Title
Immunodiagnosis of sexually transmitted disease.

Permalink
https://escholarship.org/uc/item/4s29739j

Journal
The Yale journal of biology and medicine, 58(5)

ISSN
0044-0086

Author
Schachter, J

Publication Date
1985-09-01

Peer reviewed

Immunodiagnosis of Sexually Transmitted Disease

JULIUS SCHACHTER, Ph.D.

*Department of Laboratory Medicine, University of California,
San Francisco, California*

Methods for detecting microbial antigens in clinical specimens offer an alternative to culture in the diagnosis of some sexually transmitted diseases. Developers of the immunologic methods are faced with a number of problems in evaluating the new tests. Traditionally, these tests are compared to culture as the "gold standard." Unfortunately, culture for *Neisseria gonorrhoeae* or *Chlamydia trachomatis*—the two agents most commonly sought—is considerably less sensitive than 100 percent. Immunologic methods may appear to produce false positives when the paired specimens are actually false-negative cultures. Another source of discordant results is sampling variation. These considerations, however, will not account for all false-positive results. Even the best non-culture methods have a low rate of false-positive results. If a new test has a specificity of 97 percent, it, by definition, yields approximately 3 percent false-positive reactions. In low-prevalence settings this false-positive rate will create problems in interpreting the results. For example, in a population with 3 percent prevalence of infection, a positive result in a 97 percent specificity test could only have a predictive value of 50 percent. Most testing for STD agents is performed in low-prevalence settings. None of the currently available immunodiagnostic procedures has a performance profile that suggests it will be satisfactory for diagnostic use in the low-prevalence setting.

The purpose of this article is to discuss immunologic methods of diagnosing sexually transmitted disease (STD); this will not be a review of the literature. The focus will be on evaluation of the new methods, what can be expected from such tests, and how they should be used. Much of the article will be based on the experience of the author's laboratory in evaluation of immunologic methods for diagnosing gonococcal and chlamydial infections.

We will simply accept the premise that commercially available reagents for detecting antigens of *Neisseria gonorrhoeae* or *Chlamydia trachomatis* in clinical specimens do work. The basic questions are, "How well do they work?" and "How may these tests best be used?" At this writing, there are two procedures that are available: immunofluorescent (FA) procedures, usually using monoclonal antibodies and requiring microscopic evaluation, or enzyme immunoassay (EIA) procedures where color indicator reactions may be read either visually or by a spectrophotometer to signal the presence of a detected antigen.

EVALUATION OF THE NEW TEST—HOW GOOD IS THE "GOLD STANDARD"?

Non-culture methods are usually compared to culture as the gold standard for diagnosing infection [1,2]. When a new test is being evaluated, it is typical to express

443

TABLE 1
Evaluation of a Non-Culture Diagnostic Test

		Culture +	Culture −	
New Test	+	a	b	(a + b)
	−	c	d	(c + d)
		(a + c)	(b + d)	

Element	Formula	Definition
Sensitivity	a/(a + c)	Ability of test to identify people who are culture-positive
Specificity	d/(d + b)	Ability of test to identify people who are culture-negative
Predictive Value (+)	a/(a + b)	Probability of being culture-positive given a positive test
Predictive Value (−)	d/(c + d)	Probability of being culture-negative given a negative test
False positives	b	Culture-negative specimens that test positive
False negatives	c	Culture-positive specimens that test negative
Agreement:	(a + d)/(a + b + c + d)	

the performance pattern of the test in terms of sensitivity, specificity, predictive value of a positive and negative test and agreement seen when the new test is compared to the gold standard (Table 1).

The gold standard is usually assumed to have a sensitivity and specificity of 100 percent. In other words, in the gold standard test there are neither false positives nor false negatives. Unfortunately, that is a false assumption. It is true that, with correct laboratory technique (i.e., no cross-contamination and correct identification of fermentation patterns and so on), that there will be no false positives generated by culture. However, there can be a problem with sensitivity of culture. It is clear that culture methods for chlamydial infections or gonococcal infections do not detect 100 percent of infections.

With gonococcal infections, culture failure may be caused by vancomycin susceptibility; ultimately there is a potential problem with susceptibility to any antibiotic that may be included in a selective medium [3].

