RHEUMATOLOGY

Original article

doi:10.1093/rheumatology/kex170

A multicentre study to improve clinical interpretation of proteinase-3 and myeloperoxidase anti-neutrophil cytoplasmic antibodies

Xavier Bossuyt[1,2], Niels Rasmussen[3], Pieter van Paassen[4], Bernard Hellmich[5,6], Bo Baslund[7], Pieter Vermeersch[2,8], Daniel Blockmans[1,9], Jan-Willem Cohen Tervaert[10], Elena Csernok[5,6,*] and Jan Damoiseaux[11,*]

Abstract

Objective. The objective of this multicentre study was to improve the clinical interpretation of PR3- and MPO-ANCAs as an adjunct for the diagnosis of ANCA-associated vasculitis (AAV) by defining thresholds and test result intervals based on predefined specificities and by calculating test result interval-specific likelihood ratios (LRs).

Methods. Eight different PR3- and MPO-ANCA immunoassays from seven companies were evaluated using 251 diagnostic samples from AAV patients and 924 diseased controls.

Results. Thresholds for antibody levels were determined based on predefined specificities (95, 97.5, 99 and 100%) and used to delimit test result intervals. Test result interval-specific LRs were determined. For all assays, the LR for AAV increased with increasing antibody level. For all but one immunoassay, high antibodies levels (associated with LR >55) were found in a substantial fraction (>65%) of patients. The area under the curve (AUC) of receiver operating characteristics analysis of a diagnostic approach in which positive results were confirmed by IIF or another immunoassay was not substantially higher than the AUC of performing immunoassay only. The highest AUC was found when immunoassay was combined with another immunoassay or with IIF.

Conclusion. To diagnose AAV based on PR3- and MPO-ANCA, it is useful to define thresholds for antibody levels and to assign test result interval-specific LRs. Higher antibody levels are associated with a higher likelihood for disease. Such information improves clinical interpretation.

Key words: anti-neutrophil cytoplasmic antibodies, ANCA, proteinase-3, myeloperoxidase, PR3-ANCA, MPO-ANCA, vasculitis, granulomatosis with polyangiitis, microscopic polyangiitis

CLINICAL SCIENCE

Rheumatology key messages

- Defining thresholds for antibody levels improves clinical interpretation of PR3 and MPO-ANCA.
- The likelihood ratio for ANCA-associated vasculitis increases with increasing PR3 and MPO-ANCA antibody level.

[1]Department of Microbiology and Immunology, KU Leuven, [2]Laboratory Medicine, University Hospitals Leuven, Leuven, Belgium, [3]Department of Autoimmunology and Biomarkers, Statens Seruminstitute, Copenhagen, Denmark, [4]Department of Internal Medicine, Section Nephrology and Immunology, Maastricht University Medical Center, Maastricht, The Netherlands, [5]Department of Internal Medicine, Rheumatology and Immunology, University Teaching Hospital Kirchheim, [6]Vasculitis-Center Tübingen-Kirchheim, Kirchheim, Germany, [7]Department of Rheumatology, Rigshospitalet, Copenhagen, Denmark, [8]Department of Cardiovascular Sciences, KU Leuven, [9]General Internal Medicine, University Hospitals Leuven, Leuven, Belgium, [10]Faculty of Health, Medicine and Life Sciences, Maastricht University and [11]Central Diagnostic Laboratory, Maastricht University Medical Center, Maastricht, The Netherlands

Submitted 1 November 2016; revised version accepted 24 March 2017

*Elena Csernok and Jan Damoiseaux contributed equally to this study

Correspondence to: Xavier Bossuyt, Laboratory Medicine, University Hospitals Gasthuisberg, Herestraat 49, 3000 Leuven, Belgium. E-mail: xavier.bossuyt@uzleuven.be

