Diagnostic Performance and Predictive Value of Rheumatoid Factor, Anti-citrullinated Peptide Antibodies, and the HLA Shared Epitope for Diagnosis of Rheumatoid Arthritis, Ilse E.A. Hoffman,1* Isabelle Peene,1 Hans Pottel,2 Ann Union,2 Frank Hulstaert,2 Lydie Meheus,2 Katleen Assay (LIATM) for the detection of anti-pepA and anti-pepB antibodies which can be identified by tests such as a line immunoblot (INNO-LIA) for the detection of anti-citrullinated peptide antibodies (ACPAs), however, often low.

Rheumatoid factor (RF) is currently the most accepted laboratory test for rheumatoid arthritis (RA) and is part of the revised American College of Rheumatology (ACR) classification criteria for RA (1). The specificity of RF is, however, often low (2–5). A newer diagnostic marker for RA is anti-citrullinated peptide antibodies (ACPAs), which can be identified by tests such as a line immunoblot (LIATM) for the detection of anti-PEP A and anti-PEP B antibodies (6), the anti-cyclic citrullinated peptide ELISA (7), an ELISA using citrullinated recombinant rat fibrilin (8), and an ELISA using deaminated fibrinogen (9). ACPAs have excellent specificity (89–100%) for RA, with good sensitivity (41–80%) (3–7, 10–15). Furthermore, the HLA shared epitope (SE) has been described, which is found more often in RA patients than in controls (16–18). Most studies of these newer tests have used control populations consisting of selected groups of patients with defined diseases and healthy controls. This does not represent real-life clinical practice because the composition of the control group does not reflect the natural prevalences of diseases in cases for which serologic markers for RA are requested. Data about specificity, positive predictive value (PPV), and negative predictive value (NPV) are thus hard to interpret. We designed the present study to reflect everyday rheumatology practice.

In this prospective study, we included 1003 consecutive patients in three academic and nonacademic centers: the Department of Rheumatology, Ghent University Hospital (Ghent, Belgium); the Locomotor Center, Elisabeth Hospital (Sijsele, Belgium); and the Department of Rheumatology, St-Augustinus Hospital (Wilrijk, Belgium). Patients were entered in the study when they presented with a new diagnostic problem for which RA was included in the differential diagnosis. This means that the rheumatologist would typically request RF determination, although the patients did not necessarily have early arthritis. Diagnoses were established after 1 year of follow-up. The clinicians were unaware of the test results obtained through the study. No ACPA results were available to the clinicians during the follow-up period. The study was conducted after receipt of approval by the local ethics committees. Oral informed consent was obtained from all patients.

RF was determined with the Waaler Rose (RF WR) and with the latex fixation method (RF LF). Anti-PEP A and anti-PEP B antibodies were detected by a research LIA (INNO-LIA; Innogenetics) (6). We used scan values to obtain continuous data. This test can also be reported as positive or negative by use of a reference strip. In that case, cutoffs cannot be varied, and the results may differ from the ones reported here. The HLA SE was determined by INNO-LiPA (line probe assay) technology (Innogenetics). Details of the methods are available in the Data Supplement that accompanies the online version of this Technical Brief at http://www.clinchem.org/content/vol50/issue12/.

We computed sensitivities, specificities, PPVs, and NPVs together with their 95% confidence intervals (CIs) (19). We compared sensitivities and specificities by use of the McNemar test. ROC curve analysis was performed. Statistical analysis was performed with SPSS 10.0.

Clinical diagnoses were established by the treating rheumatologist after 1 year of follow-up. The distribution was as follows: definite RA (n = 153), probable RA (n = 72), potential RA (n = 75), non-RA (n = 629), and lost to follow-up (n = 74). Of the patients with clinically “definite” RA, 144 fulfilled the revised ACR criteria for RA (2) and were further considered as RA patients. We used ACR criteria to improve the comparability of our results, although RF is part of these criteria and the diagnostic value of RF might thus be overestimated. However, by starting with a clinical diagnosis of RA, we minimized this problem. The non-RA patients were taken as the control group.

The mean age of the RA patients was 58.0 years (range, 21–84 years), which was significantly higher than the mean age of the non-RA patients (51.4 years; range, 12–88 years; P < 0.001). The male-to-female ratio was 50:94 in the RA group and 213:414 in the non-RA group (not significantly different). There were no significant differences between the two groups for duration of symptoms (mean of 19.3 months in the RA group and 15.9 months in the non-RA group).

