Improved Serodiagnostic Testing for Lyme Disease: Results of a Multicenter Serologic Evaluation

The diverse clinical manifestations of Lyme disease (1-3) have led to frequent confusion in clinical diagnosis, a confusion compounded by problems in the accuracy and precision of diagnostic serologic tests (4-11) and the difficulty of isolating the causative organism (12-14), *Borrelia burgdorferi* . In 1990, more than 20 commercially prepared serologic test kits for Lyme disease were being sold in the United States, but no nationally standardized reference test was available. A collaborative evaluation of a selected sample of the commercial test kits by the Centers for Disease Control and Prevention (CDC) and the Association of State and Territorial Public Health Laboratory Directors (ASTPHLD) demonstrated poor concordance of results among these test kits and among a selected group of state health department laboratories (11). Because of the lack of a rigorously defined reference serum panel, conclusions could not be drawn about the sensitivity and specificity of the test kits evaluated. An unexpected finding in this study was the low concordance in test results between CDC and two consulting academic reference center laboratories. A number of other studies also have demonstrated low concordance of Lyme disease serologic test results obtained by a variety of laboratories (4-10).

As a result of those findings, the study described here was designed to fulfill the following objectives: 1) to assemble a serum panel from patients who had clinically well-defined Lyme disease (preferably confirmed by isolation of *B. burgdorferi*); healthy controls, and persons residing in non–endemic-disease areas whose potentially cross-reactive specimens had yielded equivocal ELISA results in earlier CDC tests; 2) to test this panel in a blinded fashion by several recognized Lyme disease reference and research laboratories; and 3) to compare the accuracy and precision of tests as a prelude to developing national recommendations for standardized serologic testing for antibodies to *B. burgdorferi* .

Tests were performed by five academic centers active in Lyme disease research (the Marshfield Clinic, Marshfield, Wisconsin; University of Medicine and Dentistry of New Jersey–Robert Wood Johnson Medical School, New Brunswick, New Jersey; State University of New York at Stony Brook, Stony Brook, New York; Tufts/New England Medical Center, Boston, Massachusetts; and the University of Connecticut Health Center, Farmington, Connecticut) and CDC's Division of Vector-Borne Infectious Diseases, National Center for Infectious Diseases, based in Ft. Collins, Colorado.

Serum samples from Lyme disease case-patients were obtained from the participating academic investigators (n = 72) and from the CDC Lyme disease reference serum collection (n = 37). All case-patient serum samples (total = 109) were from patients who met the CDC clinical case definition for surveillance of Lyme disease (15). The clinical manifestations in these patients ranged from acute erythema migrans (EM) to late neurologic disease accompanied by Lyme arthritis. *B. burgdorferi* had been cultured by the method of Berger et al. from 14 of 34 (41%) acute-phase specimens provided by CDC (14). Duplicate specimens (n = 85) were randomly selected from the 109 case-patient samples for precision analysis, making a total of 194 case-patient samples in the panel.

Control serum samples were provided by CDC from unpaid healthy blood donors (n = 113) who resided in areas where Lyme disease is not endemic (Cincinnati, Ohio, and Atlanta, Georgia; travel histories were not available from these donors, however. Duplicate specimens (n = 87) also were randomly selected, resulting in 200 noncase samples in the serum panel. Additional control samples were obtained from persons who resided in areas where Lyme disease was not endemic but whose physicians submitted their serum for Lyme disease testing to CDC through their state health department (n = 113). These specimens from patients with suspected cases had borderline (equivocal) seroreactivity in the whole cell sonicate (WCS) enzyme-linked immunoassay (ELISA) used by CDC before 1992 and are referred to hereafter as "WCS-suspects" (16). The addition of duplicate specimens (n = 87) brought this group to 200 equivocally seroreactive samples.

Serum was separated and frozen by the original collectors and shipped frozen to CDC's facilities in Ft. Collins, Colorado. The specimens were divided into aliquots and coded; code labels were applied

by CDC staff not involved in serologic testing of the specimens (n = 594). The panels were then refrozen and shipped on dry ice for blind testing by participating investigators. All specimens were received frozen. To calculate test sensitivity and specificity, only the result of the sample with the lower random code number of each pair was used.

Each laboratory employed the testing method that it used routinely at the time this study was undertaken (1992). CDC used an ELISA with a WCS antigen prepared from highly passaged strain B31 (gift of A. Barbour, University of Texas Health Sciences Center, San Antonio, Texas) and an ELISA with a strain B31 flagellar antigen (FLA) then being evaluated (16, 17). The other five participating investigators used ELISA tests that employed a WCS antigen of *B. burgdorferi* . Four used assays developed in their own laboratories, and one used a commercially available test kit (18-22). Three investigators also tested all specimens by Western blotting using published methods (19, 20). Two of these three performed immunoblotting for IgM and IgG antibodies separately. One laboratory tested for IgM and IgG together.

