Computer-Assisted Pattern Recognition of Autoantibody Results

Steven R. Binder,1* Mark C. Genovese,2 Joan T. Merrill,3 Robert I. Morris,4 and Allan L. Metzger4

Bio-Rad Laboratories, Hercules, California 94547; Division of Immunology and Rheumatology, Stanford University, Stanford, California 94305; Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma 73104; and RDL Reference Laboratory, Los Angeles, California 90034

Received 29 July 2005/Returned for modification 6 September 2005/Accepted 5 October 2005

Immunofluorescence-based anti-nuclear antibody (ANA) screens are increasingly used in the initial evaluation of autoimmune disorders, but these tests offer no “pattern information” comparable to the information from indirect fluorescence assay-based screens. Thus, there is no indication of “next steps” when a positive result is obtained. To improve the utility of immunofluorescence-based ANA screening, we evaluated a new method that combines a multiplex immunofluorescence assay with a k nearest neighbor (kNN) algorithm for computer-assisted pattern recognition. We assembled a training set, consisting of 1,152 sera from patients with various rheumatic diseases and nondiseased patients. The clinical sensitivity and specificity of the multiplex method and algorithm were evaluated with a test set that consisted of 173 sera collected at a rheumatology clinic from patients diagnosed by using standard criteria, as well as 152 age- and sex-matched sera from presumably healthy individuals (sera collected at a blood bank). The test set was also evaluated with a HEp-2 cell-based enzyme-linked immunosorbent assay (ELISA). Both the ELISA and multiplex immunofluorescence assay results were positive for 94% of the systemic lupus erythematosus (SLE) patients. The kNN algorithm correctly proposed an SLE pattern for 84% of the antibody-positive SLE patients. For patients with no connective tissue disease, the multiplex method found fewer positive results than the ELISA screen, and no disease was proposed by the kNN algorithm for most of these patients. In conclusion, the automated algorithm could identify SLE patterns and may be useful in the identification of patients who would benefit from early referral to a specialist, as well as patients who do not require further evaluation.

Screening for anti-nuclear antibodies (ANA) has been performed for many years to identify patients with autoimmune diseases. The strong association of ANA with systemic lupus erythematosus (SLE) is well established, and this finding satisfies 1 of the 11 American College of Rheumatology (ACR) criteria available for establishing the diagnosis. However, the high sensitivity of the screen is not accompanied by a strong positive predictive value. In a study of 1,010 patients in which the estimated sensitivity of the ANA test for SLE was 100%, the positive predictive value was 11% for SLE and 11% for other rheumatic diseases (30). An evidence-based meta-analysis established that the ANA screen had an excellent negative likelihood ratio (0.11), but the positive likelihood ratio for SLE was only 2.2, and the positive likelihood ratio for other major rheumatic diseases was below 2.0, indicating that the screen was not useful (32).

Almost all of the literature that addresses the clinical performance of the ANA screen is based on the detection of antibodies by immunofluorescence (indirect fluorescence assay [IFA]), which requires microscopic examination of positive results at multiple dilutions. Recently, there has been a recommendation to increase the cutoff for reporting a positive test for the ANA so that the false-positive rate is roughly 5% in the nondiseased population (17). Nevertheless, the 1:40 cutoff is still widely used, in order to maintain a high level of sensitivity. Even higher cutoffs have been recommended for patients under 18 years of age (23).

Over the past 10 years, enzyme-linked immunosorbent assays (ELISA) have been introduced to identify patients with ANA (37). These tests can be automated, and they do not require highly trained operators who can recognize microscopic patterns. However, these newer tests yield only numeric results and cannot produce pattern information, which has traditionally aided in the interpretation of IFA results.

Recently, methods for the simultaneous measurement of multiple anti-nuclear antibodies using arrays have been described (7, 14, 19, 25, 26). These methods show a more selective response than traditional tissue-based screens, since a limited set of antigens must be defined. They offer the advantages of complete automation, consistent performance, and more precise measurement of antibody levels. Further, these methods provide numeric results for multiple antibodies, providing an opportunity to identify antibody patterns associated with SLE using a computer-assisted algorithm.