Obviously there are sampling problems, caused in part by temporal variations in the number of organisms that are present in clinically relevant sites. Thus, when samples are collected at a time of low-level infection, the infection may not be detected. Collection of multiple swabs or sampling of multiple sites has been shown to increase positivity rates [4,5]. For the sake of convenience, it has been recommended that for gonococcal culture two swabs be collected from the cervix and processed as a single specimen [6]. For chlamydial culture, increased isolation rates are seen if a cervical and a urethral swab are collected and processed as one specimen. These seem reasonable approaches and are realistic compromises with culturing multiple swabs from the cervix, rectum, and so forth. However, if such measures of maximizing

TABLE 2
Repeat Cervical Culture in Untreated Gonorrhea

Day of Cycle	No. Positive/No. Tested (%)
1–5	231/290 (80)
>6	884/1,199 (74)
Total	1,115/1,489 (75)

Data from [7]

culture results are necessary, it attests to the potential lack of reliability of a single swab culture and questions the assumption that matched swab specimens always yield the same results.

One way of assessing the sensitivity of culture is to determine the reproducibility of positive cultures after short intervals. One such study, of repeat cultures of women who had been culture-positive for *N. gonorrhoeae* and were called back for therapy, found approximately 75–80 percent positivity at the second culture [7] (Table 2). While it is likely that some of the negatives were due to spontaneous clearance of the organism, it is far more likely that most represent false-negative culture results.

It is probable that culture for chlamydial infection is equally insensitive. Oriel and Ridgway found that 83 percent of women who had been identified as source contacts for men with chlamydial urethritis yielded the organism from the cervix [8]. It is likely that the majority of the negative attempts in that study were also culture misses. In a longitudinal study of pregnant women in San Francisco, we found approximately 75 percent of cultures are positive when untreated women are retested. Most of the women who had had negative attempts were culture-positive when third or fourth isolation attempts were performed. Thus, it is likely that for these asymptomatic women, whose cervical infections were detected during routine screening, the sensitivity of culture is only about 75 percent.

The relative insensitivity of culture presents a dilemma for anyone evaluating a non-culture method for diagnosis. In typical analysis the "b" cell (Table 1) is described as the false positives by the non-culture method. But when culture is <100 percent sensitive, that cell in fact could reflect increased sensitivity of the non-culture method. Since any non-culture method is likely to have a certain rate of false positives, the "b" cell is likely to contain a summation of missed cultures ("false" false positives) and non-culture method false positives ("true" false positives). Currently there is no way of assessing this distribution within the "b" cell, although one could apply a series of different tests and accept any two as indicating a positive result, but that would probably only be valid if the tests were based on completely different criteria. It is likely that the problem of culture sensitivity becomes more important in screening asymptomatic individuals.

WHICH GOLD STANDARD?

Another question that must be dealt with when evaluating a new test is: What gold standard is being used? We have already seen that culture is <100 percent sensitive. It is imperative for an evaluator to describe carefully the culture procedure being used. The importance of defining that gold standard can be seen in Table 3, which compares different methods for isolating *C. trachomatis*. These data were generated in screening asymptomatic women for chlamydial infection of the cervix. In that setting one can see

TABLE 3

Comparison of Isolation Methods for *Chlamydia trachomatis*

Method (All Blind Pass)	% Isolates Obtained[a]
96-well plate; iodine	48
96-well plate; fluorescent antibody	73
Vials; iodine	92

[a]Cervical screening, asymptomatic population, 48/618 (7.8 percent) positive in any of the three methods; inoculum centrifuged onto monolayers (2,700 *g* for one hour), cycloheximide-treated McCoy cells, blind passage

that the use of 96-well microtiter plates is far more sensitive when the monolayers are stained with an FA procedure for detection of inclusions than when iodine stain is used. However, the use of one-dram vials (with a larger surface area of cell monolayer), even when iodine stain is used, is more sensitive than any microtiter plate procedure. Most discrepancies were found with specimens that had <10 inclusions per coverslip in the one-dram vials. If more agent was found, the specimens tended to be positive by any method.

