Phospholipid Autoantibodies and the Antiphospholipid Antibody Syndrome: Diagnostic Accuracy of 23 Methods Studied by Variation in ROC Curves with Number of Clinical Manifestations

Jan-Christian Wasmuth,* Desamparados Oliver y Miñarro, Angela Homrighausen, Ludger Leifeld, Jürgen K. Rockstroh, Tilman Sauerbruch, and Ulrich Spengler

Background: We analyzed the diagnostic accuracies for the diagnosis of antiphospholipid syndrome (APS) of 23 antiphospholipid antibody (APL-Ab) assays targeted at different antigen preparations and immunoglobulin isotypes.

Methods: In 144 patients with suspected APS, anti-cardiolipin (aCL) and anti-β2-glycoprotein I (aβ2GPI) antibodies were measured with 23 different ELISAs from three manufacturers. Data were analyzed by ROC curves. In the absence of an accepted criterion standard, the endpoint “diagnosis of APS” was varied according to the number (two through five) of signs and symptoms of APS.

Results: Although the presence of lupus anticoagulant was associated significantly with APL-Ab in 10 of 23 assays ($P = 0.01–10^{-4}$) and recurrent arterial or venous occlusions were significantly associated with APL-Ab of IgM isotype in 5 of 6 assays ($P = 0.02–10^{-4}$), sensitivity for detection of APS did not exceed 67%. With the exception of IgA APL-Ab, the diagnostic accuracy of the assays improved when the diagnosis of APS was based on an increasing number of simultaneous features of APS. For most methods, areas under the ROC curves were $>0.8$ irrespective of the method’s subclass specificity and antigen preparation (aCL or aβ2GPI), if the clinical diagnosis of APS was based on four or more signs and symptoms of APS.

Conclusion: Despite considerable heterogeneity in the individual test results, a single test of IgG or IgM isotype targeted at either aCL or aβ2GPI antibodies has excellent diagnostic accuracy when the criterion for diagnosis requires four or more typical manifestations of APS.

© 2002 American Association for Clinical Chemistry

Recurrent arterial or venous thromboembolism is the clinical hallmark of the antiphospholipid syndrome (APS),¹ which can occur as a complication of various autoimmune diseases (1). However, the possible spectrum of clinical manifestations varies considerably and involves spontaneous abortions, thrombocytopenia, prolonged in vitro bleeding tests (lupus anticoagulant), livedo racemosa, and various neurologic symptoms (2, 3). APS seems to be caused by autoantibodies that are directed against phospholipid antigens in the membranes of endothelial cells and platelets, and the great variability and heterogeneity seen at clinical presentation probably reflect the heterogeneity at the antibody level.

Among the antiphospholipid antibodies (APL-Abs), the appearance of anticardiolipin (aCL) antibodies is particularly associated with the APS (1), although antibodies can also interfere with the function of other pro- or anticoagulatory proteins, such as prothrombin, (activated) protein C, protein S, annexin V, factor XII, thrombomodulin, and tissue-type plasminogen activator (4–7).

¹ Nonstandard abbreviations: APS, antiphospholipid syndrome; APL-Ab, antiphospholipid antibody; aCL, anti-cardiolipin; aβ2GPI, anti-β2-glycoprotein I; SLE, systemic lupus erythematosus; LAC, lupus anticoagulant; and AUC, area under the curve.
in the diagnosis of APS, we measured aCL and aβ₂GPI antibodies react with cardiolipin in the presence of a 50-kDa serum cofactor, which was identified as β₂-glycoprotein I (β₂GPI) by sequencing of the N-terminal amino acids (8). β₂GPI enhances the binding of aCL antibodies to cardiolipin in systemic lupus erythematosus (SLE) (9).

Accumulating evidence suggests that antibodies to β₂GPI (aβ₂GPI antibodies) are the actual pathogenic principle that leads to APS (10, 11). This has produced controversy regarding which antigen preparation should be used for the detection of APL-Abs in clinical practice. A recent consensus conference suggested that correct identification of APS-associated autoantibodies is of outmost importance to make a correct diagnosis (12), but it is still unclear whether certain isotypes of aCL antibodies are linked to the presence of distinct clinical manifestations (13). For example, IgA isotypes have been proposed to be particularly linked with an increased risk of thrombocytopenia, skin manifestations, and thrombotic events (14, 15). However, others did not find such associations (16, 17). Moreover, the prevalence of IgA-aCL antibodies may show significant heterogeneity with respect to different ethnic backgrounds (18).