Introduction

ANCAs are associated with small-vessel vasculitis, including granulomatosis with polyangiitis (GPA) and microscopic polyangiitis (MPA; for recent reviews, see [1–3]). Traditionally, IIF on ethanol-fixed cells is used to screen for ANCA, and positive IIF results are confirmed by immunoassays that are specific for PR3 and MPO. A large European Vasculitis Study Group (EUVAS) multicentre study, in which various ANCA detection methods were evaluated, revealed high variability between IIF methods and a high performance of PR3-ANCA and MPO-ANCA as an adjunct for diagnosis of ANCA-associated vasculitis (AAV) [4]. Based on these findings and the broad availability of (automated) immunoassays for PR3- and MPO-ANCA, it is expected that immunoassays will be used for assisting in the diagnosis of AAV, without IIF.

As it is highly likely that immunoassays will be increasingly used as the main and only laboratory tool for assisting in the diagnosis of AAV, it is important to extract a maximum of clinically relevant information from test results generated by such assays. Most studies that have addressed the clinical usefulness of PR3-ANCA and MPO-ANCA for the diagnosis of AAV typically used a single cut-off value. However, a lot of information is lost when only binary results (positive/negative) are considered. It is reasonable to assume that the likelihood for AAV increases with increasing antibody levels of PR3- and MPO-ANCA [5].

In the present study, we used the large data set from the multinational EUVAS study [4] to document how the likelihood ratio (LR) for AAV depends on the antibody level. In particular, we defined thresholds and test result intervals based on predefined specificities (fractions of disease controls that test negative) and calculated test result interval-specific LRs. Such knowledge might add value to a specific test result. The concept of LR (probability of a specific result in patients divided by the probability of the same result in controls) was used because it provides information that is independent of prevalence or pre-test probability [6].

Besides, we also addressed the question of whether a single set of immunoassays (PR3-ANCA and MPO-ANCA) is sufficient for aiding in the diagnosis of AAV or whether additional laboratory testing (e.g. IIF or another set of immunoassays) is warranted.

Methods

Patients

The patients included were from a multicentre EUVAS study described by Damoiseaux *et al.* [4]. In short, patients with AAV and controls with inflammatory disease were recruited at Klinikum Bad Bramstedt (BB; Germany), Statens Serum Institute Copenhagen (Denmark), University Hospitals Leuven (Belgium) and Maastricht University Medical Center (The Netherlands). The study group included 186 patients with GPA, 65 patients with MPA and 924 disease controls. GPA patients fulfilled the ACR classification criteria [7] and the Chapel Hill Consensus Definitions [8, 9]. The diagnosis of MPA was based on Chapel Hill Consensus Definitions [8, 9]. The samples included were diagnostic samples, i.e. obtained at the time the diagnosis was established. The disease controls recruited in Copenhagen, Leuven and Maastricht (n = 735) were consecutive patients in whom ANCA analysis was requested but in whom AAV was later excluded. Patients in whom IBD, autoimmune liver disease, or both were considered were excluded. The disease controls in BB included cohorts of patients with SLE (n = 59), RA (n = 89), SSc (n = 11) and SS (n = 30). More detailed clinical information on patients and controls is given by Damoiseaux *et al.* [4]. For analysis, patients with GPA and MPA were grouped as AAV. This study was approved by the ethics committee of each participating centre (ethics committees of the University Hospitals Leuven, Maastricht University Medical Center, Klinikum Bad Bramstedt and Statens Seruminstitute). As the study was performed on left-over samples (secondary use), no informed consent was needed.

