Statement for IDSA Lyme Disease Review Panel
April 13, 2009

Joel Spinhirne
The Serano Group, Inc.
1767 12th Street #361
Hood River, OR 97031

Phone (866) 440-8204
Fax (815) 642-0056
Email: jspinhirne@seranogroup.org

Background: President of the Serano Group, Inc., a non-profit corporation analyzing and researching health issues. President of Verim Research, a firm providing independent analysis of scientific and technical issues.

This comment elaborates and supports comments submitted by Dr. David Volkman for the guideline review committee. This statement is made without communication with or endorsement by Dr. Volkman.

Issues: Unwarranted reliance on serological tests for diagnosis of Lyme borreliosis

Manipulation of literature to support guideline dogma

Statement

Dr. Volkman has provided the panel evidence that the IDSA Guidelines Committee for Lyme disease conform to dogmatic opinion and ideologies, often quoting the authors' self-generated literature that ignores available science. A portion of his statement emphasizes the committee's unwarranted reliance on serological tests and their dismissal of undeniably established cases of seronegative Lyme borreliosis. He also recounts conflicting financial interests of the guideline authors and their institutions, primarily in revenues generated by in the manufacture and sale of serological-based test kits.

I wish to present evidence that at least one of the guideline authors published a study using questionable methodology to support committee opinion that serological tests are a predominant component of Lyme borreliosis diagnosis, while minimizing the specificity and importance of PCR testing for *Borrelia burgdorferi*. The study escaped appropriate correction or rejection by journal peer-review. As recounted in Dr. Volkman's statement, journals have been unwilling to correct the IDSA committee authors' errors or challenge their unsupported statements.

One of the guideline authors, J. Stephen Dumler, was the lead investigator of a 2005 study evaluating Lyme disease tests. I challenged his invalid methodology and

unsupported findings in a letter to the *Journal of Clinical Microbiology*. Instead of addressing the issues of the flawed methodology, Dumler used his response to promote dogma, the 2000 version of the IDSA Lyme disease guidelines, and to attack an independent lab.

The basic flaw of the study was not acknowledged: the study defined true-positives using the same test being evaluated, greatly inflating statements of test sensitivity. His primary analysis had no correlation to any real-world phenomena other than a positive test result. His reply did clarify one area of obfuscation: leaving aside Dumler's invalid definition of true-positives and true-negatives, when presenting his results Dumler lumped together and did not differentiate poorly-defined true-positives and poorly-defined true-negatives in his results.

Following his reply, I responded with the following letter to the Journal of Clinical Microbiology:

EVALUATION OF LYME DISEASE TESTS

I would like to thank Dr. Dumler for his reply to my letter expressing concerns regarding the Coulter et. al. study evaluating Lyme disease testing (4). Although he raises many subjective issues, I would prefer to focus on the study in question and the authors' presentation of results. I am uncertain why he states it was not the study's intention to determine sensitivity of Lyme disease tests when sensitivity figures are stated four times in the abstract alone.

This is not an issue of overdiagnosis or underdiagnosis; instead, this is a basic issue of proper presentation of results and their analysis.

Dr. Dumler claims that no history or physical findings can define a true positive for Lyme disease, yet the accepted case definition for an infection with *Borrelia burgdorferi* is the Centers for Disease Control definition that relies heavily on history and physical findings. This case definition is designed to determine near certain cases of Lyme disease and is already biased toward missing clinically significant cases (1). Its intent is to identify only nearly-certain surveillance cases. Instead of using this established definition, it appears Dumler is claiming that his study's only recourse is to define true positives by using the same tests he is attempting to evaluate. Obviously, this methodology is invalid and produces a statistic that conveys little information. Sensitivity will always be 100%. Further clarification of the study is highly relevant at this time because of the inappropriate application of this study to support published clinical guidelines (6).

The underlying fallacy of using positive test results to define their own true positives is so apparent it hardly needs explanation. If one test had been evaluated using this methodology, sensitivity would obviously be 100%. Any

larger number of evaluated tests will always produce a sensitivity of 100% in one combination, and highly inflated sensitivities in other combinations. It is only mildly informative to know how many tests are required to reach the 100% sensitivity figure, regardless of the researchers' pre-observation subjective confidence in the tests.

