Accuracy of Anti–Ribosomal P Protein Antibody Testing for the Diagnosis of Neuropsychiatric Systemic Lupus Erythematosus

An International Meta-Analysis

Fotini B. Karassa,1 Antonella Afeltra,2 Ales Ambrozic,3 Deh-Ming Chang,4 Filip De Keyser,5 Andrea Doria,6 Mauro Galeazzi,7 Shunsei Hirohata,8 Ilse E. A. Hoffman,5 Murat Inanc,9 Loreto Massardo,10 Alessandro Mathieu,11 Chi Chiu Mok,12 Gabriella Morozzi,7 Giovanni Sanna,13 Alberto J. Spindler,14 Athanasios G. Tzioufas,15 Taku Yoshio,16 and John P. A. Ioannidis17

Objective. To quantitatively evaluate the diagnostic accuracy of antibodies to ribosomal P protein (anti-P) for neuropsychiatric systemic lupus erythematosus (NPSLE) in general, for psychosis, mood disorder, or both, and for other diffuse manifestations.

Methods. This international meta-analysis combined standardized data from 1,537 lupus patients contributed by 14 research teams. Weighted estimation of sensitivity and specificity with fixed-effects and random-effects models, as well as summary receiver operating characteristic (SROC) curve analysis, was used to summarize test performance. The robustness of the overall estimates was examined in sensitivity analyses that included additional studies published up to November 1, 2004 in the Medline, EMBase, and Cochrane databases.

Results. Combining the data from the 14 teams, the weighted sensitivity and specificity estimates for the diagnosis of NPSLE were 26% (95% confidence interval [95% CI] 15–42%) and 80% (95% CI 74–85%), respectively. For psychosis, mood disorder, or both, the sensitivity and specificity were 27% (95% CI 14–47%) and 80% (95% CI 74–85%), respectively. For other diffuse manifestations, the sensitivity was 24% (95% CI 12–42%), and the specificity was 80% (95% CI 73–85%). The proportion of patients with anti-P antibodies did not vary markedly across different presentations of NPSLE. Between-study heterogeneity was substantial, but the SROC curves were consistent with the weighted estimates. In further analyses that included another 24
published studies, only the sensitivity for psychosis and/or mood disorder was slightly improved, but it was still suboptimal (42% [95% CI 30–53%]); the specificity remained essentially the same (81% [95% CI 76–85%]).

**Conclusion.** Anti-P antibody testing has limited diagnostic value for NPSLE, and it is not helpful in differentiating among various disease phenotypes.

Neuropsychiatric manifestations occur in approximately one-half of patients with systemic lupus erythematosus (SLE) and may cause substantial impairment of quality of life as well as disability (1–3). Moreover, multiple neuropsychiatric events during the disease course are associated with adverse long-term prognosis (4,5) and may lead to death, with a mortality rate of 7–19% (2,5,6). Neuropsychiatric SLE (NPSLE) encompasses a multitude of symptoms involving the central, peripheral, and autonomic nervous systems as well as psychiatric disorders (7). Recently, an ad hoc committee of the American College of Rheumatology (ACR) proposed a standard nomenclature for 19 neuropsychiatric syndromes associated with SLE (7), yet NPSLE is difficult to diagnose and is challenging to treat. Secondary factors, such as drugs, metabolic abnormalities, or infections, can also cause neuropsychiatric disturbances in lupus patients (3,7). Manifestations reflecting diffuse cerebral involvement pose the foremost difficulty in differentiating their exact origin, since psychiatric disorders may merely be reactive psychological disturbances (2,3,7).

During the last 2 decades, several studies have explored the utility of antibodies to ribosomal P proteins (anti-P) in detecting NPSLE (6,8–35). These antibodies are directed toward 3 large-subunit ribosomal phosphoproteins, called P0 (38 kd), P1 (19 kd), and P2 (17 kd), which share a common linear determinant in the carboxyl-terminal 22-amino acid sequence (36). Early studies claimed that serum anti-P antibodies were highly accurate for the diagnosis of SLE-mediated psychosis and depression (9,26), but subsequent reports were less optimistic (11–13,18,20,25,27,31). Other studies expanded the spectrum of neuropsychiatric features that could be correlated with anti-P to include active disease, diffuse manifestations, or NPSLE overall (6,25,28,30), making even more unclear their clinical value for this entity. Methodologic shortcomings, including the criteria used to define NPSLE, the approaches adopted for detecting anti-P antibodies, and the small sample size of isolated studies, may have contributed to the uncertainty.

Because SLE is a relatively uncommon disease and NPSLE is even more uncommon, no single study can reliably assess the operating characteristics of anti-P antibodies. Yet, a rigorous appraisal of a diagnostic test may reduce the number of unwanted clinical consequences related to misleading estimates of the accuracy of that test. Ideally, one would like to assess the diagnostic accuracy of a test across a large study population and use similar, standardized, and reproducible methods. In the absence of a single very large study that could do this, an attractive alternative is to standardize data across existing cohorts of lupus patients. Therefore, the aim of this study was to evaluate the diagnostic performance of anti-P antibodies for NPSLE in general, for diffuse NPSLE manifestations, and for particular psychiatric syndromes (psychosis, mood disorder, or both) in the context of an international collaborative meta-analysis, with standardization of the data contributed by a large number of investigators.