ANCA detection methods

Eight antigen-specific ANCA assays from seven commercial manufacturers were included: QuantaLite ELISA and QuantaFlash automated chemiluminescence assay from Inova Diagnostics (San Diego, CA, USA); EliA $PR3^S$ and EliA MPO^S fluorescence enzyme immunoassays from Thermo-Fisher Scientific (Waltham, MA, USA); MPO and PR3 multiplex on BioPlex 2200 Vasculitis kit from Bio-Rad Laboratories Inc. (Hercules, CA, USA); a second-generation capture PR3- and MPO-ANCA ELISA from Euro-Diagnostica AB (Malmö, Sweden); third-generation anti-PR3 h-ELISA and first-generation anti-MPO ELISA from Orgentec (Mainz, Germany); third-generation anti-PR3-hn-hr-ELISA and first-generation MPO-ANCA from Euroimmun AG (Lübeck, Germany); and CytoBead ANCA assays from Medipan/Generic Assays GmbH (Berlin, Germany). The tests were performed according to the manufacturers' instructions, and results were expressed in the individual kit units.

The highest level of reactivity from the PR3- and MPO-ANCA determinations was selected for analysis. This was feasible because all manufacturers, except for two, apply similar cut-off values for PR3- and MPO-ANCA. For the two manufacturers that do not apply the same cut-off for PR3- and MPO-ANCA, the differences were minor, that is, 3 *vs* 5 IU/ml for EliA and 5 *vs* 10 U/ml for the Orgentec ELISAs.

IIF was performed at Bad Bramstedt (Germany) on ethanol-fixed neutrophils in combination with additional tests on formalin-fixed neutrophils and HEp-2 cells to better discriminate between P-ANCA (or atypical-ANCA) and ANA, as previously described (Csernok and Moosig [3]). The highest level of reactivity, independent of ANCA pattern, was selected for analysis.

Statistical analysis

Receiver operating characteristic curve analysis was performed using Medcalc (Version 16.1; Medcals Software

bvba, Medcalc Ostend, Belgium). Comparisons of the area under the curve (AUC) were performed using Analyse-it for Microsoft Excel 3.90 (method of Hanley and McNeil) and Medcalc (Version 16.1). For AUC, 95% CIs from Medcalc are given. LRs (and 95% CIs) were calculated with Medcalc (Version 16.1).

Results

Sensitivity and LR at defined specificity

For each immunoassay, threshold values that corresponded to a specificity of 95, 97.5, 99 and 100% (if applicable) were determined. The corresponding sensitivities and LRs for AAV were calculated and are presented in Table 1. The sensitivities, specificities and LRs for the cut-off value proposed by the manufacturer are also given in Table 1.

The LRs for a positive test result increased with increasing threshold value and were comparable between the different assays. They were ~18, 32–35 and 66–82 for a threshold that corresponded to a specificity of 95, 97.5 and 99%, respectively. The sensitivity (fraction of patients that test positive) that corresponded to a specificity of 95% varied between 84 and 92%, whereas the sensitivity that corresponded to a specificity of 99% varied between 64 and 80%, depending on the assay. For the CytoBead assay, the LR for a negative test result at a threshold that corresponded to a specificity of 95% was 0.17. This was higher than for the other assays (0.09–0.12) and is related to a lower sensitivity (84%, *vs* 88–92% for the other assays).

Test result interval-specific LR

Next, we used the threshold values that corresponded to a specificity of 95, 97.5, 99 and 100% (if applicable) to define test result intervals. For each test result interval, the interval-specific LR for AAV was calculated, as well as the fraction of patients and controls that had a result within this interval. The results are presented in Table 2. The LR for the lowest test result interval (<threshold that defines a specificity of 95%) varied between 0.09 and 0.12, except for the CytoBead assay, for which the LR was 0.17.

The LR for the next test result interval (i.e. values between the thresholds that define specificities of 95 and 97.5%) varied between 0.64 and 3.36. Most of the corresponding 95% CIs included 1, indicating that the LRs of those intervals (95–97.5% specificities) were not proved to be different from 1 for most of the tests. A small proportion (1.6–8.3%) of patients had a result within this interval.

The LR for the next test result interval (i.e. values between thresholds that define specificities of 97.5 and 99%) varied between 3.4 and 14.7, and 5–23% of patients had a test result within this interval.