Dumler's study provides an indication only of how well the evaluated tests agree with each other within their own universe of tests with positive results, completely disassociated from any real world correlations. A test being evaluated should never be used in any manner to determine its own true positive. The results will always be biased and vary only in degree of bias. When only six tests are evaluated, with some contributing few unique positives as in Dumler's study, bias is fairly extreme.

The sensitivity and selectivity of the evaluated tests have not been determined with any convincing level of confidence (2, 5). A properly conducted and presented study regarding testing for Lyme disease could contribute to our base of knowledge. Dumler's study, as presented, unfortunately, contributes little (except for making a weak case for perhaps not ordering plasma PCR tests) when it could have contributed much more. Very few studies validating Lyme tests have been conducted in the last ten years, particularly in the U.S.

If the authors had compared tests, such as serologies, to the accepted gold standards, positive culture or PCR (3), we would have an estimate of the specificities of these other tests. Of course, this does nothing to estimate sensitivity of any of the tests studied. At our current level of knowledge and technique, our estimates of sensitivity must depend on the accepted case definitions, instead of an artificial definition of infection based on the same tests we are attempting to evaluate.

Rather than produce highly biased, nearly meaningless statistics of test correlation I would encourage the study's authors to publish their study's results in terms of true positives, false positives, true negatives, and false negatives using the source data they summarized in their Table 1 where true positives and true negatives were lumped together to produce an "agreement" figure. If the message of their study is that Lyme disease tests have little clinical relevance, this is the fact that should be emphasized rather than the published statements of highly-biased sensitivity and selectivity figures.

As presented, the study does have one important message: if a physician required a positive test for diagnosis and only used the two-tier ELISA followed by Western Blot testing procedure (currently these are the only tests routinely ordered in clinical settings), 11 cases of certain infection would be entirely missed in the 86 subjects. Those 11 (13%) were positive by our present day gold standards, culture and PCR results, yet had negative

serologies. This statistic furthermore ignores all subjects not positive on any test, a number not stated in the study as presented, but quoted in the abstract as 49/86 (57%). If only 20% of these subjects (who were selected because they exhibited symptoms suggestive of Lyme disease) were truly infected, the physician relying on two-tier serology testing would miss 21/86 (24%) cases of a potentially disabling or fatal disease.

If the study's authors cannot bring themselves to use the established case definition of Lyme disease to define a true positive, they could perhaps use terms like positive-agreement, negative-agreement, and so on. The authors' audience wants to know the test results for the cohorts of probable, possibly, and unlikely infected subjects so readers can form their own opinions as to the validity and usefulness of the tests in question. The study's discussion scenario of a clinician requiring a positive test for diagnosis and searching for the most efficient set of tests to produce a possible positive result is irrelevant, especially when *B. burgdorferi* cultures are available only in research settings.

The authors are encouraged to carefully review their figures. Dumler's reply to my earlier letter states that the agreement figure of 32 subjects came from, '8 subjects initially seropositive with probable Lyme disease and also 43 initially seronegative assessed as NOT "probable" for Lyme disease.' The sum of 8 plus 43 does not equal 32.

Joel Spinhirne
Verim Research
Hood River, Oregon 97031

Phone: (815) 846-1642
Fax: (815) 642-0056
Email: joel@verimresearch.com

REFERENCES

1. 1997. Case definitions for infectious conditions under public health surveillance. Centers for Disease Control and Prevention. MMWR Recomm Rep **46:**1-55.
2. **Brown, S. L., S. L. Hansen, and J. J. Langone.** 1999. Role of serology in the diagnosis of Lyme disease. Jama **282:**62-6.
3. **Bunikis, J., and A. G. Barbour.** 2002. Laboratory testing for suspected Lyme disease. Med Clin North Am **86:**311-40.
4. **Coulter, P., C. Lema, D. Flayhart, A. S. Linhardt, J. N. Aucott, P. G. Auwaerter, and J. S. Dumler.** 2005. Two-year evaluation of Borrelia burgdorferi culture and supplemental tests for definitive diagnosis of Lyme disease. J Clin Microbiol **43:**5080-4.