Each participating laboratory submitted the raw data of its results, along with a dichotomous interpretation of those results as either positive or negative. By prior agreement, ELISA results that fell into a range ordinarily reported as "equivocal" by that laboratory were treated as negative for this analysis. Statistical analyses undertaken at CDC included calculations of sensitivity (true positives correctly identified), specificity (true negatives correctly identified), precision (frequency of obtaining the same result on duplicate analysis of a specimen), and a measure of concordance (agreement among investigators) of results among the tests using the kappa statistic.

The accuracy and precision of the serologic tests as performed in 1992 by all six laboratories is summarized in Table 1. The test methods of investigators 1, 2, and 3 produced essentially equivalent results, with moderately high sensitivity (73% to 79%) for the aggregate of all case-patient samples tested and high specificity (98% to 99.5%). Precision was high in these three laboratories for both blood donor samples (97% to 99%) and the WCS-suspects samples submitted from areas where Lyme disease is nonendemic (94% to 98%). Precision was somewhat lower for the case-patient samples (82% to 91%).

Table 1. Accuracy and precision of serologic tests for Lyme disease performed in 1992

	Accuracy (%)		Precision (%)		
Investigator	Sensitivity	Specificity	Case patients	Non-case-patients	WCS suspects[a]
CDC (WCS)	93	71	93	77	69
CDC (FLA)	92	82	92	79	62
1	73	99.5	89	99	98
2	76	99	82	99	97
3	79	98	91	97	94
4	49	91	79	94	93
5	40	72	63	74	77

[a] Specimens from patients with suspected cases that had borderline (equivocal) seroreactivity in an enzyme-linked immunosorbent assay with whole-cell sonicate antigen (WCS).

The performance of the other three laboratories, including CDC's, was poor. Both CDC ELISA tests had high sensitivity (92% to 93%), but low specificity (71% to 82%). Precision for case-patient specimens was fairly high (92% to 93%), but low for both non-case-patient (77% to 79%) and WCS-suspects groups (62% to 69%). The method of investigator 4 gave very low sensitivity (49%), moderately high specificity (91%), poor precision with Lyme disease case-patient specimens (79%), but good precision with blood donor and WCS-suspects samples (93% to 94%). Investigator 5, who used a commercial test, obtained results with low accuracy and precision.

Concordance was high (kappa statistic 0.700) between the results of investigators 1, 2, and 3. The CDC FLA test showed moderate concordance (kappa 0.400) with results from investigators 1, 2, and 3. The results of investigator 4 showed moderate concordance with those of investigators 1 and 2 (kappa 0.400) and low concordance (0.400) with the other results. The results of investigator 5 had low concordance with all other results. The CDC WCS test showed moderate concordance with the FLA test, but low concordance with results of the ELISA tests of the other laboratories.

The three investigators with the best results all used Western blot to supplement their ELISA. Two of these three investigators submitted their dichotomous test interpretation with and without using Western blot results. The sensitivity improved by 20% for one investigator and by 30% for the other when Western blot results were included. The improvement resulted from identifying as

positive by Western blot those case-patient specimens from which an equivocal result was obtained by ELISA and which by study design would have been counted as negative by ELISA results alone. Specificities were not affected by Western blot analysis in this group of three investigators, since the serum panel in this study did not contain cross-reactive sera; and the negative controls and WCS-suspects had negative results by both ELISA and Western blot.

Test sensitivity from the three laboratories with the best test specificity (98%) was analyzed according to the clinical manifestations in the case-patients (Table 2). As expected, the sensitivities of the tests were lowest in specimens from patients with early disease, 59% to 66% for erythema migrans and 63% to 75% for early neurologic disease. Sensitivities were much higher for samples of patients with late disease. Sensitivities of 89% to 95% were obtained for Lyme arthritis patients and 91% to 100% for persons with late neurologic disease, primarily encephalopathy or polyneuropathy.

The emergence of a disease can outstrip the development of reliable methods for its laboratory diagnosis. The serodiagnosis of Lyme disease has been fraught with problems of precision and accuracy. This study provided an opportunity for selected academic research centers and CDC to compare the performance of their individual tests by using a serum panel from clinically well-characterized patients and controls from non–endemic-disease areas. The clinical diagnosis of early Lyme disease was supported by the isolation of *B. burgdorferi* from skin biopsy specimens (14), when possible. The panel, which was coded blind had a sufficiently large number of samples (n = 335) to provide adequate statistical power for the comparison.