We describe here the use of an array method to identify autoantibodies in sera from a large cohort of diseased and nondiseased individuals. These data were then used to develop a k nearest neighbor (kNN) (pattern recognition) algorithm for the identification of antibody patterns associated with SLE, antibody patterns associated with other connective tissue diseases, and patterns associated with the absence of rheumatic disease.

MATERIALS AND METHODS

Study objective and design. The objective of this study was to compare three methods for screening sera to detect connective tissue disorders (CTDs), and specifically to detect SLE. The three methods we examined were (i) an ELISA method, (ii) a multiplex autoimmunity assay, and (iii) a pattern recognition algorithm operating on the multiplex assay results.
We acquired two sets of sera with known clinical diagnoses from a number of sources. We used the first set of sera to train the pattern detection algorithm, and we used the second set of sera to evaluate the performance of all three methods. Our endpoint was the positive and negative likelihood ratios for the different methods. We describe the diagnostic methods, sera, and statistical methods in the following sections.

**Diagnostic methods.** The ELISA method was the Helix ANA screen (West Sacramento, CA), which utilizes a mixture of Hep-2 cell extract and several supplementary antigens as the antigenic target (15). For this method, we reported sera with antibody index (AI) values of 1.0 or higher as positive.

The multiplex autoimmune assay method was the BioPlex 2200 ANA screen (Bio-Rad Laboratories, Hercules, CA), which employs dyed magnetic beads to perform measurements of 13 autoimmune antibodies simultaneously in one tube (10, 16, 18, 28). Briefly, 13 different dyed beads were precoated with SSA-52, SSA-60, SSB, Sm, SmRNP, RNP-68, RNP-A, Jo-1, ribosomal protein, Scl-70 (topoisomerase), centromere B, chromatin, and double-stranded DNA (dsDNA). A series of recombinant antigens from Diarect (Freiberg, Germany) were used for SSA-52, RNP-68, RNP-A, Jo-1, Scl-70, and centromere B. Affinity-purified antigens from Arotec (New Zealand) were used for Sm and SSB. Affinity-purified SSA-60, SmRNP, and ribosomal protein were prepared at Helix (West Sacramento, CA). Chromatin and dsDNA were prepared at Bio-Rad Laboratories by standard methods (4–6). The bead mixture was combined with patient sample and diluent and allowed to incubate for 20 min at 37°C. After a wash cycle, anti-human immunoglobulin G antibody conjugated to B-phycoerythrin was added to the dyed beads and allowed to incubate for 10 min at 37°C. Following removal of excess conjugate, the mixture passed through the detector, which identified the bead type, based on the fluorescence of the dyed bead. The amount of antibody bound to the bead was determined by the fluorescence of B-phycoerythrin. Raw data were initially measured as relative fluorescence intensity and then converted to fluorescence intensity using a predyed internal standard bead. A series of calibrators was analyzed with the patient samples to convert the FR into international units per milliliter for anti-dsDNA and AI for the other antibodies measured. For anti-dsDNA, levels of 10 IU/ml or higher were positive; for the other analytes, AI values of 1.0 or higher were positive. For this method, we reported the screening result as positive when 1 or more of the 13 antibody results was positive. The specificity of this method was established in a recent study of 510 healthy subjects by multiple methods (28).