5. **Mouritsen, C. L., C. T. Wittwer, C. M. Litwin, L. Yang, J. J. Weis, T. B. Martins, T. D. Jaskowski, and H. R. Hill.** 1996. Polymerase chain reaction detection of Lyme disease: correlation with clinical manifestations and serologic responses. Am J Clin Pathol **105:**647-54.
6. **Wormser, G. P., R. J. Dattwyler, E. D. Shapiro, J. J. Halperin, A. C. Steere, M. S. Klempner, P. J. Krause, J. S. Bakken, F. Strle, G. Stanek, L. Bockenstedt, D. Fish, J. S. Dumler, and R. B. Nadelman.** 2006. The clinical assessment, treatment, and prevention of lyme disease, human granulocytic anaplasmosis, and babesiosis: clinical practice guidelines by the Infectious Diseases Society of America. Clin Infect Dis **43:**1089-134.

Similar to Dr. Volkman's rejection of challenges to IDSA committee dogma, my letter was rejected by Andrew Onderdonk, editor of the *Journal of Microbiolog*y. He did not address Dumler's inability to present arithmetically valid results, indicating the current standards of journal publishing.

Dumler has refused numerous requests to provide the positive and negative results from his three defined cohorts, the basic measurement of his study. As stated in my letter, it would be of interest to know how many of his "probable" Lyme borreliosis cases produced positive test results and how many of the "unlikely" cases produced negative tests, but this has never been disclosed.

There is straightforward methodology to evaluate a test for an infectious disease. Subjects are classified as "infected" or "not infected" and then we record their test results. After that, we can analyze the results and discuss the limitations and strengths of the study.

Dumler chose not to do this. Instead, he obscured his results and manipulated the rest of his analysis to promote the IDSA committee's position that PCRs should not be performed in deference to serological tests, the same tests in which committee members' institutions have large financial interests. The sensitivity of serological tests was greatly inflated in the process.

The most prominent graphic in the paper repeatedly classified subjects as having "No evidence of Lyme disease". Dumler's study would classify a patient with a CDC case of Lyme disease with a known summer tick bite near Lyme, Connecticut, with a physician observed EM rash and flu-like symptoms as having "No evidence of Lyme disease", simply because the patient never produced a positive test result. This overwhelmingly illustrates the extent the IDSA guideline authors can and will go to support their unwarranted conclusions. The inability of peer-review or editorial responsibility to regulate is also amply illustrated by this case.

The Dumler study is only one incident, but it illustrates a common modus operandi of the IDSA guideline authors:

- Publishing studies in journals to support a foregone guideline conclusion using invalid techniques all the while proclaiming high objectivity

- Quoting their own studies in upcoming guidelines, as was done with this study in the 2006 version of the guidelines

- Repeating claims of objectivity by appealing to ignorance in areas not adequately investigated

- Supporting commercial products linked to the guideline authors (several institutions receive income from serology test kit manufacturers)

- Appealing to cultural biases of clinicians by attacking the objectivity of patients and supposed malingering

The irresponsibility of neglecting PCR positive cases of Lyme borreliosis, as recounted in my letter, is unacceptable. The financial burden of ignoring Lyme borreliosis cases that produce much harder to treat late-stage cases or undiagnosed cases is an enormous burden on our health care system.

The guideline errors raised by Dr. Volkman—denial of persistent infection, seronegative cases, inadequate treatments, and flawed diagnosis—can be shown to originate or receive support from deliberate manipulations the guideline authors. Not all of these manipulations are as blatant as Dumler's, but the means, motives, and opportunity is amply illustrated by this case.

I encourage the review panel to acknowledge the bias and manipulation of the IDSA guideline authors and to demand complete guideline revision by new, unbiased authors. Public health concerns overrule the ideologies and financial interests of the authors.

Joel Spinhirne
The Serano Group, Inc.