We performed ROC curve analyses to compare the diagnostic accuracies of RF LF, RF WR, and anti-PEP A and anti-PEP B antibodies (Fig. 1). RF LF and RF WR had higher areas under the curves than anti-PEP A and anti-PEP B antibodies. The areas under the ROC curves were 0.84 (95% CI, 0.80–0.88) for RF LF, 0.82 (95% CI, 0.77–0.87) for RF WR, 0.78 (95% CI, 0.73–0.84) for anti-PEP A anti-
bodies, and 0.79 (95% CI, 0.74–0.84) for anti-pepB antibodies. In the high specificity region, the two ACPAs performed better than the two RF assays (Fig. 1). The sensitivities, specificities, PPVs, and NPVs (with their corresponding 95% CIs) of the serologic tests are listed in Table 1. Performances using different clinical classifications are given in Tables 1 and 2 of the online Data Supplement. At low specificity, RF had higher sensitivities than anti-pepA and anti-pepB antibodies. In contrast, when a high specificity was required, the best sensitivities were obtained for anti-pepA and anti-pepB antibodies. We will focus now on the data at a cutoff corresponding to a specificity of 98%.

To further compare sensitivities and specificities, we performed McNemar tests. Specificities did not differ significantly (all $P > 0.200$). However, at a specificity of 98%, the sensitivity of anti-pepA antibodies was significantly better than the sensitivity of RF LF, RF WR, and anti-pepB antibodies (all $P < 0.001$). The sensitivity of anti-pepB antibodies was significantly better than the sensitivity of RF WR ($P = 0.001$) but did not differ significantly from the sensitivity of RF LF ($P = 0.233$). Furthermore, the sensitivity of RF LF was better than that of RF WR ($P = 0.015$).

We also tested combinations of two serum markers (Table 3 in the online Data Supplement). All “AND” combinations had specificities of at least 98.4%, but sensitivities were low. Only the combination “anti-pepA and anti-pepB antibodies” reached a sensitivity $>40%$ (specificity, 47.6%; specificity, 98.4%). The AND combinations of one RF with one ACPA gave high specificities and very high PPVs.

Although some “OR” combinations had high sensitivities (for example, the combination “RF LF or anti-pepA” and “RF LF or anti-pepB”), this also led to a loss of specificity compared with the individual markers. The usefulness of combinations is limited because some individual serum markers already have excellent performance. However, the number of positive antibodies can modulate the degree of certainty for diagnosis. For example, when both RF LF and anti-pepA antibodies were present, we found a PPV for RA of 98.0%; in contrast, when both RF LF and anti-pepA antibodies were absent, the probability of not having RA was 92.5%. We found

![Fig. 1. ROC curve analysis for RF WR (dashed line), RF LF (thick dotted line), anti-pepA antibodies (thin solid line), and anti-pepB antibodies (thick solid line).](image)

### Table 1. Diagnostic performance of serum markers at different cutoffs chosen to obtain different specificities.

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Cutoff</th>
<th>Sensitivity (95% CI), %</th>
<th>Specificity (95% CI), %</th>
<th>PPV (95% CI), %</th>
<th>NPV (95% CI), %</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF LF</td>
<td>≥25 kilounits/L</td>
<td>68.8 (61.2–76.4)</td>
<td>90.1 (87.8–92.4)</td>
<td>61.5 (54.0–69.0)</td>
<td>92.6 (90.5–94.7)</td>
</tr>
<tr>
<td>RF WR</td>
<td>≥25 kilounits/L</td>
<td>66.7 (59.0–74.4)</td>
<td>90.9 (88.7–93.1)</td>
<td>62.7 (55.1–70.3)</td>
<td>92.2 (90.1–94.3)</td>
</tr>
<tr>
<td>Anti-pepA Ab</td>
<td>≥0.148</td>
<td>63.6 (55.7–71.5)</td>
<td>90.6 (88.3–92.9)</td>
<td>60.7 (56.0–68.5)</td>
<td>91.6 (89.4–93.8)</td>
</tr>
<tr>
<td>Anti-pepB Ab</td>
<td>≥0.169</td>
<td>64.3 (56.4–72.2)</td>
<td>89.8 (87.4–92.2)</td>
<td>59.0 (51.3–66.7)</td>
<td>91.7 (89.5–93.9)</td>
</tr>
<tr>
<td>RF LF</td>
<td>≥50 kilounits/L</td>
<td>54.9 (46.8–63.0)</td>
<td>94.6 (92.8–96.4)</td>
<td>69.9 (61.4–78.4)</td>
<td>90.1 (97.8–92.4)</td>
</tr>
<tr>
<td>RF WR</td>
<td>≥25 kilounits/L</td>
<td>56.9 (48.8–65.0)</td>
<td>95.4 (93.8–97.0)</td>
<td>73.9 (65.7–82.1)</td>
<td>90.6 (88.4–92.8)</td>
</tr>
<tr>
<td>Anti-pepA Ab</td>
<td>≥0.237</td>
<td>62.9 (55.0–70.8)</td>
<td>95.1 (93.4–96.8)</td>
<td>74.4 (66.7–82.1)</td>
<td>91.8 (89.4–93.9)</td>
</tr>
<tr>
<td>Anti-pepB Ab</td>
<td>≥0.268</td>
<td>61.5 (53.5–69.5)</td>
<td>94.9 (92.3–96.6)</td>
<td>73.3 (65.4–81.2)</td>
<td>91.6 (89.2–94.0)</td>
</tr>
<tr>
<td>RF LF</td>
<td>≥100 kilounits/L</td>
<td>41.7 (33.6–49.8)</td>
<td>97.8 (96.7–98.9)</td>
<td>81.1 (72.2–90.0)</td>
<td>88.0 (85.6–90.4)</td>
</tr>
<tr>
<td>RF WR</td>
<td>≥100 kilounits/L</td>
<td>32.6 (24.9–40.3)</td>
<td>98.6 (97.7–99.5)</td>
<td>83.9 (74.3–93.5)</td>
<td>86.5 (84.0–97.0)</td>
</tr>
<tr>
<td>Anti-pepA Ab</td>
<td>≥0.448</td>
<td>58.7 (50.6–66.8)</td>
<td>98.1 (97.0–99.2)</td>
<td>87.5 (80.9–94.1)</td>
<td>91.3 (89.2–93.4)</td>
</tr>
<tr>
<td>Anti-pepB Ab</td>
<td>≥1.215</td>
<td>48.3 (41.1–56.5)</td>
<td>98.1 (97.0–99.2)</td>
<td>85.2 (80.9–94.1)</td>
<td>89.3 (87.0–91.6)</td>
</tr>
<tr>
<td>RF LF</td>
<td>≥200 kilounits/L</td>
<td>20.8 (14.2–27.4)</td>
<td>99.0 (98.2–99.8)</td>
<td>83.3 (81.1–95.5)</td>
<td>84.5 (81.9–87.1)</td>
</tr>
<tr>
<td>RF WR</td>
<td>≥200 kilounits/L</td>
<td>18.8 (12.4–25.2)</td>
<td>99.4 (98.8–100)</td>
<td>87.1 (75.3–98.9)</td>
<td>84.2 (81.6–86.8)</td>
</tr>
<tr>
<td>Anti-pepA Ab</td>
<td>≥2.516</td>
<td>41.3 (33.2–49.4)</td>
<td>99.0 (98.2–99.8)</td>
<td>90.8 (73.8–97.8)</td>
<td>88.1 (85.8–90.4)</td>
</tr>
<tr>
<td>Anti-pepB Ab</td>
<td>≥3.407</td>
<td>37.1 (29.2–35.0)</td>
<td>99.0 (98.2–99.8)</td>
<td>89.8 (82.1–97.5)</td>
<td>87.4 (85.0–89.8)</td>
</tr>
</tbody>
</table>