The subsequent test result interval (i.e. values between the thresholds that define specificities of 99 and 100%) contained the majority of the patients (57–75%, except for the CytoBead assay, for which it was 36%). The LR associated with this interval varied between 56 and 77, except for the CytoBead assay, for which it was 37. With all assays except BioPlex and Euroimmun, a threshold could be defined above which no controls were found positive (i.e. LR of $+\infty$). Depending on the assay, 1.6–29% of the patients had a value within this interval.

For two immunoassays (QuantaFlash and the EliA), a breakdown of the test result interval-specific LRs for each of the four participating centres is shown in supplementary Table S1, available at *Rheumatology* Online. This shows that, although there are differences between the centres, the overall trend between the different centres is similar.

Post-test probabilities can be estimated based on pre-test probabilities and LRs [6]. Figure 1 illustrates the post-test probabilities as a function of pre-test probability for different test result intervals (exemplified for EliA and capture ELISA from Euro-Diagnostica). On the graph, one can easily read the post-test probability for AAV for any given pre-test probability and PR3-ANCA/MPO-ANCA test result. For example, for a pre-test probability of 21%, which corresponds to the prevalence of pauci-immune crescentic glomerulonephritis in patients with haematuria, proteinuria and creatinine >3 mg/dl (obtained from Jennette *et al.* [8]), the post-test probability for AAV was 3, 25, 48, 95 and 100%, for a test result (obtained by assays from Euro-Diagnostica) of, respectively, <2.8 IU/l, between 2.8 and 4.5 IU/ml, between 4.5 and 12.2 IU/ml, between 12.2 and 247.0 IU/ml and >247.0 IU/ml. The results are summarized in Table 3. This table also illustrates how the post-test probabilities depend on the antibody level for several other clinical presentations (e.g. sinus and pulmonary involvement and different levels of renal impairment). The pre-test probabilities for AAV associated with the clinical presentations were obtained from the literature [10–12].

Comparison of diagnostic strategies

The test result intervals as outlined above were applied to classify and categorize the results and to calculate the AUC for the eight different immunoassays. The AUC of the immunoassays was compared with the AUC of the IIF methods performed in BB and in Copenhagen (C). The results are summarized in Table 4.

The AUC of IIF-BB was significantly higher than the AUC of IIF-C (see also [4]). The AUC of all immunoassays was higher than the AUC of IIF-C (data not shown; see [4]), whereas the AUC of BioPlex and Euroimmun immunoassays was significantly higher than the AUC of IIF-BB (Table 4). The AUC of the CytoBead immunoassay was significantly lower than the AUC of most other immunoassays.

The AUCs of all immunoassays were statistically significantly higher than the AUCs of screening with IIF-C and confirming a positive IIF-C result with the respective immunoassays (Table 4). The AUCs of all but one of the immunoassays were marginally higher than the AUCs of screening with IIF-BB and confirming with the respective immunoassay, but these differences were not statistically significant (Table 4).