**PATIENTS AND METHODS**

**Eligibility criteria.** The meta-analysis included lupus patients with and without NPSLE who had undergone serum anti-P antibody testing by immunoblotting, a standard enzyme-linked immunosorbent assay (ELISA), or both (37–39).

To ensure consistency, participating investigators were asked to comply with the following rules. Patients had to fulfill the ACR criteria for the classification of SLE (40) and had to be evaluated for the presence or absence of neuropsychiatric lupus syndromes according to the ACR nomenclature and case definitions (7). Patients with a neuropsychiatric syndrome during any time in the course of SLE were classified into 3 subgroups: those with psychosis, mood disorders, or both; those with other diffuse (2,6) manifestations (including acute confusional state, generalized seizures, cognitive dysfunction, anxiety disorder, and headache other than migraine or cluster headache), and those with focal (2,6) neurologic events (including cerebrovascular disease, partial seizures, migraine or cluster headache, myelopathy, demyelinating syndrome, movement disorder, aseptic meningitis, and syndromes of the peripheral nervous system) (7). When both diffuse and focal events occurred in the same patient, the designation was made according to the predominant manifestation. Severe, sustained, or progressive presentations requiring more-aggressive
treatment with cytotoxic immunosuppressive agents were considered to be predominant.

Collaborating investigators provided a clear description of the immunoassay(s) used for anti-P determination, with sufficient detail to permit replication (41). When both immunoblotting and ELISA had been used, data were reported separately for each method. Patients who had undergone testing for anti-P multiple times were considered to have this autoantibody specificity if at least 1 of the determinations yielded positive results. Investigators were also asked to specify whether immunoassays were performed without knowledge of the clinical condition of the patients and whether the diagnosis of NPSLE, as well as the assignment of neuropsychiatric syndromes, was accomplished without knowledge of the anti-P status of the participants.

**Organization of the international database.** Research teams who have previously published data on cohorts of SLE patients were invited to participate in this meta-analysis, provided that the study patients met the eligibility criteria defined above. Collaborating teams were identified through searches of the Medline, EMBASE, and Cochrane databases conducted in January 2003, using combinations of index terms (systemic lupus erythematosus, rheumatic diseases, connective tissue disease, or autoimmune disease, as well as ribosomal, antiribosomal, anti-P, or antineuronal), cited references of eligible studies and review articles, abstracts of major rheumatology conferences, and consultation with experts in the field. We e-mailed invitations to investigators working on SLE. The meta-analysis was also announced at an autoimmune disease–related scientific meeting (42). Pertinent data were contributed on a standard reporting form. The database remained open until July 2004.

Research teams from 14 centers (8 European, 4 Asian, and 2 South American) agreed to participate. We accepted data that were already available as well as data that were prospectively generated specifically by some of the participating teams for the purposes of the collaborative project. The effort was coordinated by the Clinical and Molecular Epidemiology Unit of the Department of Hygiene and Epidemiology at the University of Ioannina School of Medicine. The coordinating center was responsible for giving instructions to participating investigators on how to standardize and summarize their individual-level databases. The contributed data sets were assessed for potential errors or inconsistencies and then assembled at the coordinating center, which was also responsible for conducting the analyses. Queries were clarified through communications with the participating investigators.

**Data synthesis and statistical analysis.** Measures of diagnostic performance included sensitivity and specificity of anti-P antibodies for various forms of NPSLE. The main analysis involved the following 4 comparisons: NPSLE overall and each subgroup of NPSLE (psychosis and/or mood disorder, other diffuse manifestations, and focal events) versus the non-NPSLE group; all diffuse manifestations versus focal events; and psychosis and/or mood disorder versus other diffuse manifestations. These analyses address the discriminatory ability of the test for NPSLE in general, for each disease subtype, and for different neuropsychiatric presentations. To further pursue the possibility that anti-P may be specifically associated with particular psychiatric disorders (8,9,16,22,26), we evaluated the diagnostic accuracy of anti-P antibody for patients with psychosis and/or mood disorder versus all other lupus patients.

Test performance was estimated separately from studies that used immunoblotting for the detection of anti-P antibodies and from studies that used ELISA. In the overall analysis, when both immunoblotting and ELISA data were available from the same study, the results from the ELISA were used for the calculations. Diagnostic accuracy was also evaluated for subgroups defined by race.

Summary estimates were obtained with 2 meta-analytic methods: weighted independent estimation of sensitivity and specificity, and summary receiver operating characteristic (SROC) curve analysis.

Sensitivity and specificity estimates for each comparison were independently combined across studies, using both fixed-effects (Mantel-Haenszel) and random-effects (DerSimonian-Laird) models (43,44). Fixed-effects models weigh each study by the inverse of its variance. Random-effects models also incorporate between-study variation. The random-effects approach tends to provide wider confidence intervals (CIs) and is preferable in the presence of between-study heterogeneity. Except where indicated otherwise, random-effects estimates are provided below. Between-study heterogeneity was examined with Fisher’s exact test.