The pattern recognition method was a k nearest neighbor algorithm (39) operating on the numerical results from the multiplex autoimmune assay. This software was developed in-house by one of the authors (S. R. Binder); it is an optional component included with the instrument software. Other researchers have used the same pattern recognition method for many diagnostic purposes, such as identifying volatiles associated with infectious disease (41), tracking tumor growth rates (12), and detecting disease clustering using data from children with high blood lead levels (34). Another pattern recognition method, the artificial neural network, has been utilized for classification of giant cell arteritis (2), and for the prediction of lupus nephritis in patients with SLE (24). Before the kNN algorithm is used for prediction, it must be “trained” using a number of sera from patients with known diagnoses. We used 1,152 samples (Table 1) for this purpose. When the trained algorithm is presented with a new serum sample, it compares the characteristics of the new sample to previously encountered samples with similar characteristics to predict the diagnosis of the unknown sample. The characteristics of the sample were the log-transformed FR (AI) values for the 13 autoimmune antibodies examined by the multiplex assay. The previously encountered samples, which are considered when assigning a pattern to a new serum sample, are known as the k nearest neighbors. The algorithm was developed and tested using the “leave one out” strategy, where one sample is removed from the training set and then tested as “an unknown” against the remaining samples (39). Various parameters required for successful performance were chosen with the assistance of a commercial optimization algorithm (ECHIP, Hookes, DE). Only neighbors showing at least 30% cumulative agreement were considered (neighbors based on geometric distance [from the new sample to the neighbor for each antibody]). We selected 11 as the value for k and 30% as the value for cumulative agreement. The final step in the prediction process is to assign a pattern to the new serum sample based on the diagnoses of its k nearest neighbors. This was accomplished by accepting diagnoses that were found in at least 25% (at least three) of the neighbors. The kNN algorithm can identify patterns associated with the following connective tissue diseases (“targeted CTDs”): mixed connective tissue disease (MCTD), SLE, Sjögren’s syndrome, scleroderma, and polymyositis. In addition, it can report that no association could be established (insufficient neighbors or no consistent type of neighbors). If a new serum sample was associated with more than one disease, we reported all associations that met our agreement criterion.

### Serum sample selection

We collected 1,152 sera to use as a “training set” for the pattern recognition algorithm. As the predictive ability of the pattern recognition algorithm is related to the diversity of the samples in the training set, we collected sera from a number of institutions. All sera were identified by clinical diagnoses, and all other identifying information was stripped from the samples prior to inclusion in the study. Two authors (R. I. Morris and A. L. Metzger) collected sera (n = 570) from individuals seen in their private rheumatology practice between 1996 and 2002. Additional sera (n = 381), collected between 2000 and 2002, were obtained from The Registry for the Antiphospholipid Syndrome, which has its principal site at the Oklahoma Medical Research Foundation. These samples included autoimmune sera from individuals from different regions in the United States that were obtained, after informed consent, for use in a collaborative, blinded repository and deidentified clinical database. Both of these collections contained patients with SLE, Sjögren’s syndrome, myositis, scleroderma, rheumatoid arthritis, and MCTD. In addition, both collections included sera from patients with osteoarthritis, fibromyalgia, antiphospholipid syndrome, and vasculitis, as well as sera from individuals with no known disease. The diagnoses of SLE and rheumatoid arthritis were established according to the ACR criteria (1, 13, 35). Because polymyositis and scleroderma are relatively rare diseases and not all patients with these diseases produce the antibodies of interest (Jo-1 and Scl-70, respectively), additional sera (n = 31) from patients with these conditions were acquired from the Oklahoma Medical Research Foundation and from SLR Research (San Diego, CA). Sera (n = 24) from patients with juvenile-onset diabetes were acquired from Intergen (Purchase, NY). Finally, since autoantibodies are commonly encountered in individuals without any known autoimmune disease, we included samples from patients with positive ANA screens (IFA titer of 1:40 or greater) but without any known autoimmune disease (n = 146), obtained from the Foundation for Blood Research (Scarborough, ME).

To test the performance of the diagnostic methods, we used a separate collection of 325 sera; these are referred to collectively as the “test set.” One author (M. C. Genovesi) collected 173 sera as part of a prospective rheumatologic study conducted at Stanford University. Consecutive patients seen in a rheumatology clinic were enrolled in this study after informed consent was obtained. The protocol for this study was approved by the institutional review board at the Stanford University School of Medicine. The diagnoses of SLE, rheumatoid

