth control method or no birth control method (nonbarrier birth control methods include oral

contraceptives, the intrauterine device, sterilization, and all natural family planning methods).

Specimen collection. Specimens for gonorrhea testing and Pap smear were collected before obtaining specimens for chlamydia testing. Chlamydia test specimens were collected after first removing excess mucus from the cervical os and surrounding mucosa with a large cotton swab. Chlamydia test specimens were collected by taking six sequential swabs from the endocervix. The first of these swabs was placed into chlamydia culture transport medium and stored at 4°C for same-day transport (Washington) or at -70°C for biweekly transport (Oregon) to the University of Washington Chlamydia Laboratory. The sequence of specimen collection for the five nonculture tests was randomized. Specimens were collected by using collection kits and procedures as outlined in the package inserts for each of the tests. Clinicians were discouraged from enrolling someone in this study when a Pap smear was obtained with a cytobrush, due to frequent bleeding in such patients. All DFA specimens were obtained with a swab.

Cell culture. Specimens for chlamydia isolation were cultured by using cycloheximide-treated McCoy cells in 96-well microtiter plates as previously described (29). Blind passages were not performed. Chlamydia inclusions were detected by using a fluorescein-labeled monoclonal antibody that binds to a species-specific epitope on the major outer membrane protein.

Nonculture tests. Specimens collected for each of the nonculture tests were transported and processed according to directions provided by the manufacturer in the package inserts.

Confirmatory-supplemental testing. All specimens that gave positive results by the Syva EIA were analyzed by Syva's cytospin confirmatory DFA procedure according to directions provided in the package insert. All specimens that gave positive results by the Sanofi EIA were analyzed by a cytospin DFA confirmation procedure provided by the manufacturer. All specimens that gave positive results by the Gen-Probe assay were analyzed by Gen-Probe's probe confirmation assay according to the directions provided in the package insert. All specimens that gave positive results by the Abbott Chlamydiazyme EIA were analyzed by the blocking antibody procedure according to directions provided in the package insert. Conventional methods for confirming DFA tests (e.g., Syva DFA) with the original specimen are not available.

Data analysis. Sensitivity and specificity estimates were obtained by assuming cell culture as the "gold standard." Ninety-five percent confidence intervals (CIs) were calculated based on the binomial distribution of the observed values. Standard errors for the pairwise comparisons in Table 2 were based on a robust-variance estimation approach (23). *P* values for statistical tests of significance were calculated without adjusting for the 10 paired comparisons of sensitivity among the five tests (Table 2). Using the method of least significant differences, we also ensured an overall significance level of $\alpha = 0.05$ by requiring a *P* value to be less than 0.005 before considering a difference to be statistically significant (Table 2) (16).

RESULTS

Performance of screening tests without confirmatory-supplemental testing. A total of 4,980 clients gave endocervical samples that were tested for chlamydia by cell culture. The prevalence of chlamydia in this population of women as determined by cell culture was 3.9% (194 of 4,980). The majority (98.1%) were also tested by four or five of the nonculture tests. The sensitivities of the nonculture tests were calculated relative to cell culture as the gold standard (Table 1). The results of a pairwise statistical comparison of test sensitivities are given in Table 2. Sensitivities of the five tests ranged from 75.3% (95% CI, 68.6 to 81.2%) for Pace 2 to 61.9% (95% CI, 55.6 to 68.7%) for Chlamydiazyme (Table 1). The sensitivities of the Pace 2, MicroTrak DFA, and MicroTrak EIA tests were highest (71.7 to 75.3%) and did not differ significantly from one another ($P \geq 0.23$). The Chlamydiazyme test had the lowest sensitivity (61.9%), which was significantly less than that of each of the three most sensitive tests ($P \leq 0.003$). The sensitivity of the Sanofi EIA test was intermediate (66.8%) and significantly less than that of either the Pace 2 test ($P = 0.006$) or the MicroTrak DFA test ($P = 0.026$) but not significantly less than that of the MicroTrak EIA test ($P = 0.195$) or significantly greater than that of the Chlamydiazyme test ($P = 0.109$). All of the foregoing significant differences remained significant at $P < 0.05$ after taking into account the multiple tests of significance, except for the difference in sensitivities between the Sanofi EIA and the Gen-Probe and MicroTrak DFA tests (Table 2).