Because sensitivity and specificity are interdependent, independent weighting may sometimes underestimate both measures. Hence, we used SROC curve analysis to account for this mutual dependence (45,46). The method fits a curve describing the tradeoff between sensitivity and specificity across studies, with different characteristics and thresholds for an abnormal test result. The regression is calculated as follows: $D = \alpha + \beta S$, where $D$ is the difference in the logits of the true-positive rate (sensitivity) and the false-positive rate (1 − specificity), and $S$ is the sum of these logits. When $\beta$ is not significantly different from 0, the SROC curve is symmetric around the diagonal that runs from the top left corner to the bottom right corner of the diagram. Conversely, when $\beta$ is significantly different from 0, the SROC curve is not symmetric, and the overall diagnostic performance varies in different parts of the curve, with an uneven tradeoff between sensitivity and specificity across studies. This may indicate significant between-study variation in the selected test threshold, study population, or other parameters. SROC curves should not be extrapolated outside the range of observed values. Both nonweighted and weighted SROC curves were estimated (46,47); nonweighted curves consider all studies equally in the calculations, whereas weighted curves weigh each study by the variance of $D$.

**Inclusion of other published data.** Sensitivity analyses were conducted to examine whether the addition of further relevant published studies affected our summary estimates of the operating characteristics of anti-P antibodies. Only the following 2 comparisons were examined, since articles focused on these patient groups: the entire group of NPSLE
patients versus the non-NPSLE patient group, and patients with psychosis and/or mood disorder versus either the non-NPSLE patients or all other lupus patients. Finally, we evaluated the available data to compare active NPSLE versus non-NPSLE.

Eligible studies published in any language were retrieved during the stage of identification of pertinent articles and collaborating investigators, as described above. We updated the literature search of the 3 computerized databases in November 2004 to identify additional relevant studies published up to November 1, 2004. Meeting abstracts were not included because the results may not be final and may not have been subjected to formal peer review. Duplicate or overlapping data were counted only once. The inclusion criteria were similar to those of the collaborative meta-analysis, with no restriction on patient age or study location. Nevertheless, in these analyses, we did not use the stringent criteria regarding the method of antibody determination and classification of neuropsychiatric disease; studies were combined regardless of the assay used to detect anti-P antibodies and regardless of the criteria used to diagnose NPSLE.

Other sensitivity analyses. We also performed sensitivity analyses to assess the robustness of the quantitative estimates derived from the collaborative meta-analysis. These analyses were limited to studies that used the ACR criteria for NPSLE syndromes and limited to studies that specified blinding.

Software. Analyses were conducted with the use of the following software: SPSS, version 12.0 (SPSS, Chicago, IL), Meta-Test, version 0.6, New England Medical Center, Boston, MA, 1997 (Joseph Lau, Tufts–New England Medical Center, Boston, MA) and Meta-Analyist, version 0.991 (Joseph Lau, Boston, MA).

RESULTS

General characteristics. We sent inquiries to 104 investigators working on SLE. Of those 104 investigators, 65 did not reply, 18 did not have any data and could not produce such data for the project, and 4 declined to participate. Of the last group, 2 investigators had published studies that were included in the sensitivity analysis.

The collaborative meta-analysis considered 1,537 lupus patients from 14 teams of investigators. Of these, 1,295 patients underwent both anti-P antibody testing by immunoblotting or standard ELISA and evaluation for NPSLE according to the ACR case definitions. The median sample size per study was 91 patients (interquartile range [IQR] 48–162). Women accounted for 80–97% of each study population. Although more than one-half of the participants were of European descent, patients of other ancestries were also included (Table 1). The mean age of the patients at study entry ranged from 29.8 years to 41.6 years, and the median of the mean disease durations across study cohorts was 7.3 years (IQR 6.2–7.8).

Most studies used a solid-phase ELISA, with highly purified synthetic peptides of the carboxy-terminal 22–amino acid sequence (n = 4), a multiple-antigen peptide format (n = 3), and purified native (n = 2) or recombinant (n = 3) proteins as coating antigen to detect anti-P antibodies. Seven studies designated a positive anti-P result as >2 SD (n = 1) or >3 SD (n = 6) above the mean value obtained in a normal population, whereas 5 studies reported results according to the suggested threshold for the commercial ELISA systems they used. Only 4 studies used Western blotting on cell extracts from various sources for the detection of this autoantibody specificity. A single study used a line immunoassay, which is an ELISA-based multianalyte assay (Table 1).

The median prevalence of anti-P antibodies was 18.2% (IQR 9.7–28.6%). These antibodies were more prevalent in lupus patients of Asian descent than among those of other racial ancestries. The study-specific frequencies of anti-P antibodies were 23.8–45.5% in 320 patients of Chinese, Japanese, Taiwanese, and Filipino ancestry and 6.4–25.4% in 1,212 patients of other ancestry.