To clarify the clinical importance of different APL-Abs in the diagnosis of APS, we measured aCL and aβ₂GPI antibodies with 23 reagent sets targeted either at different immunoglobulin isotypes or antigenic preparations, respectively, in 144 Caucasoid patients with the clinical suspicion of APS. In these patients, clinical manifestations of APS were validated by a standardized questionnaire, and test accuracy was studied by ROC analysis with “diagnosis of APS” as endpoint. To study the relationship between the clinical heterogeneity in the presentation of APS and the presence and heterogeneity of APL-Abs, the endpoint, diagnosis of APS, was varied according to the numbers (two through five) of APS-related features that were present simultaneously.

Materials and Methods

Participants
The study was based on serum samples from 160 Caucasoid patients for whom analysis of APL-Abs was requested between November 1996 and August 1997 because their treating physicians suspected APS. This period had been chosen based on the anticipated numbers of ordered tests to limit the analytical costs. Blood samples had been ordered by different departments of the Bonn University Hospital: Internal Medicine (n = 84); Neurology (n = 31); Dermatology (n = 17); Gynecology (n = 21); and Pediatrics (n = 7).

To take into account differences in diagnostic standards, the patient records were reevaluated by the study personnel, using standardized criteria and a unified protocol (19) to identify clinical manifestations of APS: (a) recurrent venous thrombosis (verified by either ultrasound or phlebography); (b) thrombocytopenia (<100 g/L in both EDTA and citrate blood to exclude pseudothrombocytopenia); (c) recurrent spontaneous abortions (more than two); (d) neurologic symptoms (such as transient ischemic attacks and hemiplegia) and signs of arterial or venous occlusion verified by Doppler ultrasound, angiography, or magnetic resonance imaging angiography; (e) peripheral arterial occlusion (verified by angiography or operative thrombectomy); (f) amaurosis fugax; (g) livedo reticularis; (h) ulceration of the legs; (i) SLE or other autoimmune processes; and (j) presence of a prolonged activated partial thromboplastin time as a marker of lupus anticoagulant (LAC), confirmed according to the guidelines of the International Society on Thrombosis and Hemostasis (20) (unexplained prolongation of the activated partial thromboplastin time that was not reversed after mixing and dilution of the patient’s plasma with normal platelet-free plasma; confirmatory tests were dilute Russell’s viper venom time and kaolin clotting time).

For 16 patients, the records were incomplete and did not allow a definite decision on the presence of APS-associated manifestations. These 16 patients were excluded from the analysis, and 144 well-characterized patients were entered into the analysis (50 males and 94 females; mean age ± SD, 44 ± 12 years). We studied clinical features and APL-Abs at a second time point in 53 of these 144 patients. Blood samples from 10 healthy, symptom-free controls without any known underlying disease (3 males and 7 females; mean age ± SD, 42 ± 14 years of age) were included as negative controls. All study procedures were done in accordance with the current revision of the Helsinki Declaration of 1975. Finally, the study design was oriented on the 2000 guidelines for evaluation of diagnostic accuracy (21), and samples were deidentified.

Laboratory Procedures

Serum samples were collected and frozen at −80 °C. For quantitative determination of aCL and aβ₂GPI antibodies, ELISAs from three different manufacturers were used (summarized in Table 1). All assays were used according to the manufacturers’ instructions according to standard ELISA procedures. In-house calibrators were used for standardization. Negative controls were included in the assays to ensure negative test results and were included in the statistical analysis. Absorbances were determined with a photometer at 450 nm (Spectra Mini; Tecan). The cutoff values predefined by the manufacturers and the analytical precisions of the ELISAs are given in Table 1. The laboratory staff was blinded with respect to the clinical presentation of the patients.

Statistical Analysis

To characterize each ELISA with respect to its accuracy to detect patients with an APS, sensitivity and specificity were calculated from 2 × 2 contingency tables for each possible cutoff value and plotted as ROC curves. The area under the ROC curve (AUC), W, and its standard error, SE_W, were calculated according to the methods of Hanley and McNeil (22) and Beck and Schultz (23). The signific-
The results of both 
\( z \) and 
\( \text{SE} \) were calculated as:

\[
\frac{W_\text{A} - W_\text{B}}{(\text{SE}_{W_\text{A}})^2 - 2r(\text{SE}_{W_\text{A}}\text{SE}_{W_\text{B}})}
\]

where \( r \) represents the estimated correlation between \( W_\text{A} \) and \( W_\text{B} \). [For calculation of \( r \) and further details, refer to Hanley and McNeil (24) and Zweig and Campbell (25).]

The results of both \( z \) statistics are given as \( P \) (two-sided).