We ran parallel tests with plates and vials, splitting the inoculum between the two systems. The comparative tests were performed over several months while we evaluated different staining methods and different procedures for blind passage. No matter what variables we used, the vials were always superior. In one parallel series, other than that shown in Table 3, we found that in screening asymptomatic women we recovered *Chlamydia* from 93 of 1,229 (7.6 percent) cervical specimens. Ninety of the specimens were positive in vials, while 69 were positive in plates. Of 52 specimens that yielded <10 inclusions per vial coverslip, 19 (37 percent) were missed in the plates. Of the total 21 misses in the plates, 19 (90 percent) were from the low-titer inoculum group.

When FA techniques were used to stain the coverslips in vials, the procedure was even more sensitive. In some laboratories it has been found that all positive chlamydial cultures could be detected in the first pass when the more sensitive FA procedure was used, but that has not been our experience in screening asymptomatic individuals. A second pass, even when FA was used for staining both primary culture and passage, resulted in 13 percent more isolates being obtained after the passage. Thus, one must accept the gold standard for chlamydial culture, at the moment, to be the use of one-dram vials, a blind passage, and the FA technique for staining inclusions. The MicroTrak® (Syva Co., Palo Alto, CA) *Chlamydia trachomatis* Culture Confirmation Test is the commercially available reagent for which there is an extensive published experience. It may be considered the reagent of choice [9]. Other FA reagents for staining inclusions in cell culture are available (Ortho Chlamydia Direct Identification Test®; Ortho Diagnostic Systems, Inc., Carpinteria, CA) but have not been as well evaluated.

This is not to say that other procedures, such as using microtiter plates or not doing a blind passage, are inappropriate. If our studies had been performed in an STD clinic, where more symptomatic patients were being tested, the differences in procedure

would not be as great. Selecting an asymptomatic screening setting artificially exaggerates the differences in sensitivity. Whether or not a blind pass is useful depends on a cost-benefit analysis. In symptomatic patients we find that approximately 10 percent of isolates are obtained in the second pass. Increasing the recovery rate of chlamydiae from 20 percent to 22 percent in a population may well be desirable, but it is probably far more practical to test another 80 patients and detect another 16 (20 percent) infections rather than do the blind pass to identify another two infections. In screening asymptomatic women, however, we find that up to 40 percent of the isolates are detected in the blind pass.

The above discussions obviously stress the importance of recognizing the fallibility of the gold standard but also the necessity for adequate description. It should lead a reader of the literature to a better understanding of how new tests are truly performing. If a non-culture method for detecting chlamydial infection is being compared to culture in an asymptomatic population and no blind passage is done, one can predict that the results will be markedly skewed by the relative insensitivity of the gold standard. If the tests agree, then the non-culture assay is relatively insensitive as well. If there are many false positives detected by the non-culture assay, one must wonder how many of those false positives would become true positives with a second pass (i.e., being moved from the "b" cell to the "a" cell in Table 1).

ANALYSIS OF THE EVALUATION; GONOZYME AS AN EXAMPLE

The application of predictive value analysis is crucial in determining how a test should be used. It is usually more important to stress the PV (+). With STD, in most clinics, PV (−) and agreement are always high, because most patients do not have the disease and the majority of the population is in the "d" cell (Table 1). For example, if the population has a true prevalence of 10 percent as they enter a clinic, someone standing at the door saying, "You don't have it," will be in agreement with culture >90 percent of the time and the predictive value of that negative judgment will also be >90 percent.

Obviously there are other practical considerations. When we analyzed the Gonozyme® (Abbott Laboratories, Chicago, IL) test for diagnosing gonococcal infections in the male urethra, we found that the test performed very well, but it did not present any marked improvement over the inexpensive and rapid Gram stain procedure (Table 4). Thus, there appeared to be no reason to recommend that test for diagnosis of symptomatic male urethritis.