Approximately one-third of the 1,537 lupus patients had NPSLE that manifested as syndromes described in the ACR case definitions (median prevalence 32% [IQR 12–42%]). In 1 study (Table 1), neuropsychiatric involvement was determined according to prespecified criteria other than the ACR case definitions. Eight research teams provided individual patient data; in these studies, 8% of patients had >1 neuropsychiatric disorder, but only 5% had both focal and diffuse presentations. The other 6 teams directly collected data on only the most prominent manifestation. More than one-half of the NPSLE patients presented with disorders reflecting diffuse cerebral involvement (median prevalence 54.5% [IQR 47.6–68.2%]). The median prevalence of psychosis, mood disorder, or both was 24.9% (IQR 17.1–38.4%). In most studies, NPSLE was diagnosed without knowledge of the anti-P antibody status, and test interpreters were blinded to the clinical condition of the patients (Table 1).

Diagnostic performance of anti-P antibody testing. Substantial heterogeneity was found in both the sensitivity and the specificity of anti-P antibody testing
<table>
<thead>
<tr>
<th>Study ID</th>
<th>Investigator, country, year (ref.)</th>
<th>Study setting</th>
<th>No. of patients</th>
<th>% women</th>
<th>Ethnicity (%)†</th>
<th>Mean age, years</th>
<th>Mean disease duration, years</th>
<th>Anti-P antibody assay</th>
<th>Prevalence of NPSLE, %</th>
<th>NPSLE manifestation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Doria A, Italy, 2004</td>
<td>University</td>
<td>101</td>
<td>88</td>
<td>Italian (98), African (2)</td>
<td>29.8</td>
<td>6.7</td>
<td>WB/ELISA</td>
<td>21</td>
<td>8 6 7 T, C</td>
</tr>
<tr>
<td>2</td>
<td>Morozzi G, Galeazzi M, Italy, 2004</td>
<td>University</td>
<td>20§</td>
<td>90</td>
<td>Italian (85), Chinese/Filipino (15)</td>
<td>35.7</td>
<td>7.6</td>
<td>ELISA</td>
<td>15</td>
<td>0 0 3 T, C</td>
</tr>
<tr>
<td>3</td>
<td>Afeltra A, Italy, 2004</td>
<td>University</td>
<td>43</td>
<td>88</td>
<td>Italian</td>
<td>41.6</td>
<td>8</td>
<td>ELISA</td>
<td>93</td>
<td>2 16 22 T, C</td>
</tr>
<tr>
<td>4</td>
<td>Mathieu A, Italy; Sanna G, UK, 2000 (24)</td>
<td>University</td>
<td>68¶</td>
<td>96</td>
<td>Italian</td>
<td>38.4</td>
<td>7.7</td>
<td>ELISA</td>
<td>49</td>
<td>7 9 17 T, C</td>
</tr>
<tr>
<td>5</td>
<td>Hoffman I, De Keyser F, Belgium, 2004 (14)</td>
<td>University</td>
<td>235#</td>
<td>88</td>
<td>Belgian, Dutch, Slovak, English</td>
<td>40</td>
<td>7.2</td>
<td>LIA</td>
<td>59</td>
<td>33 32 51 NS</td>
</tr>
<tr>
<td>6</td>
<td>Tzioufas A, Greece, 2000 (30)</td>
<td>University</td>
<td>185</td>
<td>96</td>
<td>Greek</td>
<td>34.7</td>
<td>4.3</td>
<td>ELISA</td>
<td>9</td>
<td>2 7 8 NS</td>
</tr>
<tr>
<td>7</td>
<td>Ambrozic A, Slovenia, 2003</td>
<td>University</td>
<td>150</td>
<td>91</td>
<td>Slovenian</td>
<td>38.1</td>
<td>7.8</td>
<td>WB</td>
<td>39</td>
<td>11 14 33 T, C</td>
</tr>
<tr>
<td>8</td>
<td>Inanc M, Turkey, 2004</td>
<td>University</td>
<td>218</td>
<td>89</td>
<td>Turkish</td>
<td>38.5</td>
<td>7.8</td>
<td>ELISA</td>
<td>23</td>
<td>20 5 26 T, C</td>
</tr>
<tr>
<td>9</td>
<td>Chang D-M, Taiwan, 2003</td>
<td>Community</td>
<td>80</td>
<td>91</td>
<td>Taiwanese</td>
<td>35</td>
<td>9.4</td>
<td>ELISA</td>
<td>6</td>
<td>1 3 1 NS</td>
</tr>
<tr>
<td>10</td>
<td>Mok CC, China, 2004</td>
<td>Community</td>
<td>33</td>
<td>97</td>
<td>Chinese</td>
<td>36.2</td>
<td>7</td>
<td>WB/ELISA</td>
<td>33</td>
<td>3 5 3 T, C</td>
</tr>
<tr>
<td>11</td>
<td>Hirohata S, Japan, 2003</td>
<td>University</td>
<td>50</td>
<td>80</td>
<td>Japanese</td>
<td>40.8</td>
<td>2.6</td>
<td>ELISA</td>
<td>32</td>
<td>5 7 4 T, C</td>
</tr>
<tr>
<td>12</td>
<td>Yoshio T, Japan, 2003 (35)</td>
<td>University</td>
<td>154**</td>
<td>90</td>
<td>Japanese</td>
<td>34.6</td>
<td>4.7</td>
<td>ELISA</td>
<td>40</td>
<td>14 24 24 T, C</td>
</tr>
<tr>
<td>13</td>
<td>Massardo L, Chile, 2002 (21)</td>
<td>University</td>
<td>141††</td>
<td>90</td>
<td>Chilean</td>
<td>33</td>
<td>7</td>
<td>WB/ELISA</td>
<td>9</td>
<td>5 1 6 T, C</td>
</tr>
<tr>
<td>14</td>
<td>Spindler AJ, Argentina, 2003</td>
<td>University</td>
<td>59</td>
<td>92</td>
<td>Argentinean</td>
<td>36</td>
<td>7.3</td>
<td>ELISA</td>
<td>44</td>
<td>11 4 11 T, C</td>
</tr>
</tbody>
</table>