### Table 1: Comparison of the diagnoses of individuals who provided sera in the training set and test set

<table>
<thead>
<tr>
<th>Physician’s diagnosis</th>
<th>SLE</th>
<th>Sjögren’s syndrome</th>
<th>MCTD</th>
<th>Polymyositis</th>
<th>Scleroderma</th>
<th>Myositis</th>
<th>SLE</th>
<th>Sjögren’s syndrome</th>
<th>MCTD</th>
<th>Polymyositis</th>
<th>Scleroderma</th>
</tr>
</thead>
<tbody>
<tr>
<td>No CTD</td>
<td>458 (39.8)</td>
<td>21 (1.8)</td>
<td>19 (1.6)</td>
<td>46 (4.0)</td>
<td>51 (4.4)</td>
<td>1354</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nontargeted CTD</td>
<td>253 (22.0)</td>
<td>90 (7.7)</td>
<td>41 (3.4)</td>
<td>21 (1.8)</td>
<td>45 (4.2)</td>
<td>1354</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Targeted CTD</td>
<td>441 (38.3)</td>
<td>90 (7.7)</td>
<td>41 (3.4)</td>
<td>21 (1.8)</td>
<td>45 (4.2)</td>
<td>1354</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2: Results for the test set (n = 325)*

<table>
<thead>
<tr>
<th>Physician’s diagnosis</th>
<th>Total no. of sera</th>
<th>No CTD</th>
<th>Nontargeted CTD</th>
<th>Targeted CTD</th>
<th>SLE</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELISA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiplex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kNN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No CTD</td>
<td>190</td>
<td>42</td>
<td>37</td>
<td>24</td>
<td>10</td>
<td>55</td>
</tr>
<tr>
<td>Nontargeted CTD</td>
<td>90</td>
<td>37</td>
<td>24</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Targeted CTD</td>
<td>45</td>
<td>42</td>
<td>41</td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLE</td>
<td>34</td>
<td>32</td>
<td>32</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>325</td>
<td>121</td>
<td>83</td>
<td>55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Positive results produced by ELISA compared to positive results produced by the multiplex method and to positive disease associations proposed by the kNN algorithm.
TABLE 3. Output of the kNN interpretative algorithm for the samples collected at a rheumatology clinic compared to rheumatologist-assigned diagnoses

<table>
<thead>
<tr>
<th>Physician’s diagnosis</th>
<th>Total no. of samples</th>
<th>No. of patients with positive antibody results</th>
<th>No. of samples with the following kNN output&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Total no. of output results</th>
</tr>
</thead>
<tbody>
<tr>
<td>No CTD</td>
<td>38</td>
<td>5</td>
<td>SLE 0 MCTD 0 Scleroderma 0 SS 0 PMYO 0 NA 5</td>
<td>5</td>
</tr>
<tr>
<td>Nontargeted CTD</td>
<td>90</td>
<td>24</td>
<td>SLE 8 MCTD 1 Scleroderma 1 SS 1 PMYO 14 NA 25</td>
<td>25</td>
</tr>
<tr>
<td>SLE</td>
<td>32</td>
<td>30</td>
<td>SLE 25 MCTD 3 Scleroderma 3 SS 0 PMYO 3 NA 37</td>
<td>37</td>
</tr>
<tr>
<td>SLE and MCTD</td>
<td>2</td>
<td>2</td>
<td>SLE 2 MCTD 1 Scleroderma 0 SS 0 PMYO 0 NA 3</td>
<td>3</td>
</tr>
<tr>
<td>MCTD</td>
<td>4</td>
<td>4</td>
<td>MCTD 4 Scleroderma 2 SS 1 PMYO 6 NA 6</td>
<td>6</td>
</tr>
<tr>
<td>Scleroderma</td>
<td>4</td>
<td>3</td>
<td>Scleroderma 0 SLE 0 SS 0 PMYO 0 NA 0</td>
<td>0</td>
</tr>
<tr>
<td>Sjögren’s syndrome</td>
<td>2</td>
<td>1</td>
<td>Sjögren’s syndrome 1 SLE 0 SS 0 PMYO 0 NA 2</td>
<td>2</td>
</tr>
<tr>
<td>Polymyositis</td>
<td>1</td>
<td>1</td>
<td>Polymyositis 0 SLE 0 SS 0 PMYO 0 NA 1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>173</td>
<td>70</td>
<td>Total 40 SLE 6 MCTD 7 Scleroderma 5 SS 2 PMYO 22 NA 82</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Note that the total number of output results from the kNN interpretative algorithm exceeds the number of patients with positive antibody results, because some samples were associated with more than one disease.