Table 1 Sensitivity and likelihood ratios for specified specificities

Manufacturer	Threshold	Sensitivity (95% CI)	Specificity (95% CI)	LR (+) (95% CI)	LR (–) (95% CI)
Cut-off manufacturer					
ELISA Quantalite, Inova	>10.0	90.4 (86.1, 93.8)	95.0 (93.4, 96.3)	18.2 (13.7, 24.1)	0.1 (0.07, 0.1)
20 units	>19	86.5 (81.6, 90.4)	97.5 (96.3, 98.4)	34.7 (23.1, 52.2)	0.14 (0.1, 0.2)
	>37.0	77.3 (71.6, 82.3)	99.0 (98.2, 99.6)	79.4 (41.3, 152.5)	0.23 (0.2, 0.3)
	>159.1	17.1 (12.7, 22.4)	100.0 (99.6, 100.0)	∞	0.83 (0.8, 0.9)
	>20.3	86.5 (81.6, 90.4)	97.9 (96.8, 98.8)	42.0 (26.9, 65.8)	0.14 (0.1, 0.2)
CLIA QuantaFlash, Inova	>12.5	90.0 (85.6, 93.5)	94.9 (93.3, 96.2)	17.7 (13.4, 23.5)	0.1 (0.07, 0.2)
20 CU	>23.8	87.3 (82.5, 91.1)	97.5 (96.3, 98.4)	35.1 (23.3, 52.6)	0.13 (0.09, 0.2)
	>74.3	72.1 (66.1, 77.6)	99.0 (98.2, 99.6)	74.0 (38.5, 142.5)	0.28 (0.2, 0.3)
	>1049.8	8.8 (5.6, 13.0)	100.0 (99.6, 100.0)	∞	0.91 (0.9, 0.9)
	>19.3	88.5 (83.8, 92.1)	96.7 (95.3, 97.7)	26.4 (18.6, 37.4)	0.12 (0.08, 0.2)
FEIA EliA, Thermo-Fisher	>2.1	90.0 (85.6, 93.5)	95.0 (93.4, 96.3)	18.1 (13.6, 24.0)	0.1 (0.07, 0.2)
PR3: equivocal 2–3 IU/ml	>5	81.7 (76.3, 86.3)	97.5 (96.3, 98.4)	32.8 (21.8, 49.3)	0.19 (0.1, 0.2)
MPO: equivocal 3.5–5 IU/ml	>16	63.8 (57.5, 69.7)	99.0 (98.2, 99.6)	65.4 (33.9, 126.2)	0.37 (0.3, 0.4)
	>142	6.8 (4.0, 10.6)	100.0 (99.6, 100.0)	∞	0.93 (0.9, 1.0)
	>5	81.7 (76.3, 86.3)	97.5 (96.3, 98.4)	32.8 (21.8, 49.3)	0.19 (0.1, 0.2)
BioPlex 2200, BioRad	>0.5	90.8 (86.6, 94.1)	95.0 (93.4, 96.3)	18.3 (13.7, 24.2)	0.096 (0.07, 0.1)
1 AI	>1.3	88.5 (83.8, 92.1)	97.4 (96.2, 98.3)	34.1 (22.9, 50.7)	0.12 (0.08, 0.2)
	>6.2	64.9 (58.7, 70.8)	99.0 (98.2, 99.6)	66.7 (34.6, 128.5)	0.35 (0.3, 0.4)
	>8	0.0 (0.0, 1.5)	100.0 (99.6, 100.0)	∞	1 (1.0, 1.0)
	>1	89.6 (85.2, 93.1)	96.0 (94.5, 97.2)	22.4 (16.3, 30.8)	0.11 (0.07, 0.2)
ELISA Euro-Diagnostica	>2.8	88.1 (83.4, 91.8)	95.1 (93.5, 96.4)	18.1 (13.5, 24.1)	0.13 (0.09, 0.2)
PR3: equivocal 5–7 IU/ml	>4.5	84.9 (79.8, 89.1)	97.4 (96.2, 98.3)	32.7 (21.9, 48.7)	0.16 (0.1, 0.2)
MPO: equivocal: 5–7 IU/ml	>11.9	80.1 (74.6, 84.8)	99.0 (98.2, 99.6)	82.2 (42.8, 158.0)	0.2 (0.2, 0.3)
	>246.9	8.8 (5.6, 13.0)	100.0 (99.6, 100.0)	∞	0.91 (0.9, 0.9)
	>7.4	84.5 (79.4, 88.7)	98.3 (97.2, 99.0)	48.8 (29.9, 79.5)	0.16 (0.1, 0.2)
ELISA Orgentec	>4.3	88.1 (83.4, 91.8)	94.9 (93.3, 96.2)	17.3 (13.1, 23.0)	0.13 (0.09, 0.2)
PR3: 10 U/ml	>7.9	86.5 (81.6, 90.4)	97.5 (96.3, 98.4)	34.7 (23.1, 52.2)	0.14 (0.1, 0.2)
MPO: 5 U/ml	>22.4	76.1 (70.3, 81.2)	99.0 (98.2, 99.6)	78.1 (40.6, 150.2)	0.24 (0.2, 0.3)
	>230	1.6 (0.4, 4.0)	100.0 (99.6, 100.0)	∞	0.98 (1.0, 1.0)
	>10.3	84.1 (78.9, 88.4)	98.3 (97.2, 99.0)	48.6 (29.8, 79.1)	0.16 (0.1, 0.2)
ELISA Euroimmun	>6.9	91.6 (87.5, 94.7)	94.7 (93.0, 96.1)	17.3 (13.1, 22.7)	0.088 (0.06, 0.1)
20 U/ml	>26.6	87.3 (82.5, 91.1)	97.5 (96.3, 98.4)	35.1 (23.3, 52.6)	0.13 (0.09, 0.2)
	>115.7	64.1 (57.9, 70.1)	99.0 (98.2, 99.6)	65.9 (34.1, 127.0)	0.36 (0.3, 0.4)
	>200	0.0 (0.0, 1.5)	100.0 (99.6, 100.0)	∞	1 (1.0, 1.0)
	>21.7	89.2 (84.7, 92.8)	97.0 (95.6, 98.0)	29.5 (20.4, 42.5)	0.11 (0.08, 0.2)
CytoBead	>4.6	84.1 (78.9, 88.4)	95.0 (93.4, 96.3)	16.9 (12.7, 22.5)	0.17 (0.1, 0.2)
Equivocal 4.5–5 IU/ml	>5.6	80.5 (75.0, 85.2)	97.5 (96.3, 98.4)	32.3 (21.5, 48.6)	0.2 (0.2, 0.3)
	>14.2	65.3 (59.1, 71.2)	99.0 (98.2, 99.6)	67.1 (34.8, 129.3)	0.35 (0.3, 0.4)
	>68.2	29.1 (23.5, 35.1)	100.0 (99.6, 100.0)	∞	0.71 (0.7, 0.8)
	>5	82.1 (76.8, 86.6)	97.0 (95.6, 98.0)	27.1 (18.7, 39.2)	0.18 (0.1, 0.2)