Laboratories that supplemented their primary test, an ELISA, with immunoblotting achieved greater test accuracy than those that did not. The use of Western blot as a second test enabled the best performing laboratories to increase test sensitivity without a concomitant loss of specificity. This increase in sensitivity occurred as a result of identifying as true positives by Western blot a number of those specimens from patients with clinical cases of Lyme disease that were interpreted as equivocal by ELISA and would have been otherwise considered in this study as dichotomously negative results. Although the investigators employing Western blot tested all panel specimens with this method, they did so at that time to evaluate the potential value of Western blot in Lyme disease serologic diagnosis.

The observation that Western blotting could be employed to resolve equivocal ELISA results gave additional impetus for evaluating its potential adjunctive role in Lyme disease serodiagnosis and eventually led to the finally recommended two-test approach (23). The potential utility of Western blotting, however, pointed out the lack of standardized methods for producing blots and standardized interpretive criteria.

Performance of the CDC WCS and FLA ELISA in this study that did not include known cross-reactive sera suggested that the positive cut-off value for these tests was inappropriately low, thereby increasing sensitivity at the expense of specificity. These results then explained the large number of borderline WCS ELISA results obtained by CDC when it tested the sera of patients residing in areas where Lyme disease was not endemic. This group of WCS suspects was nearly uniformly found to be negative on ELISA by the three laboratories with the best performance (Table 1) (23).

Table 2. Test sensitivity of laboratories demonstrating a test specificity of 98%

Clinical Manifestations	Sensitivity, % (positive samples/total)		
	Laboratory 1	Laboratory 2	Laboratory 3
Erythema migrans, all	59 (55/94)	60 (56/94)	66 (62/94)
Acute phase [a]	65 (11/17)	65 (11/17)	76 (13/17)
Convalescent phase [b]	57 (44/77)	58 (45/77)	64 (49/77)
Carditis	100 (2/2)	100 (2/2)	100 (2/2)
Lyme arthritis [c]	89 (58/65)	95 (62/65)	92 (60/65)
Neurologic, all	85 (28/33)	88 (29/33)	91 (30/33)
Early	63 (5/8)	63 (5/8)	75 (6/8)
Late	91 (10/11)	100 (11/11)	91 (10/11)
Late and arthritis	93 (13/14)	93 (13/14)	100 (14/14)
Total	74 (143/194)	77 (149/194)	79 (154/194)

[a] ≤ 30 days from onset of erythema migrans to blood collection.

[b] > 30 days from onset of erythema migrans to blood collection.

[c] Without neurologic signs or symptoms.

Specificity in this study was determined by testing specimens from blood bank donors. With these samples, specificity in the three laboratories that used immunoblotting was very high (98% to 99.5%). The test panel did not, however, contain specimens from patients with conditions known to produce cross-reacting antibodies (e.g., syphilis) or polyclonal B-cell activation (e.g., Epstein-Barr virus infection or systemic lupus erythematosus). Thus, reported specificities in this study are likely higher than they would have been if cross-reactive specimens were included in the evaluation. Subsequent studies that included cross-reactive sera demonstrated that Western blotting correctly identifies many false-positive ELISA reactions (23, 24).

This study confirmed in the reference and research laboratory setting the previously documented problems with accuracy and precision of serodiagnostic tests by using WCS antigens of *B. burgdorferi* (4-11). The study confirmed that a serious disparity existed between the test results obtained by CDC and those obtained by academic reference centers with the best testing performances. These results guided corrective action and led to the adoption by CDC and ASTPHLD of a two-test approach to serodiagnosis (23), which forms the basis for the future national standardization of Lyme disease serologic testing methods.

Robert B. Craven,* Thomas J. Quan,* Raymond E. Bailey,* Raymond Dattwyler,[†] Raymond W. Ryan,[‡] Leonard H. Sigal,[§] Allen C. Steere,[¶] Bradley Sullivan,[#] Barbara J.B. Johnson,* David T. Dennis,* and Duane J. Gubler*

*Centers for Disease Control and Prevention, Fort Collins, Colorado, USA; [†]State University of New York at Stony Brook, Stony Brook, New York, USA; [‡]University of Connecticut Health Center, Farmington, Connecticut, USA; [§]University of Medicine and Dentistry of New Jersey–Robert Wood Johnson Medical School, New Brunswick, New Jersey, USA; [¶]Tufts/New England Medical Center, Boston, Massachusetts, USA; [#]Marshfield Clinic, Marshfield, Wisconsin, USA

A portion of this information was presented at the VIth International Conference on Lyme Borreliosis, Bologna, Italy, June 1994 and at the Second National Conference on Serologic Diagnosis of Lyme Disease, Dearborn, Michigan, October 1994.