JOURNAL OF CLINICAL MICROBIOLOGY, Apr. 2006, p. 1616–1617 Vol. 44, No. 4
0095-1137/06/$08.00+0 doi:10.1128/JCM.44.4.1616–1617.2006

Evaluation of Tests for Lyme Disease

Coulter et al. are to be commended for their study evaluating the sensitivity of various tests for the diagnosis of Lyme disease (1). As the authors state, Lyme disease can be asymptomatic in its early stages and left untreated can develop chronic major manifestations. Studies evaluating available tests are quite important.

There are problems, however, with the data as presented. The combination of acute- and convalescent-phase serology with skin PCR is stated to have the highest sensitivity (100%), with serological testing with skin PCR almost as sensitive (92%). It would be easy for the reader miss the fact that these percentages were calculated using as a denominator the number of individuals who produced at least one positive test result from the set of tests being evaluated. These percentages would be useful if the intent of the study were to contribute to methodology for efficient selection of cases for future studies where high certainty of infection by *Borrelia burgdorferi*, the pathogen of Lyme disease, is required. The authors somewhat inappropriately expand the scope of their study when they comment the study supports the appropriateness of a published treatment guideline (3). Their definition of "sensitivity" becomes an issue.

The treating physician is likely to interpret test "sensitivity" in its broader sense: "How likely is a test to produce a positive result in an infected individual?" Using as their denominator the number of individuals in the study sample who produced at least one positive test result produces a result tangential to this question. The cohort producing at least one positive test result is quite different from the population the practitioner wants characterized: individuals infected with *B. burgdorferi*.

The dangers of misinterpretations of sensitivity statistics are obvious. Overstating sensitivities based on agreement with other tests, rather than characterization using known or highly probable infected subjects, encourages overconfidence for the treating physician in the tests being evaluated. Although the study attempts to present data from the cohort determined to be probable cases based on symptoms, there seems to be confusion in data presentation. For example, although 25 subjects were classified as probable for Lyme disease, the data show 32 subjects in this group positive on their initial serology. There are several other problems of this nature. Test results from subjects prospectively deemed unlikely to have infection were apparently not reported.

The authors are encouraged to clarify their data and emphasize their definition of sensitivity so that treating physicians do not misinterpret their findings and subsequently fail to diagnose patients with *B. burgdorferi* infection. Lack of prompt antimicrobial treatment for Lyme disease can result in a case of severe morbidity highly resistant to treatment (2). The sensitivity of our available tests should not be overestimated.

REFERENCES

1. **Coulter, P., C. Lema, D. Flayhart, A. S. Linhardt, J. N. Aucott, P. G. Auwaerter, and J. S. Dumler.** 2005. Two-year evaluation of Borrelia burgdorferi culture and supplemental tests for definitive diagnosis of Lyme disease. J. Clin. Microbiol. **43:**5080–5084.
2. **Oksi, J., M. Marjamaki, J. Nikoskelainen, and M. K. Viljanen.** 1999. Borrelia burgdorferi detected by culture and PCR in clinical relapse of disseminated Lyme borreliosis. Ann. Med. **31:**225–232.
3. **Wormser, G. P., R. B. Nadelman, R. J. Dattwyler, D. T. Dennis, E. D. Shapiro, A. C. Steere, T. J. Rush, D. W. Rahn, P. K. Coyle, D. H. Persing, D. Fish, B. J. Luft, et al.** 2000. Practice guidelines for the treatment of Lyme disease. The Infectious Diseases Society of America. Clin. Infect. Dis. **31**(Suppl. 1)**:**1–14.

Joel Spinhirne
Verim Research
Hood River, OR 97031

Phone: (503) 715-3174
Fax: (815) 642-0056
E-mail: joel@verimresearch.com

Author's Reply

I thank Dr. Spinhirne for his comments and recognition of the importance of objective, evidence-based evaluations of Lyme disease diagnostics (3). His comments seem to focus on "ambiguous" definitions for sensitivity that obscure precise infection detection. This is puzzling since the accepted definition for sensitivity is very simple: sensitivity = (number of test positives/number of true positives) × 100.