<sup>b</sup> SS, Sjögren’s syndrome; PMYO, polymyositis; NA, no association. The values in bold type show results where the kNN output and clinical assessment are in concordance.

RESULTS

The data produced by the test set were analyzed to determine the number of positive specimens obtained by the two immunological techniques and the software approach. The results of this analysis are presented in Table 2. Table 2 shows the number of positive specimens determined by the ELISA method and multiplex method; these data are compared to the number of specimens associated with a targeted CTD, based upon computer-assisted pattern recognition (kNN).

Table 3 shows the output of the kNN algorithm for 173 sera from the fully characterized patients enrolled in the study at the Stanford University Clinic. An appropriate kNN result was proposed for 53 of the 70 antibody-positive patients (75.7%), including 27 of the 32 patients (84.4%) with SLE who were positive for at least one antibody (two were antibody negative by both methods). No further statistical analysis was performed due to the small number of patients in each group.

Table 4 displays the performance characteristics of the three diagnostic methods for SLE and for all targeted connective tissue diseases. There is a significant difference in the sensitivity of the three methods for SLE (Q = 6.0; P = 0.049) and for targeted CTDs (Q = 7.6; P = 0.022). Table 4 shows that the sensitivity of all methods is greater for detecting SLE than for detecting the presence of any targeted CTD. The sensitivity of the ELISA method is greater than the sensitivity of the multiplex method for targeted CTDs, but examination of the confidence intervals reveals that this difference is not significant. The two methods differ chiefly in the use of a HEp-2 cell extract in the ELISA kit, which theoretically permits the detection of many other antibodies that will be missed by the limited number of antigens available in the multiplex method.

TABLE 4. Performance of three diagnostic methods in detecting two types of disease conditions<sup>a</sup>

<table>
<thead>
<tr>
<th>Test</th>
<th>Disease</th>
<th>Positive LR&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Negative LR</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELISA screen</td>
<td>TCTD&lt;sup&gt;c&lt;/sup&gt;</td>
<td>2.87 [2.38–3.45]</td>
<td>0.097 [0.032–0.290]</td>
<td>0.935 [0.811–0.983]</td>
<td>0.674 [0.615–0.728]</td>
</tr>
<tr>
<td></td>
<td>SLE</td>
<td>2.89 [2.39–3.48]</td>
<td>0.087 [0.023–0.336]</td>
<td>0.941 [0.789–0.990]</td>
<td>0.674 [0.615–0.728]</td>
</tr>
<tr>
<td>Multiplex</td>
<td>TCTD</td>
<td>5.92 [4.40–7.96]</td>
<td>0.128 [0.056–0.293]</td>
<td>0.891 [0.756–0.959]</td>
<td>0.849 [0.801–0.888]</td>
</tr>
<tr>
<td></td>
<td>SLE</td>
<td>6.25 [4.67–8.36]</td>
<td>0.069 [0.018–0.266]</td>
<td>0.941 [0.789–0.990]</td>
<td>0.849 [0.801–0.888]</td>
</tr>
<tr>
<td>kNN</td>
<td>TCTD</td>
<td>12.1 [7.70–19.1]</td>
<td>0.187 [0.099–0.351]</td>
<td>0.826 [0.680–0.917]</td>
<td>0.932 [0.894–0.957]</td>
</tr>
<tr>
<td></td>
<td>SLE</td>
<td>12.5 [7.94–19.8]</td>
<td>0.158 [0.070–0.355]</td>
<td>0.853 [0.682–0.945]</td>
<td>0.932 [0.894–0.957]</td>
</tr>
</tbody>
</table>

<sup>a</sup> Ninety-five percent confidence intervals are shown in brackets.

<sup>b</sup> LR, likelihood ratio.

<sup>c</sup> TCTD, targeted CTD.
In the test set reported, this benefit was not observed for any of the SLE patients and was only observed for one scleroderma sample. The kNN algorithm has lower sensitivity than both the ELISA and multiplex methods for detection of SLE and for detection of targeted CTDs, but these differences are also not statistically significant. However, it should be noted that the kNN algorithm is providing disease-specific information, unlike the other antibody screens.