Sensitivity and LRs (with 95% CIs) that correspond to a specificity of 95, 97.5 and 99% (and 100%) for eight different immunoassays for PR3- and MPO-ANCA. The sensitivity, specificity and LRs for a threshold that is close to the cut-off point proposed by the manufacturer are also given. The highest level of reactivity from the PR3- and MPO-ANCA determinations was selected for analysis. LR: likelihood ratio.

Subsequently, we evaluated whether confirming a positive immunoassay test result (using the threshold that corresponds to a specificity of 95%) with IIF or with another immunoassay adds diagnostic value and whether combining the immunoassay with IIF or with another immunoassay on all samples adds diagnostic value. As examples of a second immunoassay, we selected Euroimmun and QuantaFlash, two assays with a high AUC for AAV. The results are summarized in Table 4.

The AUC of confirming a positive immunoassay test result by IIF or by another immunoassay was not significantly higher than the AUC of performing only the immunoassay (i.e. without confirming a positive immunoassay). Overall, the highest AUC was found for the approach in which two immunoassays were performed on all samples (which gives information on double positivity, single positivity and double negativity). This reached a statistically significant difference for combining Euroimmun or QuantaFlash with Euro-Diagnostica, Orgentec or Medipan (CytoBead). Supplementary Fig. S1, available at *Rheumatology* Online, illustrates the receiver operating characteristic curves for the combination