* References and publication dates (when the contributed data were derived from published studies) are provided; otherwise, the year the data were collected and sent to the coordinating center are shown. See Patients and Methods for a full description of the 3 subgroups of neuropsychiatric systemic lupus erythematosus (NPSLE). Anti-P = anti-ribosomal P; WB = Western blotting; ELISA = enzyme-linked immunosorbent assay; LIA = line immunoassay.
† Percentages are given for studies that included patients of different ethnicities, when known.
‡ NPSLE was diagnosed without knowledge of the results of the anti-P antibody testing (T), and test interpreters were blinded to the clinical data (C). NS = not specified.
§ In this study, 3 patients had indeterminate results for anti-P antibodies and were not included in the quantitative synthesis.
¶ In this study, 5 patients who were not tested for anti-P antibodies were not included in the quantitative synthesis.
# In this study, sufficient clinical information for NPSLE was available for 196 patients; the presence or absence of NPSLE was assessed using prespecified criteria other than the American College of Rheumatology case definitions (7); and data for disease duration were available for 197 patients.
** Only 44 patients were included in the published study.
†† In this study, 2 patients in addition to the ones listed under NPSLE manifestations had NPSLE, but the type of involvement was not known.
using ELISA (Table 2). In the random-effects model, the overall weighted sensitivity and specificity estimates for the diagnosis of NPSLE were 26% (95% CI 15–42%) and 80% (95% CI 74–85%), respectively (Table 2).

Diagnostic performance for neuropsychiatric disease appeared to be somewhat better in studies that used Western blotting to detect anti-P antibodies (summary sensitivity 36% [95% CI 16–63%]; summary specificity 84% [95% CI 70–92%]), but significant between-study heterogeneity was still present ($P = 0.0001$ for heterogeneity in sensitivity estimates and $P = 0.0007$ for heterogeneity in specificity estimates), and data were too limited to be conclusive (4 studies; 424 patients). Test performance was poor for NPSLE in Asian patients (4 studies; 317 patients, yielding a summary sensitivity of 55% [95% CI 45–65%] and a summary specificity of 68% [95% CI 59–76%]). The weighted specificity tended to be higher in all other lupus patients, which were mostly of European descent, but there was low sensitivity (9 studies; 1,023 patients, yielding a summary sensitivity of 17% [95% CI 9–32%] and a summary specificity of 85% [95% CI 81–88%]).

SROC analyses suggested similar performance for identifying SLE-induced neuropsychiatric disease. Weighted and nonweighted curves were practically coincident (Figure 1A). Anti-P antibodies had an almost equally meager discriminating ability for the diagnosis of either psychiatric syndromes or other forms of neuropsychiatric involvement in SLE (Table 2). Weighted random-effects independent estimates stand very close to the weighted SROC curves for these comparisons (Figures 1B–D), suggesting that they are appropriate approximations of the overall diagnostic performance. Statistically significant asymmetry was found in all these curves (Figure 1), indicating that an improvement in specificity was accompanied by a disproportionately large decrease in sensitivity.

Within the group with NPSLE (Table 2), anti-P antibody testing could not accurately discriminate patients presenting with diffuse manifestations from those presenting with focal events (summary sensitivity 26%; summary specificity 70%) (Figure 2A) or patients presenting with psychiatric disorders from those presenting with any other diffuse symptom (summary sensitivity 28%; summary specificity 75%) (Figure 2B). Test characteristics remained unchanged for the identification of patients with psychiatric disorders compared with all other lupus patients (with or without neuropsychiatric dysfunction) (Table 2). Significant asymmetry was found in the corresponding SROC curve (Figure 2C), implying that an improvement in specificity was accompanied by an uneven, large decrease in sensitivity.

