JOURNAL OF CLINICAL MICROBIOLOGY, Nov. 2000, p. 4301–4302 Vol. 38, No. 11
0095-1137/00/$04.00+0

Discrepant Analysis Is an Inappropriate and Unscientific Method

In a recent guest commentary, McAdam (6) using a hypothetical example demonstrated that estimates of test performance parameters (sensitivity and specificity, etc.) could be increased markedly and inappropriately through the use of discrepant analysis. He correctly concluded that discrepant analysis is an inherently biased and unsatisfactory approach to parameter estimation. Although discrepant analysis has been pilloried in leading journals, including *Lancet*, the *Journal of Clinical Epidemiology*, *Statistics in Medicine*, and *Clinical Infectious Diseases*, etc., the editors of the *Journal of Clinical Microbiology* (JCM) continue to publish articles (over 100 in the last 6 or 7 years) based on discrepant analysis. For example, an article in the March issue (9) of JCM employs discrepant analysis techniques to estimate test performance parameters of PCR-based tests.

It is important that diagnostic tests (like therapeutics) are properly evaluated before they are marketed for routine clinical use. Evaluation of DNA amplification tests for many infectious diseases, including those caused by *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, *Clostridium difficile*, *Mycobacterium tuberculosis*, *Legionella* spp., *Toxoplasma gondii*, and *Helicobacter pylori*, etc., is based on discrepant analysis, an "unsatisfactory" method according to McAdam and a "ploy" to exaggerate claims of performance indices according to Hilden's invited commentary in *Lancet* (5). In the past, having entered clinical use without rigorous and sound evaluation, many diagnostic tests have proven less desirable and sometimes useless in subsequent studies. Among such tests are the dexamethasone suppression test for depression, the indirect immunofluorescence assay for Lyme disease, the carcinoembryonic antigen marker test for colon cancer, and iodine 125-labeled fibrinogen scans for deep venous thrombosis (8).

Articles that employ discrepant analysis estimates have been harshly criticized by many prominent researchers in diagnostic testing issues. Unfortunately, a great majority of these articles were published in this journal, JCM. With growing awareness of these methodological issues, it is hoped that the editors of JCM will in the future reject articles whose estimates of sensitivity and specificity are based on discrepant analysis. I also hope that they will take some corrective action or give warning of the results in previously published JCM articles. This is not to say that these tests are bad or good but rather to acknowledge the fact that discrepant analysis may call into question the validity of the results and conclusions of those published articles, which is consistent with this journal's stated policy on warnings and retractions. Whether these tests are truly good or bad can only be determined by performing the appropriate analysis. Thus, McAdam is correct to conclude that "if a newer, better test requires harder methods of analysis, we are obliged to make the effort to accurately test the test." As many new tests continue to flood the market, particularly with the new opportunities from DNA technology, a rigorous and thorough evaluation should occur before, and not after, such tests are disseminated.

The problem with discrepant analysis, in addition to its inherent bias, is that it is fundamentally unscientific. This lack of scientific credibility results from the following. First, it violates the most fundamental principle of diagnostic testing, which is that the new test should not be used in the determination of the true disease status. In discrepant analysis, the definition of true disease status is based, in part, on the outcome of the new test under investigation and its sister test (3, 4). For example, in the evaluation of the plasmid-based LCR test for *Chlamydia*, the definition of true chlamydial status of individuals is based, in part, on the outcome of the plasmid-based LCR test itself and its sister test, the major outer membrane protein-based LCR. This is analogous to LCR being simultaneously the judge and the defendant in a court of law (3). In the words of J. Hilden (5), an invited *Lancet* commentator, discrepant analysis is a situation in which "the defendant decides the procedure of the court." Second, as has been demonstrated repeatedly, even under the ideal situation where the resolution of discrepant results is performed by a perfect test, discrepant analysis estimates are upwardly biased (3, 4, 7). Thus, even under the best of conditions, the ultimate outcome of using discrepant analysis is to produce upwardly biased estimates. As such, it is untenable as a standard truth-seeking procedure. Third, as pointed out by McAdam (6) and others, the resolution of discrepant results is usually determined by a dependent and sister test. Moreover, such resolution tests have not been evaluated properly nor approved (3).

Lastly, there is not a single statistical textbook or journal that treats discrepant analysis as a legitimate statistical approach for estimating sensitivity and specificity parameters. In fact, the opinion of statisticians on discrepant analysis is very harsh. For example, Colin Begg, a prominent researcher in diagnostic testing, appropriately characterized discrepant analysis as "conceptually and logically flawed," "fundamentally unscientific," and "not a truth seeking methodology" (1, 3). In his invited *Lancet* commentary, Hilden implicitly equated discrepant analysis with "discrepant behavior" and explicitly called it a "ploy" to exaggerate claims of performance indices. He characterized discrepant analysis as "poor science" and a procedure based on "faulty logic" and "fallacious statistical arguments." Note that the criticism against discrepant analysis and its proponents and defenders comes not only from statisticians but also from independent physicians, epidemiologists, microbiologists, pathologists, and others.