Table 2 Test result-specific likelihood ratios

Manufacturer cut-off	Interval	Fraction of patients	Fraction of controls	LR	95% CI
QuantaLite, Inova	0.0, 10.1	0.099	0.951	0.10	0.07, 0.15
20 Units	10.1, 19.0	0.036	0.024	1.50	0.70, 3.22
	19.0, 37.0	0.091	0.014	6.49	3.34, 12.64
	37.0, 159.1	0.603	0.011	55.79	29.88, 104.19
	159.1, 250.0	0.171	0.000	∞	19.50, ∞
QuantaFlash, Inova	0.0, 12.5	0.095	0.949	0.10	0.07, 0.15
20 CU	12.5, 23.8	0.032	0.026	1.22	0.56, 2.70
	23.8, 78.2	0.155	0.015	10.23	5.64, 18.53
	78.2, 1049.8	0.631	0.010	64.85	33.62, 125.09
	1049.8, 3500	0.087	0.000	∞	9.82, ∞
FEIA EliA, Thermo-Fisher	0.0, 2.1	0.099	0.950	0.10	0.07, 0.15
PR3: equivocal 2–3 IU/l	2.1, 5.0	0.083	0.025	3.35	1.89, 5.96
MPO: equivocal 3.5–5 IU/l	5.0, 16.0	0.179	0.015	11.80	6.58, 21.14
	16.0, 142.0	0.571	0.010	58.73	30.39, 113.50
	142.0, 180.0	0.067	0.000	∞	7.52, ∞
BioPlex 2200, BioRad	0.0, 0.5	0.091	0.950	0.10	0.06, 0.14
1 AI	0.5, 1.3	0.024	0.024	1.00	0.41, 2.44
	1.3, 6.2	0.234	0.016	14.44	8.33, 25.01
	6.2, 8.0	0.651	0.010	66.89	34.69, 128.95
ELISA Euro-Diagnostica	0.0, 2.8	0.119	0.951	0.13	0.09, 0.17
Equivocal 5–7 IU/ml	2.8, 4.5	0.028	0.023	1.22	0.53, 2.85
	4.5, 11.9	0.052	0.016	3.18	1.53, 6.60
	11.9, 247.0	0.710	0.010	73.01	37.92, 140.53
	247.0, 3000.0	0.091	0.000	∞	10.28, ∞
ELISA Orgentec	0.0, 4.4	0.119	0.950	0.13	0.09, 0.18
PR3: 10 U/ml	4.4, 7.9	0.016	0.025	0.64	0.22, 1.83
MPO: 5 U/ml	7.9, 22.4	0.103	0.015	6.82	3.61, 12.86
	22.4, 230.0	0.746	0.010	76.68	39.86, 147.48
	230.0, 450.0	0.016	0.000	∞	1.56, ∞
ELISA Euroimmun	0.0, 7.4	0.083	0.949	0.09	0.06, 0.13
20 U/ml	7.4, 26.7	0.052	0.026	1.99	1.03, 3.85
	26.7, 115.7	0.222	0.015	14.68	8.31, 25.93
	115.7, 200.0	0.643	0.010	66.07	34.26, 127.40
CytoBead	0.0, 4.6	0.163	0.948	0.17	0.13, 0.23
Equivocal 4.5–5 IU/ml	4.6, 5.6	0.036	0.027	1.32	0.63, 2.79
	5.6, 14.3	0.151	0.015	9.96	5.49, 18.09
	14.3, 68.2	0.361	0.010	37.11	18.98, 72.51
	68.2, 180.0	0.290	0.000	∞	33.32, ∞

Threshold values that corresponded to a specificity of 95, 97.5, 99 and 100% (if applicable; see Table 1) were used to define test result intervals. For each test result interval, the interval-specific LR and the fraction of patients and controls that had a result within this interval are shown. The highest level of reactivity from the PR3- and MPO-ANCA determinations was selected for analysis. The 95% CIs for the LRs were computed using Medcalc. LR: likelihood ratio.