It is more useful in practice to consider likelihood ratios than sensitivity and specificity, because likelihood ratios provide a better indication of the influence of a positive or negative test on clinical decision-making (11). Table 4 lists the positive likelihood ratios and 95 percent confidence intervals of each diagnostic method for SLE and for CTDs when the nondisease comparison group contains nontargeted CTDs. The positive likelihood ratios are consistently larger for the kNN method than for the multiplex method, and the positive likelihood ratios for the multiplex method are consistently greater than for the ELISA method. This implies that a positive result from the kNN algorithm is more informative in a clinical context than a positive result from either of the other two methods.

**DISCUSSION**

The data presented here support the ability of a $k$ nearest neighbor algorithm to classify antibody patterns on the basis of their resemblance to samples from patients with targeted connective tissue diseases or their lack of resemblance to such samples. The algorithm uses case-based reasoning, an algorithm that is easy to understand because it compares results from patients directly to results from other patients. Recently, case-based reasoning has been used to classify patients with early arthritis (40).

Computer-assisted diagnosis of rheumatic disorders has been evaluated for over four decades. A review published in 1991 identified 14 different software systems with applications to rheumatology, of which 7 were developed for general internal medicine and 7 were developed specifically for rheumatologic problems (20). Several of the rheumatology-focused systems have been further refined (3, 21, 27). All of these systems rely primarily on clinical symptoms. The approach described here has obvious limitations due to the use of serological data alone, and it is designed to support a diagnosis rather than supply one. It may benefit from the uniformity of data produced by automated instrumentation, since results often vary greatly between laboratories and between methods (9).

A comparison of the performance of IFA and a number of ELISA methods for ANA screening was recently presented (8). All of the ELISA methods used purified antigens in combination with HEP-2 cell extracts or recombinant antigens. Results were compared against diagnosis using high-prevalence and low-prevalence populations. The IFA was performed at a 1:160 dilution to minimize false-positive results in accordance with current recommendations (17); as a result, several ELISA methods were more sensitive than IFA. In the study reported here, we show that a multiplex method can produce results comparable to ELISA for disease state samples (Table 2). For patients without known CTDs, there were less positive results, and the use of the kNN algorithm as an additional screen further reduced the number of positive samples that might require follow-up testing and referrals.

The performance reported here for the multiplex screen relates specifically to the targeted diseases as defined above. Other diseases (such as rheumatoid arthritis, primary biliary cirrhosis, and dermatomyositis) may produce characteristic autoantibodies, but these disease-associated antibodies are not reportable as part of a defined ANA panel in the United States, due to FDA definitions. Further, an array-based technology cannot replace the use of HEP-2 cell-based ANA screens for autoimmune hepatitis or for monitoring patients with juvenile rheumatoid arthritis (32).

The test population reported here is small, and additional studies are in progress to evaluate the utility of pattern recognition in a larger prospective study. The samples that were collected at a blood bank were not from clinically characterized patients, and some of the antibody-positive sera may reflect the presence of actual disease. The blood bank study offers an estimate of specificity that can be readily repeated in other laboratories and in other countries (28).

**Conclusion.** The ANA screen has been previously used as a qualitative tool to satisfy one of the ACR criteria. The use of multiple antibody results, coupled with computer-generated pattern interpretation, demonstrates another option that could aid in the identification of patients who may benefit from an early referral to a specialist and in the identification of patients for whom a referral might not be needed.

**ACKNOWLEDGMENTS**

We thank Morris Reichlin, Oklahoma Medical Research Foundation, for providing specimens from patients with polymyositis; Kevin Bickford, SLR Inc., for providing specimens from patients with scleroderma; and Thomas Ladue (Foundation for Blood Research) for providing additional IFA-positive sera. We also acknowledge the contribution of John Glossenger (Bio-Rad Laboratories), who programmed software for the kNN algorithm. David Buckeridge, Stanford University, assisted with the statistical presentation of the data.

**REFERENCES**


mol. 44:763–770.