**Findings of additional analyses.** Our search of the 3 databases identified a total of 306 potentially relevant articles, of which 243 studies were excluded upon reading the titles and abstracts. Another 39 studies were excluded after reviewing the complete reports: 8 were editorials, comments without original data, or review articles, 11 were case reports, 7 studies presented duplicate or overlapping data, 8 evaluated anti-P antibody testing for other SLE manifestations or other autoimmune diseases, 3 focused on isolated neuropsy-
NPSLE versus non-NPSLE (A), psychosis and/or mood disorder versus non-NPSLE (B), other diffuse neuropsychiatric manifestations versus non-NPSLE (C), and focal neurologic events versus non-NPSLE (D).

Figure 1. Summary receiver operating characteristic curves for the performance of antibodies to ribosomal P proteins in the diagnosis of various forms of neuropsychiatric systemic lupus erythematosus (NPSLE). Results are from the main analysis. Each ellipse corresponds to a study estimate of sensitivity and specificity; the area of each ellipse is proportional to the study size. Numbers beside the ellipses are study identification numbers and correspond to those shown in Table 1. Thin lines indicate nonweighted analyses; thick lines indicate weighted analyses. Shaded rectangles mark the 95% confidence intervals of the pooled sensitivity and pooled specificity obtained by random-effects calculations. × indicates exact estimates. A, NPSLE overall versus non-NPSLE. B, Psychosis and/or mood disorder versus non-NPSLE. C, Other diffuse neuropsychiatric manifestations versus non-NPSLE. D, Focal neurologic events versus non-NPSLE.

Twenty-four additional publications (6,8–13,15–20,22,23,25–29,31–34) were retrieved from the database search, representing a total of 38 studies involving 3,713 lupus patients. Nevertheless, data for the comparison of NPSLE versus non-NPSLE groups were available in only 18 of the 24 additional studies; data for other comparisons were available in even fewer reports (Table 3). The results were consistent with those derived from the collaborative meta-analysis (Table 3 and Figure 3), but between-study heterogeneity was always considerable (Table 3). The overall weighted sensitivity and specificity estimates for identifying patients with NPSLE were 28% (95% CI 22–35%) and 80% (95% CI 75–85%), respectively. The SROC curve for this comparison was located very close to the diagonal, indicating poor diagnostic performance (Figure 3A). The overall sensitivity for psychosis, mood disorder, or both was slightly
improved, but it was still suboptimal (42%), and the specificity remained essentially the same (81%). There was still significant asymmetry in the SROC curves for the diagnosis of psychiatric disorders (Figures 3B and C). Anti-P antibody testing was not more accurate when used to discriminate active NPSLE from non-NPSLE (Table 3 and Figure 3D). Weighted and nonweighted SROC curves were almost coincident in all these contrasts (Figure 3).

**Findings of other sensitivity analyses.** Analyses limited to studies that used the ACR criteria for NPSLE yielded similar results. The weighted sensitivity for NPSLE overall was 29% (95% CI 17–45%) and the weighted specificity was 79% (95% CI 73–84%). Analyses excluding studies that did not specify blinding yielded a sensitivity of 25% (95% CI 13–43%) for the diagnosis of NPSLE and a specificity of 79% (95% CI 70–86%). Likewise, the diagnostic performance of anti-
P antibodies was largely unaffected in all other comparisons (data not shown).

**DISCUSSION**

This meta-analysis demonstrated with large-scale evidence that the value of anti-P antibody testing for the diagnosis of NPSLE overall or for particular disease phenotypes is negligible. No large differences in diagnostic performance with ELISA measurements or with Western blotting were discerned. Serum anti-P antibodies are detected by ELISA in less than one-third of patients with NPSLE, while 15–25% of lupus patients without neuropsychiatric involvement have this autoantibody specificity. Testing for anti-P antibody is not useful in excluding disease-mediated psychosis or mood disorder with enough certainty, since more than 60% of cases are false negative. Also, a false-positive rate of ~20% militates against the dependence on this laboratory test for diagnosing psychiatric disorders in lupus patients.

Whereas nearly all studies suggested poor diagnostic performance, the exact test performance varied substantially. Variability beyond chance could be attributed to ethnic differences in the study patients, the clinical setting, the type of assay used, differences in test thresholds, and differences in therapy at the time of testing. Anti-P antibodies were more prevalent in Asian patients with lupus than among those of other racial ancestries. This finding is consistent with the observation that their production is influenced by certain class II major histocompatibility complex alleles (8). Despite the use of uniform criteria for defining neuropsychiatric disease, the prevalence of NPSLE differed across centers. This difference probably reflects varying referral patterns at the research sites, as well as varying practice patterns for performing anti-P antibody testing in lupus patients with possible NPSLE syndromes.

The immunoassays used for anti-P antibody determination often differed in terms of the antigenic source, the conditions of protein extraction and denaturation, the nature of the coating antigen, and the carrier proteins and coupling agents used for binding antigen to the plate. The selected cutoff value designating a positive result in enzyme immunoassays could also affect the sensitivity and specificity. Nevertheless, a standardization of anti-P antibody testing is essential to avoiding technical or analytical differences among centers. Treatment with immunosuppressive drugs at the time of testing might influence the antibody response and, therefore, could also account for the discrepancies in test performance. Heterogeneity stemming from all these sources is probably unavoidable, and it reflects actual clinical practice.