The number of false-positive results for each diagnostic test

TABLE 1. Performance characteristics of nonculture tests relative to culture

| Test | Sensitivity ^a | 95% CI (%) for sensitivity | Specificity ^a | 95% CI (%) for specificity | No. of false positives | PPV ^b |
|------------|--------------------------|-------------------------------|--------------------------|-------------------------------|---------------------------|------------------|
| Gen-Probe | 75.3 (146/194) | 68.6–81.2 | 99.6 (4,694/4,712) | 99.4–99.8 | 18 | 0.890 |
| Syva DFA | 74.5 (140/188) | 67.6–80.5 | 99.8 (4,429/4,438) | 99.6–99.9 | 9 | 0.940 |
| Syva EIA | 71.7 (139/194) | 64.8–77.9 | 99.6 (4,696/4,713) | 99.4–99.8 | 17 | 0.891 |
| Sanofi EIA | 66.8 (129/193) | 59.7–73.4 | 99.6 (4,703/4,720) | 99.4–99.8 | 17 | 0.884 |
| Abbott EIA | 61.9 (120/194) | 55.6–68.7 | 99.6 (4,692/4,713) | 99.3–99.7 | 21 | 0.851 |

^a True positive equals culture positive. Data were generated by using the manufacturer's recommended cutoffs and results prior to confirmation. Values are reported as percents (numbers of specimens positive/total numbers of specimens).

^b Prevalence of 3.9% based on cell culture.

relative to cell culture ranged from 9 to 21. Test specificities were uniformly high, exceeding 99.5% (95% CI, lower bound exceeding 99.3%) for each test (Table 1). A pairwise comparison revealed no significant differences among the nonculture tests in specificity. Given the prevalence of *C. trachomatis* infection in this population of 3.9% based on cell culture results, these specificities yielded PPVs within the range of 85.1 to 94.0%.

Performance of screening tests with confirmatory-supplemental testing. A total of 69 clients had a negative culture result but at least one positive nonculture test result (i.e., false positive based on the culture gold standard). Confirmatory or supplemental testing had been performed following a positive result for each nonculture test except for Syva DFA. Subsequent to such testing, only seven individuals remained with a false-positive test result relative to culture (Table 3). By using culture as the gold standard and defining test positives as those which remained positive after confirmatory-supplemental testing, sensitivities and specificities were recalculated (Table 4). Specificities exceeded 99.9% for all tests. Sensitivity was not affected by the confirmatory-supplemental methods for the Chlamydiazyme and Pace 2 tests, was reduced by 1.1% for MicroTrak EIA, and was reduced by 5.7% for Sanofi EIA.

To briefly summarize the results for each test method, 100% (146 of 146) of Gen-Probe-positive specimens from culture-positive individuals were confirmed while only 22% (4 of 18) of Gen-Probe-positive specimens from culture-negative individuals were confirmed, 100% (120 of 120) of Chlamydiazyme-positive specimens from culture-positive individuals were confirmed while only 10% (2 of 21) of Chlamydiazyme-positive specimens from culture-negative individuals were confirmed, 98.6% (137 of 139) of Syva EIA-positive specimens from culture-positive individuals were confirmed while only 12% (2 of 17) of Syva EIA-positive specimens from culture-negative individuals were confirmed, and 91.5% (118 of 129) of Sanofi EIA-positive specimens from culture-positive individuals were confirmed while only 18% (3 of 17) of Sanofi EIA-positive specimens from culture-negative individuals were confirmed.

Of the seven individuals who remained classified as false positive after confirmatory-supplemental testing, i.e., were confirmatory-supplemental test positive and culture negative, five had a positive result from two or more tests (including confirmatory tests) that detect different *C. trachomatis* molecules (Table 3). The remaining two were positive by Gen-Probe only but had a percent competition of greater than 99% in the probe competition assay. An additional four individuals had 10 or more elementary bodies detected by Syva DFA. These 11 individuals were most likely truly infected and had false-negative cultures.

DISCUSSION

The principal motivations for undertaking this study were to obtain comparative sensitivity and specificity data for several chlamydia tests and to evaluate available means for performing confirmatory or supplemental testing. This information was intended to be used in the process of selecting a screening test for use in the Region X Chlamydia Project, a project which performs about 175,000 tests each year in family planning and sexually transmitted disease clinics in the states of Alaska, Idaho, Oregon, and Washington. A review of the literature turned up very few controlled comparative evaluations in which the performance of two or more tests is compared to that of a gold standard or reference standard test that defines the infection status of study subjects. Rather, most published studies have evaluated only a single test compared to a reference standard test. Therefore, to compare tests requires comparing data derived at different times on patient populations that vary in terms of size, demographics, risk factors, and prevalence and using a reference standard that is known to vary widely in performance. Differences in test performance estimated in this way are ambiguous for several reasons: (i) sample sizes of most studies are small, and the resulting estimates of sensitivity and specificity are imprecise; (ii) most studies do not present the precision of the estimates; and (iii) differences in the performance of the reference standards in different studies unpredictably bias any estimated differences in performance. The present study was designed to deal with each of these issues.