However, Gram stain is not as useful for cervical infection, where improved diagnostic tests are needed. The findings with Gonozyme in diagnosing gonococcal infection of the cervix are presented in Table 4. The test obviously performs reasonably well. The 87 percent sensitivity leaves something to be desired in an STD setting, but the performance profile is quite a reasonable one in a high-prevalence setting. However, the 98 percent specificity which looks like a very good result (indeed, how much more can one expect from developers of such a test), severely restricts the use of the test in a low-prevalence setting. The results expected if a test with that performance profile is used in a 1 percent prevalence population are shown in Table 5. It becomes obvious that more false-positive results are being generated than are true positives. This clearly creates a problem because the clinician has a dilemma in how to manage these patients. Should they be treated? Should partners be treated routinely? Should

TABLE 4

Diagnosis of Gonococcal Infection by Gonozyme and Culture

A. Gonozyme in male urethritis

		Culture	
		+	−
EIA	+	70	0
	−	5	42

Prevalence = 64 percent   PV (+) = 100 percent
Sensitivity = 93 percent   PV (−) = 89 percent
Specificity = 100 percent   Agreement = 96 percent

B. Gonozyme in cervical infection

		Culture	
		+	−
EIA	+	33	1
	−	5	47

Prevalence = 44 percent   PV (+) = 97 percent
Sensitivity = 87 percent   PV (−) = 90 percent
Specificity = 98 percent   Agreement = 93 percent

Data from [10]

these positive results be reported to the public health officials as diagnosed cases of gonorrhea? Obviously these are serious questions, and clinicians would prefer not to have to deal with them. A clinician wants a diagnostic test to provide answers, not create further problems.

Of course, a clinician can always say, "Why bother generating this type of analysis?" when the test was actually evaluated in high-prevalence settings. But, in fact, that analysis was not a philosophical endeavor but is truly representative of the way some of these tests are applied. For example, the most recent available data on cervical cultures for gonorrhea performed in this country is for the year 1983 [11] (Table 6). Eighty-nine percent of cultures for gonorrhea are performed in settings other than STD clinics. The prevalence of gonococcal infection in STD clinics was approximately 17 percent, but it was only 2.5 percent in the setting where most tests are performed. A test with the performance characteristics described above would have a positive predictive value of only 53 percent in other than STD clinics. Such a predictive value may provide information of some use to the clinician; it may indicate the need for a second test, perhaps by culture. That depends on the cost of culture, the dangers of not treating, and the various dangers of different intervention modalities.

TABLE 5

Results Expected Using Gonozyme and Culture in Screening 10,000 Women with a One Percent Prevalence of Cervical Gonorrhea

Gonozyme Results	Culture Results	
	Positive	Negative
Positive	87	205
Negative	13	9,695

Data from [10]

TABLE 6
Gonorrhea Cultures Performed on Females in the United
States—1983

	No. Positive/No. Tested (%)
VD Clinics	148,429/885,231 (16.8)
Other	170,673/6,835,629 (2.5)
Total	319,102/7,720,860 (4.1)

Data from [11]

The latter dangers may be medical, financial, or social. The obvious question is whether any test with this performance profile should be used in this low-prevalence setting.

EVALUATION OF MICROTRAK® DIRECT SPECIMEN TEST

Our evaluation of the FA procedure using monoclonal antibodies specific for major outer membrane antigen of *C. trachomatis* has been somewhat less satisfactory than the published experience. Published data have indicated a sensitivity of >90 percent and specificity levels of approximately 98 percent [12]. Some results have been far better than that, with virtual perfect concordance. Some researchers have claimed sensitivities of well over 100 percent, as compared to culture. They have assumed that because a monoclonal antibody was being used, that all reactions were truly specific and that all positives not confirmed by culture results represented culture misses. That type of evaluation is naive. The FA reagent is a superb one. It stains chlamydial particles brightly and there is very little nonspecific fluorescence in the clinical specimen. Certainly it is far superior to the polyclonal reagents that had been previously used, but there are true false-positive results obtained with this reagent. Some small round bacteria have been observed to stain. We have also seen persistence of FA staining that appears to be elementary body staining, after tetracycline therapy and in the absence of symptoms. The ultimate source of these confounding reactions awaits elucidation, but suffice it to say they do exist, and the rates must be defined.