Our analysis addressed heterogeneity by using a random-effects model that incorporated the uncertainty arising from between-study differences. SROC curves, which correct for variation due to differences in test thresholds across studies, were also consistent with the independently weighted estimates, and accordingly, the results of the meta-analysis should be generalizable to diverse settings.

Specific design flaws of primary studies of diagnostic tests including lack of blinding, use of different reference tests according to the results of the experimental test, and insufficient description of the population under study can lead to biased, usually optimistic estimates of diagnostic accuracy (48). Our study had the methodologic advantage of using data from adequately described lupus cohorts in which a consistent application of standardized definitions of NPSLE syndromes, and

### Table 3. Summary results of additional analyses that included published studies from database searches*

<table>
<thead>
<tr>
<th>Comparison</th>
<th>No. of studies</th>
<th>No. of subjects</th>
<th>Weighted sensitivity (95% CI)</th>
<th>Weighted specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPSLE versus non-NPSLE</td>
<td>32</td>
<td>2,861</td>
<td>0.28 (0.22–0.35)</td>
<td>0.80 (0.75–0.85)</td>
</tr>
<tr>
<td>Psychosis and/or mood disorder versus non-NPSLE</td>
<td>25</td>
<td>1,909</td>
<td>0.42 (0.30–0.53)</td>
<td>0.81 (0.76–0.85)</td>
</tr>
<tr>
<td>Patients with psychosis and/or mood disorder</td>
<td>31</td>
<td>3,309</td>
<td>0.41 (0.31–0.52)</td>
<td>0.81 (0.77–0.85)</td>
</tr>
<tr>
<td>versus all other lupus patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active NPSLE versus non-NPSLE</td>
<td>10</td>
<td>1,025</td>
<td>0.34 (0.27–0.43)</td>
<td>0.82 (0.74–0.87)</td>
</tr>
</tbody>
</table>

* Data from the studies shown in Table 1 as well as from additional studies retrieved from a search of the Medline, EMBase, and Cochrane databases are included. Weighted sensitivity and specificity were determined according to the random-effects model. Between-study heterogeneity was statistically significant for all comparisons (P < 0.01). 95% CI = 95% confidence interval; NPSLE = neuropsychiatric systemic lupus erythematosus.
blinded interpretation of both the test results and the reference standard was ensured in most cases. In addition, the overall estimates did not materially change after we excluded the few studies that did not specify blinding or did not use the ACR case definitions for NPSLE.

We should acknowledge that the ACR criteria may not be a perfect reference standard for assessing the presence or absence of NPSLE syndromes in lupus patients. In fact, this classification system has been criticized for some lack of specificity; disorders such as headache, anxiety, mild cognitive dysfunction, mild depression, and polyneuropathy without electrophysiologic confirmation may not truly be NPSLE syndromes (1,49). Nevertheless, until revised criteria (49,50) are accepted and validated, the ACR case definitions constitute the best available tool with which to categorize neuropsychiatric events in SLE (4,51).

Another limitation of the study is that patients having both diffuse and focal NPSLE events were clas-
Based on the categorization standards adopted by the American College of Rheumatology, the accuracy of diagnostic tests in systemic lupus erythematosus (SLE) is an important consideration. However, the true diagnostic performance of anti-P antibodies may have remained unpublished. Studies that failed to show a diagnostic value for anti-P antibodies might have been misclassified as non-NPSLE patients, because anti-P antibodies had poor discriminating ability and significant effects on the estimated performance. Nevertheless, this limitation is unlikely to affect the generalizability of the findings of the primary investigations, as the observed sensitivity and specificity estimates have widely overlapped when further published studies were included in the analyses. However, these estimates have shown that the diagnostic ability of anti-P antibody in the cerebrospinal fluid needs further study, although it seems to be more limited than the ability of serum autoantibodies to express itself. This seems implausible, since nervous system involvement occurs within the first 2 years of disease onset in most patients and rarely presents late. Therefore, the assessment of the contributions of these factors to the development of neuropsychiatric systemic lupus erythematosus (NPSLE) (NPSLE) is important in the context of other important considerations.

In conclusion, anti-P antibody testing has negligible effect on the overall sensitivity of anti-P antibody, and further studies are needed to improve the accuracy of diagnostic tests in SLE, which is often complicated by neuropsychiatric features. Furthermore, the absence of neurological involvement in the study population might have affected the performance of anti-P antibodies, and the absence of neurological involvement is more frequent in patients with NPSLE (2,3,6). Therefore, the evidence is far sparser than in other fields. Furthermore, the absence of neurological involvement in the study population might have affected the performance of anti-P antibodies, and the absence of neurological involvement is more frequent in patients with NPSLE (2,3,6). Therefore, the evidence is far sparser than in other fields.