Cell culture of chlamydia was chosen as the reference stan-

TABLE 2. Pairwise statistical comparison of test sensitivities

| Comparison (test 1 vs test 2) | % Difference (test 1 – test 2) | P value |
|----------------------------------|-----------------------------------|-----------------------|
| Gen-Probe vs Syva DFA | 0.8 | 0.813 |
| Gen-Probe vs Syva EIA | 3.6 | 0.232 |
| Gen-Probe vs Sanofi EIA | 8.4 | 0.006 ^a |
| Gen-Probe vs Abbott EIA | 13.4 | <0.001 ^{a,b} |
| Syva DFA vs Syva EIA | 2.8 | 0.380 |
| Syva DFA vs Sanofi EIA | 7.6 | 0.026 ^a |
| Syva DFA vs Abbott EIA | 12.6 | <0.001 ^{a,b} |
| Syva EIA vs Sanofi EIA | 4.8 | 0.195 |
| Syva EIA vs Abbott EIA | 9.8 | 0.003 ^{a,b} |
| Sanofi EIA vs Abbott EIA | 5.0 | 0.109 |

^a Difference significant at $P \leq 0.05$.

^b Difference significant at $P < 0.005$ to maintain an overall significance level of $\alpha = 0.05$, allowing for the multiple tests of significance.

TABLE 3. False-positive tests remaining after confirmatory-supplemental testing

| Patient no. | Gen-Probe | | Syva DFA | Syva EIA | | Abbott EIA | | Sanofi EIA | |
|-------------|-----------|-----------------|------------------|----------|--------------|------------|--------------|------------|--------------|
| | Screen | Confirmation | | Screen | Confirmation | Screen | Confirmation | Screen | Confirmation |
| 720 | – | ND ^a | – | – | ND | – | ND | + | + |
| 2258 | + | + | – | – | ND | – | ND | – | ND |
| 2691 | + | + | 10+ ^b | + | + | + | + | + | + |
| 30072 | – | ND | – | – | ND | – | ND | + | + |
| 40264 | – | ND | – | + | + | – | ND | – | ND |
| 80300 | + | + | 2 | + | – | + | + | + | – |
| 90692 | + | + | – | – | ND | – | ND | – | ND |

^a ND, not done because initial result was negative.

^b Ten or more elementary bodies were detected on the DFA slide.

dard for this study. Culture was always performed on the first chlamydia swab to eliminate any swab order effect on the culture result. An advantage of using cell culture as the reference standard to classify subjects as infected or not is its high specificity. Therefore, our estimates of test sensitivities are unlikely to have been underestimated due to false-positive culture results. As discussed above, a drawback of cell culture as the reference standard is that it is less sensitive than it is specific. Knowing this, most investigators of chlamydia test performance augment their culture standard with additional testing to identify possible false-negative cultures. One commonly used method to correct for possible false-negative cultures is to create a reference standard that defines truly infected persons as those who are culture positive or are culture negative but positive by the test under evaluation and also positive by an additional test that detects a different chlamydia macromolecule (lipopolysaccharide, major outer membrane protein, or nucleic acid) (5). Revised sensitivities and specificities are then calculated by using the alternate reference standard. For chlamydia tests, there is not general agreement as to what the ideal reference standard should be (8, 9, 18). Taking a different tack, Hadgu and Qu have reported the results of this study to statistically inclined readers by using a latent class model analysis that yielded estimates of sensitivity and specificity for all the tests without designating any as a reference standard (10, 22).

In the present study, 69 subjects had a negative culture and a positive nonculture screening test. Only 5 of these 69 had a positive result from two or more tests that detect a different *C. trachomatis* molecule and would have been considered true positives by using a revised reference standard. Thus, use of an alternative reference standard to culture (i.e., positive nonculture tests detecting at least two different chlamydia macromolecules) would not have substantially changed the performance characteristics reported in Table 1.