A cut-off external criterion standard not involving APL-Ab testing does not exist for the diagnosis of APS. Therefore, we took the presence of typical manifestations of APS as a substitute for such a gold standard, and ROC analysis was performed repeatedly with different definitions of APS diagnosis, depending on the number of simultaneous signs and symptoms. Patients were arbitrarily divided into four subgroups, based on the signs and symptoms. In group I, the diagnosis of APS was assumed if two or more of the above-mentioned 10 clinical criteria were met. Similarly, patients in group II fulfilled three or more of the above criteria; in group III, four or more of the above criteria were fulfilled, and in group IV, at least five criteria were fulfilled. In each of these analyses, patients with less than the required number of features were allocated to the “disease-free” group in the \( 2 \times 2 \) contingency tables.

Statistical analysis was performed with SPSS software (Ver. 6.1.3; SPSS Inc.), and ROC analysis was performed on an electronic spreadsheet (Microsoft Excel 7.0; Microsoft Corp.).

### Results

**Table 1. Investigated ELISAs, predefined cutoff values, and interassay CVs.***

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>aCL*</th>
<th>CV, %</th>
<th>aβ2 GPI*</th>
<th>CV, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company A</td>
<td>IgG &gt;18 units</td>
<td>6.5–8.4</td>
<td>IgG &gt;15 units</td>
<td>7.1–9.5</td>
</tr>
<tr>
<td></td>
<td>IgM &gt;10 units</td>
<td>9.3–11.7</td>
<td>IgM &gt;15 units</td>
<td>8.5–12.9</td>
</tr>
<tr>
<td></td>
<td>IgA &gt;13 units</td>
<td>7.4–14.9</td>
<td>IgA &gt;10 units</td>
<td>7.6–13.0</td>
</tr>
<tr>
<td></td>
<td>Screen &gt;1 unit</td>
<td>5.3–12.0</td>
<td>Screen &gt;1 unit</td>
<td>5.4–6.5</td>
</tr>
<tr>
<td>Company B</td>
<td>IgG &gt;20 units</td>
<td>2.9–7.4</td>
<td>IgG &gt;20 units</td>
<td>3.5–5.4</td>
</tr>
<tr>
<td></td>
<td>IgM &gt;20 units</td>
<td>2.3–5.1</td>
<td>IgM &gt;20 units</td>
<td>2.5–6.6</td>
</tr>
<tr>
<td></td>
<td>IgA &gt;15 units</td>
<td>11.5–13.1</td>
<td>IgA &gt;20 units</td>
<td>4.0–6.6</td>
</tr>
<tr>
<td></td>
<td>Screen &gt;500 units</td>
<td>4.1–10.0</td>
<td>Screen &gt;500 units</td>
<td>4.1–10.0</td>
</tr>
<tr>
<td>Company C</td>
<td>IgG &gt;10 units</td>
<td>16.1</td>
<td>IgG &gt;7 units</td>
<td>18.3</td>
</tr>
<tr>
<td></td>
<td>IgM &gt;9 units</td>
<td>17.9</td>
<td>IgM &gt;7 units</td>
<td>14.0</td>
</tr>
<tr>
<td></td>
<td>IgA &gt;8 units</td>
<td>15.5</td>
<td>Not provided</td>
<td>Not provided</td>
</tr>
<tr>
<td></td>
<td>Screen &gt;9 units</td>
<td>Not provided</td>
<td>Not provided</td>
<td>Not provided</td>
</tr>
</tbody>
</table>

* Interassay CVs were indicated by the manufacturers.
* Units defined through standard ELISA gauging procedures for each assay.
* Screening assays contained antibodies of all isotypes.

**Clinical Features and ROC Analysis of Raw Data**

Of the 144 patients in our study, 25 had recurrent venous thromboembolism, 21 had peripheral arterial occlusions, 10 had ulcerations of the legs, 8 had livedo reticularis, and 37 had ischemic neurologic symptoms or signs together with radiologic evidence of cerebral vessel occlusions. Twenty-seven patients had SLE, 24 had thrombocytopenia, and 9 had detectable concentrations of LAC. Reevaluation of the patient records according to the predefined study criteria failed to confirm any objective manifestations of APS in 38 of the 144 patients. On the basis of the simultaneous presence of several features of APS, 45 patients were classified in group I (two or more features), 20 patients in group II (three or more features), 7 in group III (four or more features; 5 with LAC), and 4 in group IV (five or more features; 4 with LAC). For 53 patients (19 with at least a positive APL-Ab and 34 without any APL in the first serum analysis), a second serum sample was analyzed at a different time point. However, none of these 53 patients had to be allocated to a different group based on the second APL-Ab test or newly emerging features of APS.