Our experience is summarized in Table 7. Here too we were testing asymptomatic individuals and that works against the performance profile of a non-culture method. Sensitivity of the FA procedure was 70 percent. The 94 percent specificity obviously reflects both false positives and culture misses. One problem identified in our evaluation was that 15 percent of the smears for the FA procedure were inadequate

TABLE 7
MicroTrak® Compared to Culture for Chlamydial Cervical
Infection

Element	No. Positive/No. Tested (%)
Prevalence by culture	90/655 (13.7)
Agreement	592/655 (90)
FA sensitivity	63/90 (70)
FA specificity	529/565 (94)
PV (+) FA	63/99 (64)
PV (−) FA	529/556 (95)

(sufficient numbers of cells could not be seen microscopically). However, we have no way of knowing whether a specimen submitted for an isolation attempt is adequate. Where the FA smear was appropriately collected and positive, we could have what appears to be a false positive, if the parallel culture specimen was not adequate. The reverse could not happen, because we would see that the smear was inadequate and exclude the paired specimen from the analysis.

It is likely that the FA procedure will be more efficient in diagnosing symptomatic infection. In a small study of such patients we found sensitivity of 80 percent, specificity of 97 percent (data not shown). We concluded that isolation was the method of choice for diagnosing chlamydial cervical infection. If culture is not available, the FA procedure is useful in testing high-prevalence symptomatic populations.

WHERE SHOULD THESE TESTS BE USED—FUTURE DEVELOPMENTS

There are obviously a number of settings in which non-culture methods of diagnosis are desirable: where culture methods are expensive or difficult to perform, or are not readily available, or where transport problems may result in loss of viability. In addition, non-culture methods may be suited for screening. Manufacturers must decide whether these tests are to be used for diagnostic tests or for screening. For the latter purpose a clinician would want high sensitivity (to identify all positives) and any problems with specificity could be dealt with by a more specific confirmatory test. The latter could be another test (culture?) or perhaps a different configuration of the same antigen detection method.

Obviously the manufacturers will endeavor to improve the performance of their products. We, as scientists, can hasten that process by refusing to accept inadequate performance characteristics.

Non-culture methods of diagnosis can have a great public health impact. They hold the promise of making accurate diagnosis of gonococcal and chlamydial infection more readily available. These methods may allow us to test populations that currently escape surveillance. For example, we have spent much effort in screening asymptomatic women and trying to define risk factors for chlamydial infection [14]. We can generate this information because women routinely have pelvic examinations and cervical specimens are readily obtainable. In some of our studies we find infection rates of 15–25 percent in women who are not coming to clinics because of medical problems but are appearing for routine examinations (Pap smears, birth control advice, and so on). When such women are identified, obviously they and their partners (when available) are treated. However, we never deal with the male population on a routine screening basis because there are no noninvasive methods for sampling the urethra to detect chlamydial or gonococcal infections. One obvious possibility would be to test urine. A review on the use of urine in diagnosing STD concluded that it might be an appropriate specimen for use in screening asymptomatic males for gonococcal infection [15]. A study of asymptomatic adolescent males found that pyuria in first-catch urine specimens had a high predictive value for recovery of chlamydiae or gonococci from urethral swabs [16]. Obviously it would be desirable if the same specimen could be used for both screening (pyuria) and diagnosis (culture or antigen detection). If urine were an appropriate holding medium for gonococci and *Chlamydia,* it might be possible to screen first-catch urines for presence of polymorphonuclear leukocytes and then do the appropriate cultures on it. However, results obtained with urine as a holding medium for gonococcal specimens have been variable and urine or urine

TABLE 8
Gonozyme® on Urine Sediment versus Urethral Swab Culture

		Culture	
		+	−
EIA	+	25	2
	−	2	156

Agreement = 98 percent	Prevalence = 15 percent
Sensitivity = 93 percent	PV (+) = 93 percent
Specificity = 99 percent	PV (−) = 99 percent

sediment are clearly not the specimens of choice for diagnosing chlamydial infection [17]. It may well be that antigen detection offers a viable alternative. Rudrik and colleagues have shown that gonococcal antigens could be detected by the Gonozyme test in urine specimens from men with gonococcal urethritis [18]. We too have been evaluating this approach (Table 8) and find that the Gonozyme test has reasonably high sensitivity, specificity, and predictive value in diagnosing gonorrhea by detection of gonococcal antigen in urine sediment. The data presented in Table 8 are preliminary results taken from an ongoing study where we are assessing different methods of detecting gonococcal and chlamydial antigens in first-catch urine from men. Clearly this represents one potential use of immunologic tests for chlamydial or gonococcal infection.