The sensitivities of the nonculture tests relative to culture

ranged from 61.9 to 75.3% (Table 1). Significant differences among the tests were identified (Table 2). However, a consequence of the experimental design is that the sensitivity calculated for each test is an average taken over five different swab positions. Although we did not formally test for an effect of swab order, which was randomized to avoid confounding, sensitivities were consistently higher for specimens collected with the first swabs taken after cell culture. Specificity did not appear to be affected by swab order.

The specificities of the nonculture tests evaluated in this study relative to the culture reference standard ranged from 99.6 to 99.8%. The specificities reported in the package inserts for Gen-Probe Pace 2, Syva MicroTrak DFA, and Syva MicroTrak EIA are 97.0 to 99.7, 98.0, and 97.0%, respectively. Published estimates are generally equal to or lower than the package insert values. The higher specificity values reported here may reflect a relatively high culture sensitivity in the present study. Possibly, the use of multiple swabs may also have removed contaminating vaginal secretions containing substances that lead to false-positive tests. However, if this were true, we would have expected the specificities calculated with results from the third through fifth swabs taken after culture to be higher than those calculated with results from the first two swabs, and this was not the case.

Application of confirmatory or supplemental tests to verify a positive nonculture screening result had the effect of eliminating most of the false-positive results. At the same time, specimens that were culture positive and positive by a nonculture test remained positive with only a few exceptions (2 for Syva EIA and 11 for Sanofi EIA). (We should point out that the confirmation method for Sanofi EIA that we evaluated, a major outer membrane protein-based DFA, has never been submitted to the Food and Drug Administration for licensing by the manufacturer.) The results from this study suggest that confirmatory tests are a practical aid to clinicians in patient management, particularly in areas where the prior probability

TABLE 4. Performance characteristics relative to culture of nonculture tests after confirmation of positive results

| Test (confirmation test) | Sensitivity ^a | 95% CI (%) for sensitivity | Specificity ^a | 95% CI (%) for specificity | No. of false positives | PPV ^b |
|--------------------------------|--------------------------|----------------------------|----------------------------------|----------------------------|------------------------|------------------|
| Gen-Probe (Sanofi DFA) | 75.3 (146/194) | 68.6–81.2 | 99.92 (4,708/4,712) | 99.78–99.98 | 4 | 0.973 |
| Syva EIA (blocking antibody) | 70.6 (137/194) | 63.7–76.9 | 99.96 (4,710/4,712) ^c | 99.85–99.99 | 2 | 0.986 |
| Abbott EIA (Syva OFA) | 61.9 (120/194) | 54.6–68.7 | 99.96 (4,711/4,713) | 99.85–99.99 | 2 | 0.984 |
| Sanofi EIA (probe competition) | 61.1 (118/193) | 53.9–68.1 | 99.94 (4,712/4,715) ^d | 99.81–99.99 | 3 | 0.975 |

^a True positive equals culture positive. Data were generated by using the manufacturer's recommended cutoffs and results after confirmation. Values are reported as percents (numbers of specimens positive/total numbers of specimens).

^b Prevalence of 3.9% based on cell culture.

^c No confirmatory test was performed on one specimen that was dropped from the analysis.

^d No confirmatory test was performed on five specimens that were dropped from the analysis.

of chlamydia infection in a particular patient or subpopulation is low.

The importance of confirmatory-supplemental testing can be illustrated by applying the screening tests used in this study to a population of women where the prevalence of chlamydia infection is 2.0%. This prevalence may reflect the reality of many rural and even urban clinics where control programs have been in effect. For example, in the Region X Chlamydia Project the average prevalence across all clinic sites is about 4.0% even when using selective screening criteria to identify high-risk women for testing. However, nearly 10% of the individual clinics have a prevalence of chlamydia less than 2.0% (7a). With the specificity results from Table 1, the PPVs for the nonculture tests would range from 0.606 for Abbott EIA to 0.800 for Syva DFA. Therefore, between 20 and 40% of all positive results would be false positive in such a population.

By applying confirmatory or supplemental testing, up to 90% of false positives will not be confirmed whereas nearly all true positives will be. Provided with the confirmatory test result, clinicians can decide whether to treat or to retest individuals with a positive screening test. Based on this type of consideration, the Region X Chlamydia Project elected to institute mandatory confirmatory testing regardless of which screening test is performed. Further, the CDC currently recommend that positive screening tests be confirmed in patient populations where the prevalence is below 5% (5). The results described herein support these recommendations.