In our primary ROC analysis, the presence of LAC was associated significantly with the presence of APL-Ab for 10 of 23 ELISAs (aCL antibodies, IgG and IgM by ELISAs from companies A and B, IgA by ELISAs from companies B and C; aβ2 GPI antibodies, IgG by ELISAs from all three companies, IgA by ELISAs from company B; range of AUCs, 0.75–0.93; range of SE, 0.080–0.109; range for \( z \), 0.01–10^-4). Antibodies of the IgM isotype were significantly associated with recurrent venous thromboembolism or arterial occlusion in five of six ELISAs (aCL antibodies, ELISAs from all three companies; aβ2 GPI
antibodies, ELISAs from companies A and B; range of AUCs, 0.64–0.76; range of SE, 0.049–0.075; range of P, 0.02–10^−4), and with ulceration of the legs in two IgM ELISAs [IgM aCL antibodies by the ELISA from company B (AUC = 0.67; SE = 0.075; P < 0.05); and IgM aβ2-GPI antibodies by the ELISA from company A (AUC = 0.70; SE = 0.091; P < 0.05)]. The best diagnostic accuracy for the association of a positive antiphospholipid test result and any of the other APS-associated phenomena was found with the IgG aβ2-GPI ELISA from company C and the presence of LAC (AUC = 0.93; SE = 0.102; P < 10^−4). However, even this association had only 67% sensitivity and 91% specificity with the cutoff provided by the manufacturer. Therefore, our further analysis was focused on assay performance to detect patients who had several APS-associated features simultaneously.

**ROC analysis by stratification based on number of APS criteria**

The ROC characteristics (AUCs) improved in parallel to the number of required criteria to make a positive diagnosis of APS for all assays except the aCL antibody tests from company C [Fig. 1; also see the data supplement that accompanies the online version of this article on the Clinical Chemistry Online web site (http://www.clinchem.org/content/vol48/issue7/)]. Areas under the ROC curves >0.8 were obtained when four or more typical manifestations (group III) were set as the definition of APS (Fig. 2).

When we used the cutoff values provided by the manufacturers, sensitivities were ≤67% in groups I and II for any of the assays (range, 4–67%), whereas in groups III and IV, sensitivities reached 100% for the IgG, IgM, and screen assays (33–100%), indicating that stricter criteria for the definition of APS produced fewer false-negative results. The analyzed tests to detect IgA antibodies were inferior to the other test types regardless of aC or aβ2-GPI antibody specificity or manufacturer. Sensitivities did not exceed 25% in the best IgA tests (data not shown). Therefore, IgA tests were excluded from further statistical analysis, as were all aC antibody tests from company C because their diagnostic accuracy remained poor when the number of clinical criteria for the diagnosis of APS was varied. Considering the relatively small numbers of patients in subgroups III and IV, it was not possible to analyze whether any of the tests significantly outperformed the other tests.

**Discussion**

The clinical importance of APS has found widespread recognition. Various symptoms, such as vascular thrombosis and pregnancy losses, are associated with APS. However, the mechanisms by which the associations occur are not well understood and appear to be heterogeneous, probably reflecting the heterogeneity of APL-Abs. In an attempt to facilitate future studies, a working definition based on clinical and laboratory features most closely related to APS has been proposed by an international expert panel (12), not withstanding four prior proposals to find a consensus definition for this syndrome. However, these proposals have only imperfectly reflected the heterogeneity of APL-Abs and their relation to distinct symptoms or complexes of symptoms.

In this study, we searched for APL-Abs in 144 patients, using 23 different ELISAs based on antibodies of different isotypes and specificities. To account for differences attributable to the antigen preparations, we chose assays from three different manufacturers. To compare the diagnostic accuracy of these test methods, we used ROC analysis with surrogate definitions of APS, which varied in the numbers of APS-related features, and we assumed that the AUCs for the ROC curves would increase with the progressive allocation of patients to the disease-free group. This approach was chosen to circumvent the problem that current consensus definitions of APS could not be used as a gold standard for such an analysis because APL-Abs themselves are an integral part of the definition.