We acknowledge the contributions of the other investigators of the collaborating study teams involved in this project, as follows: Borut Bozic, PhD, and Blaz Rozman, MD, Teikyo University School of Medicine, Tokyo, Japan; Takashi Sawada, MD, and Masahiko Polat, PhD, and Bahar Artim Esen, MD, Istanbul Medical University Medical Centre, Ljubljana, Slovenia; Nuray Gurel, MD, Istanbul Medical University Medical Centre, Ljubljana, Slovenia; and S. West, MD, Hong Kong University Medical Centre, Hong Kong, SAR, China; and S. West, MD, Emlen W, Wener MH, Kotzin BL. Neuropsychiatric systemic lupus erythematosus: a comprehensive review of the literature. Rheum Dis Clin North Am 2005;31:2156–62.

ACKNOWLEDGMENTS

The authors acknowledge the contributions of the other investigators of the collaborating study teams involved in this project, as follows: Borut Bozic, PhD, and Blaz Rozman, MD, Teikyo University School of Medicine, Tokyo, Japan; Takashi Sawada, MD, and Masahiko Polat, PhD, and Bahar Artim Esen, MD, Istanbul Medical University Medical Centre, Ljubljana, Slovenia; Nuray Gurel, MD, Istanbul Medical University Medical Centre, Ljubljana, Slovenia; and S. West, MD, Hong Kong University Medical Centre, Hong Kong, SAR, China; and S. West, MD, Emlen W, Wener MH, Kotzin BL. Neuropsychiatric systemic lupus erythematosus: a comprehensive review of the literature. Rheum Dis Clin North Am 2005;31:2156–62.

REFERENCES

5. Rood MJ, Breedveld FC, Huizinga TW. The accuracy of diagnostic tests in SLE. Such an approach could limit health care costs by preventing unnecessary testing. The overall sensitivity of anti-P antibody is modified by several factors, which have a significant effect on the observable diagnostic performance of anti-P antibodies. The median disease duration in the study population was 3.5 years. A further explanation for anti-P antibody performance is the presence of other manifestations that could have been misclassified as non-NPSLE patients, because anti-P antibodies had poor discriminating ability and significant effects on the estimated performance. Nevertheless, this limitation is unlikely to affect the generalizability of the findings of the primary investigations, as the observed sensitivity and specificity estimates have widely overlapped when further published studies were included in the analyses. However, these estimates have shown that the diagnostic ability of anti-P antibody in the cerebrospinal fluid needs further study, although it seems to be more limited than the ability of serum autoantibodies to express itself. This seems implausible, since nervous system involvement occurs within the first 2 years of disease onset in most patients and rarely presents late. Therefore, the assessment of the contributions of these factors to the development of neuropsychiatric systemic lupus erythematosus (NPSLE) (NPSLE) is important in the context of other important considerations.

In conclusion, anti-P antibody testing has negligible effect on the overall sensitivity of anti-P antibody, and further studies are needed to improve the accuracy of diagnostic tests in SLE, which is often complicated by neuropsychiatric features. Furthermore, the absence of neurological involvement in the study population might have affected the performance of anti-P antibodies, and the absence of neurological involvement is more frequent in patients with NPSLE (2,3,6). Therefore, the evidence is far sparser than in other fields. Therefore, the evidence is far sparser than in other fields.

We acknowledge the contributions of the other investigators of the collaborating study teams involved in this project, as follows: Borut Bozic, PhD, and Blaz Rozman, MD, Teikyo University School of Medicine, Tokyo, Japan; Takashi Sawada, MD, and Masahiko Polat, PhD, and Bahar Artim Esen, MD, Istanbul Medical University Medical Centre, Ljubljana, Slovenia; Nuray Gurel, MD, Istanbul Medical University Medical Centre, Ljubljana, Slovenia; and S. West, MD, Hong Kong University Medical Centre, Hong Kong, SAR, China; and S. West, MD, Emlen W, Wener MH, Kotzin BL. Neuropsychiatric systemic lupus erythematosus: a comprehensive review of the literature. Rheum Dis Clin North Am 2005;31:2156–62.

ACKNOWLEDGMENTS

The authors acknowledge the contributions of the other investigators of the collaborating study teams involved in this project, as follows: Borut Bozic, PhD, and Blaz Rozman, MD, Teikyo University School of Medicine, Tokyo, Japan; Takashi Sawada, MD, and Masahiko Polat, PhD, and Bahar Artim Esen, MD, Istanbul Medical University Medical Centre, Ljubljana, Slovenia; Nuray Gurel, MD, Istanbul Medical University Medical Centre, Ljubljana, Slovenia; and S. West, MD, Hong Kong University Medical Centre, Hong Kong, SAR, China; and S. West, MD, Emlen W, Wener MH, Kotzin BL. Neuropsychiatric systemic lupus erythematosus: a comprehensive review of the literature. Rheum Dis Clin North Am 2005;31:2156–62.

REFERENCES

5. Rood MJ, Breedveld FC, Huizinga TW. The accuracy of diagnostic tests in SLE. Such an approach could limit health care costs by preventing unnecessary testing.