The present study served to clarify several complex issues involving the characterization of nonculture tests for chlamydia. First, the direct comparison of tests provided unambiguous evidence for differences in test sensitivity. Second, all of the tests were more specific than what has been reported in package inserts and in the literature generally. Third, confirmatory-supplemental tests reduced the number of false-positive results generated by the nonculture tests. For the products currently marketed with confirmatory procedures that were evaluated in this study (Abbott, Gen-Probe, and Syva), there was little or no increase in the number of false-negative results. This is of practical importance in low-prevalence settings that are becoming more common in the United States. This study also demonstrated the feasibility and utility of a multitest comparison using sequential swabs from the same study subjects. However, swab-order effects may need to be considered if more than two tests are compared simultaneously with a reference standard test. We anticipate that future studies of nonculture tests, including the newly developed nucleic acid amplification tests, will employ some of the design characteristics described here.

REFERENCES

- Barnes, R. C. 1989. Laboratory diagnosis of human chlamydial infections. *Clin. Microbiol. Rev.* **2**:119-136.
- Black, C. M. 1997. Current methods of laboratory diagnosis of *Chlamydia trachomatis* infections. *Clin. Microbiol. Rev.* **10**:160-184.
- Blanding, J., L. Hirsch, N. Stranton, T. Wright, S. Aarnaes, L. M. de la Maza, and E. M. Peterson. 1993. Comparison of the Clearview Chlamydia, the PACE 2 assay, and culture for detection of *Chlamydia trachomatis* from cervical specimens in a low-prevalence population. *J. Clin. Microbiol.* **31**:1622-1625.
- Britton, T. F., S. DeLisle, and D. Fine. 1992. STDs and family planning clinics: a regional program for chlamydia control that works. *Am. J. Gynecol. Health* **VI**:80-87.
- Centers for Disease Control and Prevention. 1993. Morbidity and mortality weekly report, vol. 42, no. RR-12. Recommendations for the prevention and management of *Chlamydia trachomatis* infections, 1993. Centers for Disease Control and Prevention, Atlanta, Ga.
- Chan, E. L., K. Brandt, and G. B. Horsman. 1994. A 1-year evaluation of Syva MicroTrak *Chlamydia* enzyme immunoassay with selective confirmation by direct fluorescent-antibody assay in a high-volume laboratory. *J. Clin. Microbiol.* **32**:2208-2211.
- Clarke, L. M., M. F. Sierra, B. J. Daidone, N. Lopez, J. M. Covino, and W. M. McCormack. 1993. Comparison of the Syva MicroTrak enzyme immunoassay and Gen-Probe PACE 2 with cell culture for diagnosis of cervical *Chlamydia trachomatis* infection in a high-prevalence female population. *J. Clin. Microbiol.* **31**:968-971.
- Fine, D. Unpublished data.
- Hadgu, A. 1996. The discrepancy in discrepant analysis. *Lancet* **348**:592-593.
- Hadgu, A. 1997. Bias in the evaluation of DNA amplification tests for detecting *Chlamydia trachomatis*. *Stat. Med.* **16**:1391-1399.
- Hadgu, A., and Y. Qu. 1998. A biomedical application of latent class models with random effects. *J. R. Stat. Soc. Ser. C* **47**:603-616.
- Iwen, P. C., T. M. H. Blair, and G. L. Woods. 1991. Comparison of the Gen-Probe Pace 2 system, direct fluorescent antibody, and cell culture for detecting *Chlamydia trachomatis* in cervical specimens. *Am. J. Clin. Pathol.* **95**:578-582.
- Kellogg, J. A. 1989. Clinical and laboratory considerations of culture vs antigen assays for detection of *Chlamydia trachomatis* from genital specimens. *Arch. Pathol. Lab. Med.* **113**:453-460.
- Kellogg, J. A., J. W. Seiple, and M. E. Hick. 1992. Cross-reaction of clinical isolates of bacteria and yeasts with the Chlamydiazyme test for chlamydial antigen, before and after use of a blocking reagent. *Am. J. Clin. Pathol.* **97**:309-312.