A further consideration in our approach was that the use of a fixed criterion is particularly prone to spectrum bias in a disease, such as APS, with highly variable manifestations, i.e., only patients with severe manifestations are diagnosed correctly. In contrast, our concept of a variation of ROC curves with the number of clinical manifestations also takes into account variable disease expression. Choosing an incremental number of diagnostic criteria as the endpoint in our ROC analysis produced higher apparent diagnostic accuracies for the tests (except for the aCL antibody ELISAs manufactured by company C). This finding together with the reliable exclusion of the negative controls is expected for any reasonable tests and, thus, strengthens the internal validity of our analysis.
Our analysis indicated that an almost-perfect ROC plot was obtained if the clinical diagnosis of APS was based on four or more features. This almost-perfect match was obtained with tests targeted at different antigen specificities (cardiolipin and $\beta_2$GPI) and from different manufacturers. In general, tests for IgG, IgM, or IgG + IgM (screen) produced equivalent results, similar to those reported previously by Bertolaccini et al. (27). However, the numbers of patients in groups III and IV were rather low; therefore, our ability to detect statistically significant differences among tests was rather limited.

It was unexpected that the aCL assays manufactured by company C failed to correctly identify patients with APS when stricter criteria were required for the diagnosis of APS. This company had added abundant $\beta_2$GPI in the dilution buffers of their aCL reagent sets to exclude false-positive reactivities attributable to $\beta_2$GPI contamination in the cardiolipin antigen preparation in the solid
phase, whereas such measures were not taken by the other manufacturers. There is accumulating evidence that \( \alpha_2 \)-GPI may be a more relevant antibody than aCL with respect to clinical manifestations of APS (10, 11, 28).

The failure of the aCL assay manufactured by company C in our analysis supports the concept that \( \alpha_2 \)-GPI antibodies are better correlated to clinical symptoms and that the diagnostic power of so-called cardiolipin tests is actually based on the cross-reaction of antibodies with contaminating \( \beta_2 \)-GPI. This interpretation is in line with recent data reported by Day et al. (29) and Gomez-Pacheco et al. (30). Cross-reactivity with contaminating \( \beta_2 \)-GPI probably also explains why \( \beta_2 \)-GPI tests achieved areas under the ROC curves similar to those of the aCL tests (excluding the aCL ELISAs from company C).

Tests targeted at IgA isotypes performed less well than tests targeted at antibodies of the other isotypes. This finding is in line with the data from Wilson et al. (16) and Selva-O’Callaghan et al. (17), who reported that IgA antibodies do not play an important role in the diagnosis of APS. However, these studies, as well as our present analysis, were done in a Caucasian population, and considerably higher prevalences of IgA-isotype APL-Abs have been reported in populations with different ethnic backgrounds (18, 31, 32). Therefore, our study does not exclude that IgA-aCL testing is of use in such populations.

No relevant association of distinct antibodies or immunoglobulin classes with single clinical symptoms was found. Although ROC analysis revealed significant AUC values for the detection of recurrent venous thromboembolism, arterial occlusion, or ulceration of the legs by IgM antibodies, sensitivities always were <50% when we used the cutoffs provided by the manufacturers. Likewise, sensitivities were ≤44% for the associations between IgG and IgM aCL antibodies or IgG \( \alpha_2 \)-GPI antibodies and the presence of LAC.

The activated partial thromboplastin time used to detect the presence of LAC in our study can be normal in a substantial proportion of LAC-positive plasmas. Thus, we cannot exclude that our approach to identifying LAC has underestimated the prevalence of LAC-positive samples. Nevertheless, our data are in contrast to some studies, which had demonstrated an association between venous or arterial thrombosis and recurrent fetal loss with the presence of APL-Abs (33); other studies, however, did not find associations between immunoglobulin isotypes and clinical manifestations of APS (34, 35) or between APL-Abs and SLE (36). This obvious paradox reflects the great variability of the APS itself, the variability of different assay methods (37), and the difficulty in comparing results from different investigators (38). Accordingly, we can not exclude a heterogeneity of our own study population, thereby hampering comparison with other groups.

In conclusion, our analysis confirms that aCL and \( \alpha_2 \)-GPI antibodies are suitable diagnostic markers for APS despite considerable variability in the diagnostic accuracy of commercially available tests. Current guidelines recommend repeated testing of phospholipid autoantibodies to make a reliable diagnosis of APS (12), whereas our analysis was based on a single test result to meet statistical requirements. However, it was reassuring that none of our 53 patients who were evaluated at a second time point had to be allocated to a different group based on clinical features or APL-Ab results. Our analysis of diagnostic accuracy suggests that a single test targeted at IgG or IgM aCL or \( \alpha_2 \)-GPI antibodies has satisfactory accuracy to detect the clinical presence of APS. Therefore, we recommend that further studies with rigorous statistical procedures to be done in different ethnic populations and on a broader clinical basis to clarify whether incremental testing with different tests is really necessary in the diagnostic routine.

References