In his guest commentary, McAdam pointed out that the signal amplification of nucleic acid amplification tests is extraordinarily efficient, so that even a single organism may be detected, at least in theory. He also warns that the great sensitivity of nucleic acid amplification tests may result in reduced specificity and thus increase the risk of false-positive results (10). Why? Because, as previously described (A. Hadgu, Letter, J. Clin. Epidemiol., in press), the "detection" of one *Chlamydia trachomatis* organism or the "detection" of one *Clostridium difficile* or *Mycobacterium tuberculosis* organism, e.g., may not necessarily constitute the presence of disease and the need for subsequent treatment. This is important in light of the fact that these tests may be susceptible to laboratory and aerosol contaminations. It is also possible that these tests could be amplifying dead chlamydial cells in situ. The implication of this is that although these tests are sensitive, the near-perfect specificity obtained by discrepant analysis should be suspect and that has serious ramifications for screening general populations.

Green et al. (2) claimed that the discrepant analysis-based estimates of specificity are typically less biased than those based on culture and that the discrepant analysis-based spec-

ificity shows little appreciable bias. However, in a subsequent article, I (3) showed that those conclusions are incorrect. In that article, I showed algebraically that the discrepant analysis-based estimates of sensitivity and specificity can generate a significant and clinically important overestimation of the true sensitivity and specificity values. This conclusion is consistent with the work of Miller (7). In summary, discrepant analysis is not only biased but also unscientific. To pursue the standards of good science and scientific publication, the editors of JCM should avoid publishing articles utilizing this flawed approach and alert and warn its readers of its use in previously published articles.

REFERENCES

1. **Begg, C.** 1999. Workshop of statistical and quantitative methods used in screening and diagnostic tests. Centers for Disease Control and Prevention, 3 to 5 May 1999.
2. **Green, T. A., C. M. Black, and R. E. Johnson.** 1998. Evaluation of bias in diagnostic-test sensitivity and specificity estimates computed by discrepant analysis. J. Clin. Microbiol. **36:**375–381.
3. **Hadgu, A.** 1999. Discrepant analysis: a biased and an unscientific method for estimating test sensitivity and specificity. J. Clin. Epidemiol. **52:**1231–1237.
4. **Hadgu, A.** 1997. Bias in the evaluation of DNA-amplification tests for detecting *Chlamydia trachomatis*. Stat. Med. **16:**1391–1399.
5. **Hilden, J.** 1997. Discrepant analysis — or behavior? Lancet **350:**902.
6. **McAdam, A. J.** 2000. Discrepant analysis: how can we test a test? J. Clin. Microbiol. **38:**2027–2029.
7. **Miller, W. C.** 1998. Bias in discrepant analysis. When two wrongs don't make it a right. J. Clin. Epidemiol. **51:**219–231.
8. **Reid, M. C., M. S. Lachs, and A. R. Feinstein.** 1995. Use of methodological standards in diagnostic test research. JAMA **274:**645–651.
9. **Van der Pol, B., T. C. Quinn, C. A. Gaydos, K. Crotchfelt, J. Schachter, J. Moncada, D. Jungkind, D. H. Martin, B. Turner, C. Peyton, and R. B. Jones.** 2000. Multicenter evaluation of the AMPLICOR and automated COBAS AMPLICOR CT/NG Tests for detection of *Chlamydia trachomatis*. J Clin Microbiol. **38:**1105–1112.
10. **Vaneechoutte, M., and J. Van Eldere.** 1997. The possibilities and limitations of nucleic acid amplification technology in diagnostic microbiology. J. Med. Microbiol. **46:**188–194.

Alula Hadgu
Centers for Disease Control and Prevention
Division of STD Prevention
1600 Clifton Rd.
Mail Stop E-63
Atlanta, Georgia 30333

Phone: (404) 639-8356
Fax: (404) 639-8611
E-mail: axh1@cdc.gov

Author's Reply

I am pleased that Dr. Hadgu agrees with my conclusions about the use of discrepant analysis. In my guest commentary, I urged authors and reviewers to carefully consider whether discrepant analysis is a reasonable statistical method and suggested alternative approaches for evaluating new assays. In my opinion, discrepant analysis is flawed and should be avoided.

It is my understanding that the editors of the *Journal of Clinical Microbiology* will let reviewers evaluate the statistical validity of the papers submitted to this journal, just as reviewers also evaluate the methods and scientific relevance of the papers. I think that this is a reasonable approach. This underscores the great importance of careful review of manuscripts, including the statistical methods. I would again urge reviewers to evaluate the validity of discrepant analysis.

Alexander J. McAdam
Department of Pathology
LMRC 5th Floor
221 Longwood Ave.
Boston, Massachusetts 02115