- Kellogg, J. A., J. W. Seiple, J. L. Klinedinst, and J. S. Levisky. 1991. Impact of endocervical specimen quality on apparent prevalence of *Chlamydia trachomatis* infections diagnosed using an enzyme-linked immunosorbent assay method. *Arch. Pathol. Lab. Med.* **115**:1223-1227.
- Kellogg, J. A., J. W. Seiple, C. L. Murray, and J. S. Levisky. 1990. Effect of endocervical specimen quality on detection of *Chlamydia trachomatis* and on the incidence of false-positive results with the Chlamydiazyme method. *J. Clin. Microbiol.* **28**:1108-1113.
- Kleinbaum, D. G., L. L. Kupper, and K. E. Muller. 1988. The LSD approach, p. 362-365. In D. G. Kleinbaum et al. (ed.), *Applied regression analysis and other multivariable methods*. PWS-Kent Publishing Company, Boston, Mass.
- Lossick, J., S. DeLisle, D. Fine, D. Masure, V. Lee, and C. Smith. 1990. Regional program for widespread screening for *Chlamydia trachomatis* in family planning clinics, p. 575-579. In W. R. Bowie, H. D. Caldwell, R. P. Jones, P.-A. Mardh, G. L. Ridgway, J. Schachter, W. E. Stamm, and M. E. Ward (ed.), *Chlamydial infections: proceedings of the seventh international symposium on human chlamydial infections*. Cambridge University Press, Cambridge, United Kingdom.
- McGee, D. L., G. H. Reynolds, D. Danielson, I. Falk, K. Persson, J. Schachter, W. E. Stamm, T. C. Quinn, and A. Hadgu. 1996. Letter to the editor. *Lancet* **348**:1307-1309.
- Mills, R. D., A. Young, K. Cain, T. M. Blair, M. A. Sitorius, and G. L. Woods. 1992. Chlamydiazyme plus blocking assay to detect *Chlamydia trachomatis* in endocervical specimens. *Am. J. Clin. Pathol.* **97**:209-212.
- Moncada, J., J. Schachter, G. Bolan, J. Engelman, L. Howard, I. Mushahwar, G. Ridgway, G. Mumtaz, W. Stamm, and A. Clark. 1990. Confirmatory assay increases specificity of the Chlamydiazyme test for *Chlamydia trachomatis* infection of the cervix. *J. Clin. Microbiol.* **28**:1770-1773.
- Østergaard, L., and J. K. Møller. 1995. Use of PCR and direct immunofluorescence microscopy for confirmation of results obtained by Syva MicroTrak *Chlamydia* enzyme immunoassay. *J. Clin. Microbiol.* **33**:2620-2623.
- Qu, Y., and A. Hadgu. 1998. A model for evaluating sensitivity and specificity for correlated diagnostic tests in efficacy studies with an imperfect reference test. *J. Am. Stat. Assoc.* **93**:920-928.
- Royal, R. M. 1986. Model robust confidence intervals using maximum likelihood estimators. *Biometrics* **54**:221-226.
- Schachter, J. 1986. Rapid diagnosis of sexually transmitted diseases: speed has a price. *Diagn. Microbiol. Infect. Dis.* **4**:185-189.
- Schachter, J., W. E. Stamm, M. A. Chernesky, E. W. Hook, R. B. Jones, F. N. Judson, J. A. Kellogg, B. LeBar, P. A. Mardh, and W. M. McCormack. 1992. Nonculture tests for genital tract chlamydial infection. What does the package insert mean, and will it mean the same thing tomorrow? *Sex. Transm. Dis.* **19**:243-244.
- Schwebke, J. R., W. E. Stamm, and H. H. Handsfield. 1990. Use of sequential enzyme immunoassay and direct fluorescent antibody tests for detection of *Chlamydia trachomatis* infections in women. *J. Clin. Microbiol.* **28**:2473-2476.
- Skulnick, M., G. W. Small, A. E. Simor, D. E. Low, H. Khosid, S. Fraser, and R. Chau. 1991. Comparison of the Clearview Chlamydia test, Chlamydiazyme, and cell culture for detection of *Chlamydia trachomatis* in women with a low prevalence of infection. *J. Clin. Microbiol.* **29**:2086-2088.
- Stamm, W. E. 1988. Diagnosis of *Chlamydia trachomatis* genitourinary infections. *Ann. Intern. Med.* **108**:710-717.
- Stamm, W. E., M. Tam, M. Koester, and L. Cles. 1983. Detection of *Chlamydia trachomatis* inclusions in McCoy cell cultures with fluorescein-conjugated monoclonal antibodies. *J. Clin. Microbiol.* **17**:666-668.
- Warren, R., B. Dwyer, M. Plackett, K. Pettit, N. Rizvi, and A.-M. Baker. 1993. Comparative evaluation of detection assays for *Chlamydia trachomatis*. *J. Clin. Microbiol.* **31**:1663-1666.