*Ab, antibody.
similar PPVs and NPVs when we considered the combination of either of the two RF tests with either of the two ACPA tests. Saraux et al. (20) suggested combining two RF assays and anti-keratin antibodies (requiring positivity of at least one test). We feel that the specificity (only 82%) and PPV they obtained are insufficient to be clinically reliable. Jansen et al. (21) proposed the combination “RF or ACPA” (more precisely, IgM-RF combined with anti-cyclic citrullinated peptide antibodies), which had a specificity of 96.7%. We obtain similar specificities for OR combinations between RF and ACPA, but in this consecutive cohort, we demonstrated that allowing for such a specificity leads to a decrease in PPV.

In the RA patients, SE was absent in 36.6% (n = 52), one copy was found in 42.3% (n = 60), and two copies were found in 21.1% (n = 30). In the control group, SE was absent in 57.6% (n = 359), one copy was found in 37.2% (n = 232), and two copies were found in 5.1% (n = 32). In accordance with previous reports (22), there was a significant association between diagnosis of RA and SE status (P < 0.001). However, the high prevalence of SE in controls limits its diagnostic usefulness. Combining SE with serologic markers did not increase diagnostic performance compared with the serologic markers alone.

At least one swollen joint at baseline was found in 95.8% of the patients later classified as having RA and in 37.7% of the patients later classified as not having RA. The presence of swollen joints thus reached a PPV for RA of 36.9%, whereas the absence of swollen joints had a NPV of 98.5% for RA. We analyzed whether the presence of swollen joints at baseline could be combined with serologic markers to predict RA. The AND combinations revealed high specificities and PPVs (Table 4 in the online Data Supplement). The combination of anti-pepA antibodies and swollen joints also remained fairly sensitive. When the pre-test probability for RA is increased by clinical findings, serum markers with lower specificities (i.e., using lower cutoff values) may become relevant.

In summary, in a setting reflecting everyday rheumatologic practice for prevalence of RA, we evaluated the diagnostic performance of two classic RF assays and two methods for ACPA detection. At high specificities, the anti-pepA antibodies had the best sensitivity. Combining one of the RF tests with one of the ACPA tests increased PPV. Combining one serologic marker with the finding of swollen joints also provides a high PPV.

Ilse Hoffman is supported by a research grant from the “Bijzonder Onderzoeksfonds” Ghent University. The study was sponsored by Innogenetics (Ghent, Belgium). We wish to thank Jeanine Discart, Liesbeth Amerijckx, and Ann Vos for technical assistance, and Katrien De Bosschere for careful reading of the manuscript.

References


Previously published online at DOI: 10.1373/clinchem.2004.034728