The implication is that the denominator used in our study imprecisely estimates a true-positive Lyme disease population. However, the definition of such a population is key, and there have been significant flaws in many investigations (1, 4; IGeneX, Inc. website [http://www.igenex.com/lymeset4.htm]; accessed 29 January 2006). Everyone agrees that objective findings must be used for scientific studies. Thus, no definition of Lyme disease can rest solely upon history or physical findings owing to inherent nonspecificity. We applied carefully researched, validated laboratory tests scrutinized by peer review and reproduced by other laboratories (3, 5) to a population for which Lyme disease was considered possible or probable for the majority. Using this approach, most would agree that identification of the bacterium by culture or PCR or the demonstration of a clear serological reaction provides an objective identification of infection. Owing to the uncertainty of clinical assessment, it seems very reasonable that this cohort is the most objectively defined "true-positive" population.

It is also important to avoid confusing agreement with sensitivity. Table 1 in our study revealed 32 subjects for whom initial serologic results and clinical assessments agreed, including 8 subjects initially seropositive with probable Lyme disease and also 43 initially seronegative assessed as NOT "probable" for Lyme disease. The intent was not to determine sensitivity but to illustrate the poor agreement between clinical and laboratory assessments, supporting our approach and the recommendations of the Infectious Diseases Society of America that call for maximizing positive and negative predictive value by integrating established clinical and laboratory studies (5). Our data also underscore the absolute requirement for an objective "gold standard." For example, when the "gold standard" cohort includes those for whom any laboratory test was reactive AND patients who were suspected to have Lyme disease, the diagnostic sensitivity of all tests and combinations is lower. With the ever-broadening clinical criteria used by some to define Lyme disease (2), the sensitivity of testing could approach zero, where no objective criterion would be helpful and a diagnosis of Lyme disease could be supported by any subjective finding deemed suitable. The unfortunate outcome is "shopping" for a laboratory result that conforms to the clinical impression but is more likely to be false positive among tests with limited specificity.

Inevitably, health care practitioners must understand algorithms for laboratory confirmation with known levels of confidence. We generated data to provide sensitivity information based strictly upon laboratory investigations to provide a framework for evidence-based laboratory utilization. It is unfortunate that highly sensitive laboratory diagnostics for all phases of Lyme disease have not yet been developed. However, it would be a disservice to evidence-based medicine to misclassify patients by broadening gold standards to those derived from questionably objective clinical manifestations, subjective "accumulated experience," or well-intentioned but unsupported opinions.

REFERENCES

1. **Cairns, V., and J. Godwin.** 2005. Post-Lyme borreliosis syndrome: a meta-analysis of reported symptoms. Int. J. Epidemiol. **34:**1340–1345.
2. **Cameron, D., A. Gaito, N. Harris, G. Bach, S. Bellovin, K. Bock, S. Bock, J. Burrascano, C. Dickey, R. Horowitz, S. Phillips, L. Meer-Scherrer, B. Raxlen, V. Sherr, H. Smith, P. Smith, R. Stricker, et al.** 2004. Evidence-based guidelines for the management of Lyme disease. Expert Rev. Anti. Infect. Ther. **2:**S1–S13.
3. **Coulter, P., C. Lema, D. Flayhart, A. S. Linhardt, J. N. Aucott, P. G. Auwaerter, and J. S. Dumler.** 2005. Two-year evaluation of *Borrelia burgdorferi* culture and supplemental tests for definitive diagnosis of Lyme disease. J. Clin. Microbiol. **43:**5080–5084.
4. **Stricker, R. B., A. Gaito, N. S. Harris, and J. J. Burrascano.** 2004. Treatment of early Lyme disease. Ann. Intern. Med. **140:**577.
5. **Wormser, G. P., R. B. Nadelman, R. J. Dattwyler, D. T. Dennis, E. D. Shapiro, A. C. Steere, T. J. Rush, D. W. Rahn, P. K. Coyle, D. H. Persing, D. Fish, B. J. Luft, et al.** 2000. Practice guidelines for the treatment of Lyme disease. Clin. Infect. Dis. **31**(Suppl. 1)**:**1–14.

J. Stephen Dumler
Division of Medical Microbiology
Department of Pathology
The Johns Hopkins University School of Medicine
720 Rutland Ave.
Ross 624
Baltimore, MD 21205

Phone: (410) 955-8654
Fax: (443) 287-3665
E-mail: sdumler@jhmi.edu