CONCLUSIONS

In summary it is clear that immunologic methods for detecting microbial antigens in specimens being studied for STDs do work. It is equally clear, however, that none of the currently available tests has a performance profile that allows any real degree of confidence in applying the test in a low-prevalence setting. The tests seem best suited for use in high-risk populations and symptomatic individuals. The non-culture methods of diagnosis obviously must be compared against culture as a gold standard, but it is imperative that those planning to apply these tests simply not accept the numbers presented to them in the literature. One must be aware of the real performance of culture and of the potential poor value of the gold standard. If a culture method is really only 75 percent sensitive and a non-culture method is 90 percent as sensitive as the culture method, we are now accepting a real sensitivity of 65 percent. We must insist on better.

REFERENCES

1. Dans PE, Rothenberg R, Holmes KK: Gonococcal serology: How useful and how much? (Editorial). J Infect Dis 135(2):330–333, 1977
2. Vecchio TJ: Predictive value of a single diagnostic test in unselected populations. New Eng J Med 274:1171–1173, 1966
3. Brorson JE, et al: Vancomycin-sensitive strains of Neisseria gonorrhoeae: A problem for the diagnostic laboratory. Br J Vener Dis 49:452, 1973
4. Paavonen J: Chlamydia trachomatis-induced urethritis in female partners of men with nongonococcal urethritis. Sex Transm Dis 6(2):69–71, 1979
5. Chapel TA, Smeltzer M: Culture of urinary sediment for the diagnosis of gonorrhea in women. Br J Vener Dis 51:25–27, 1975
6. Judson FN, Werness BA: Combining cervical and anal-canal specimens for gonorrhea on a single culture plate. J Clin Microbiol 12(2):216–219, 1980

7. McCormack WM, Reynolds GH, Cooperative Study Group: Effect of menstrual cycle and method of contraception on recovery of *Neisseria gonorrhoeae*. JAMA 247(9):1292–1294, 1982

8. Oriel JD, Ridgway GL: Studies of the epidemiology of chlamydial infection of the human genital tract. In Chlamydial Infections. Edited by P-A Mardh, KK Holmes, JD Oriel, P Piot, J Schachter. Amsterdam, Elsevier Biomedical, 1982, pp 425–428

9. Stamm WE, Tam M, Koester M, Cles L: Detection of *Chlamydia trachomatis* inclusions in McCoy cell cultures with fluorescein-conjugated monoclonal antibodies. J Clin Microbiol 17(4):666–668, 1983

10. Schachter J, McCormack WM, Smith RF, Parks RM, Bailey R, Ohlin AC: Enzyme immunoassay for diagnosis of gonorrhea. J Clin Microbiol 19(1):57–59, 1984

11. Centers for Disease Control, Division of Sexually Transmitted Diseases: STD Statistics 133: 50, Table 36. Atlanta, GA, Statistics Services Section, 1985

12. Tam MR, Stamm WE, Handsfield HH, Stephens R, Kuo C-C, Holmes KK, Ditzenberger K, Crieger M, Nowinski RC: Culture-independent diagnosis of *Chlamydia trachomatis* using monoclonal antibodies. New Eng J Med 310(18):1146–1150, 1984

13. Lipkin E, Moncada J, Schachter J: Comparison of monoclonal antibody staining and culture in diagnosing cervical chlamydial infection. In preparation

14. Schachter J, Stoner E, Moncada J: Screening for chlamydial infections in women attending family planning clinics: Evaluations of presumptive indicators for therapy. W J Med 138(3):375–379, 1983

15. Schachter J: Urine as a specimen for diagnosis of STD. Am J Med 28:93–97, 1983

16. Adger H, Shafer M-A, Sweet RL, Schachter J: Screening for *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in adolescent males: Value of first catch urine examination. Lancet ii(8409):944–945, 1984

17. Smith TF, Weed LA: Comparison of urethral swabs, urine and urinary sediment for the isolation of *Chlamydia*. J Clin Microbiol 2(2):134–135, 1975

18. Rudrik JT, Waller JM, Britt EM: Efficacy of an enzyme immunoassay with uncentrifuged first-voided urine for detection of gonorrhea in males. J Clin Microbiol 20(3):577–578, 1984