References

1. Steere, AC. Lyme disease. N Engl J Med 1989;321:586-96.
2. Logigian EL, Kaplan RF, Steere AC. Chronic neurologic manifestations of Lyme disease. N Engl J Med 1990;323:1438-44.
3. Steere AC, Bartenhagen NH, Craft JE, et al. The early clinical manifestations of Lyme disease. Ann Intern Med 1989;99:76-82.
4. Bakken LL, Case KL, Callister SM, Bourdeau NJ, Schell RF. Performance of 45 laboratories participating in a proficiency testing program for Lyme disease serology. JAMA 1992;268:891-5.
5. Hedberg CW, Osterholm MT, MacDonald KL, White KE. An interlaboratory study of antibody to *Borrelia burgdorferi* . J Infect Dis 1987;155:1325-7.
6. Hedberg CW, Osterholm MT. Serologic tests for antibody to *Borrelia burgdorferi* —another Pandora's box for medicine? Arch Intern Med 1990;150:732-3.
7. Jones JM. Serodiagnosis of Lyme disease. Ann Intern Med 1991;114:1064.
8. Lane RS, Lennette ET, Madigan JE. Interlaboratory and intralaboratory comparisons of indirect immunofluorescence assays for serodiagnosis of Lyme disease. J Clin Microbiol 1990;28:1774-9.
9. Luger SW, Krauss E. Serologic tests for Lyme disease: interlaboratory variability. Arch Intern Med 1990;15:761-3.
10. Schwartz BS, Goldstein MD, Ribeiro JMC, Schulze TL, Shahied SI. Antibody testing in Lyme disease: a comparison of results in four laboratories. JAMA 1989;262:3431-4.
11. Quan TJ, Wilmoth BA, Carter LG, Bailey RE. A comparison of some commercially available serodiagnostic kits for Lyme disease. In: Proceedings of the First National Conference on Lyme Disease Testing (Dearborn, Michigan). Washington, DC: Association of State and Territorial Public Health Laboratory Directors, 1991:61-73.
12. Steere AC, Grodzicki RL, Kornblatt AN, et al. The spirochetal etiology of Lyme disease. N Engl J Med 1983;308:733-40.
13. Benach JL, Bosler EM, Hanrahan JP, et al. Spirochetes isolated from the blood of two patients with Lyme disease. N Engl J Med 1983;308:740-2.
14. Berger BW, Johnson RC, Kodner C, Coleman L. Cultivation of *Borrelia burgdorferi* from erythema migrans lesions and perilesional skin. J Clin Microbiol 1992;30:359-61.
15. Centers for Disease Control and Prevention. Case definitions for public health surveillance. MMWR 1990;39(RR-13):19-21.
16. Russell H, Sampson JS, Schmid GP, Wilkinson HW, Plikaytis B. Enzyme-linked immunosorbent assay and indirect immun ofluorescence assay for Lyme disease. J Infect Dis 1984;149:465-70.
17. Hansen K, Åsbrink E. Serodiagnosis of erythema migrans and acrodermatitis chronica atrophicans by the *Borrelia burgdorferi* flagellum enzyme-linked immunosorbent assay. J Clin Micrbiol 1989;27:545-51.
18. Craft JE, Grodzicki RL, Steere AC. Antibody response in Lyme disease: evaluation of tests. J Infect Dis 1984;149:789-95.
19. Grodzicki RL, Steere AC. Comparison of immunoblotting and indirect enzyme-linked immunosorbent assay using different antigen preparations for diagnosing early Lyme disease. J Infect Dis 1988;157:790-97.

20. Dressler F, Whalen JA, Reinhardt BN, Steere AC. Western blotting in the serodiagnosis of Lyme disease. J Infect Dis 1993;167:392-400.
21. Fister RD, Weymouth LA, McLaughlin JC, Ryan RW, Tilton RC. Comparative evaluation of three products for the detection of *Borrelia burgdorferi* antibody in human serum. J Clin Microbiol 1989;27:2834-37.
22. Dattwyler RJ, Volkman DJ, Luft BJ, Halperin JJ, Thomas J, Golightly MG. Seronegative Lyme disease: dissociation of specific T- and B-lymphocytic response to *Borrelia burgdorferi* . N Engl J Med 1988;319:1441-6.
23. Association of State and Territorial Public Health Laboratory Directors and the Centers for Disease Control and Prevention. Recommendations. In: Proceedings of the Second National Conference on Serologic Diagnosis of Lyme Disease (Dearborn, Michigan). Washington, DC: Association of State and Territorial Public Health Laboratory Directors 1995:1-5.
24. Johnson BJB, Robbins KE, Bailey RE, et al. Serodiagnosis of Lyme disease: accuracy of a two-step approach using a flagella-based ELISA and immunoblotting. J Infect Dis 1995; submitted.