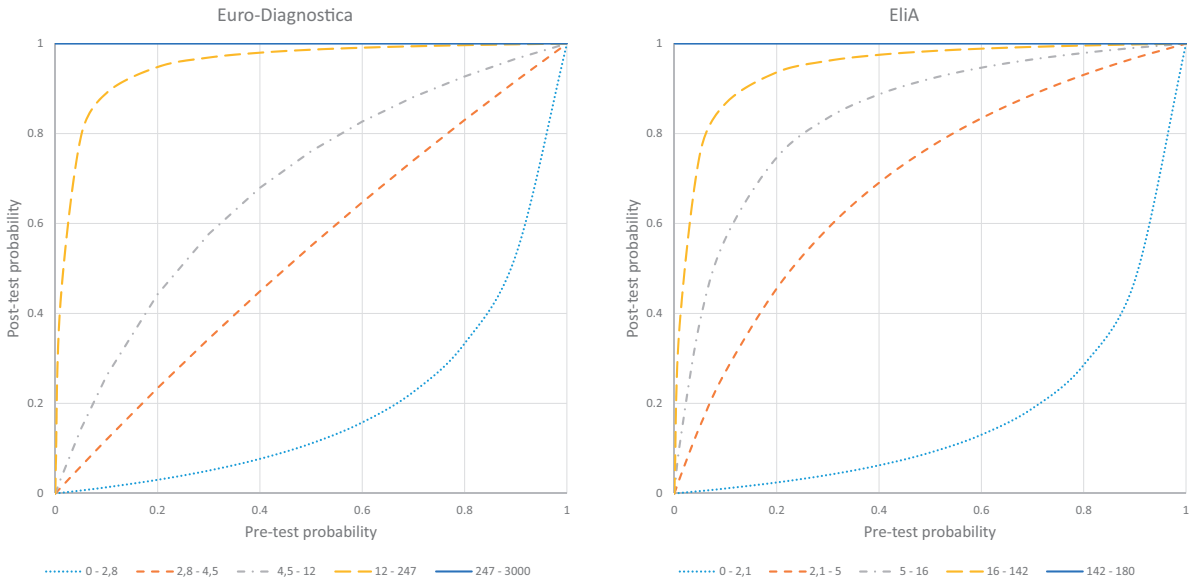
of Orgentec ELISA with IIF or Euroimmun ELISA or QuantaFlash. Supplementary Table S2, available at *Rheumatology* Online, shows the distribution of patients and controls across the combination of intervals for Orgentec and Euroimmun or IIF. Supplementary Fig. S2, available at *Rheumatology* Online, shows the AUC for the combination of Euro-Diagnostica ELISA with IIF or Euroimmun ELISA or QuantaFlash.

Discussion

PR3-ANCA and MPO-ANCA are reliable laboratory tests [4], and it is expected that they will be used as an adjunct for the diagnosis of AAV without the use of IIF [4]. As it can be anticipated that PR3-ANCA and MPO-ANCA will be increasingly used as a screening tool, it is important to gain the best possible understanding of their diagnostic potential. We evaluated the diagnostic value of eight different immunoassays using a large cohort of diagnostic samples from patients with AAV and a large disease control group. We propose the concept of test result interval-specific LRs to document the diagnostic accuracy. LRs are independent of disease prevalence and can be applied to test result intervals [6].

In order to define test result intervals, one could use the cut-off proposed by the manufacturer and multiples of the cut-off to define intervals. This, however, has the disadvantage that the patient and control populations used to

Fig. 1 Post-test probability as a function of pre-test probability



Post-test probability (*y*-axis) as a function of pre-test probability (*x*-axis) for different test result intervals. The data are given for the capture ELISA from Euro-Diagnostica (in international units per millilitre) and for EliA (in international units per millilitre).

Table 3 Post-test probability for ANCA-associated vasculitis as a function of defined pre-test probabilities and specific test result

Clinical condition	Pre-test probability	Post-test probability				
		0.0–2.8 IU/ml	2.8–4.5 IU/ml	4.5–12.2 IU/ml	12.2–247.0 IU/ml	247.0–3000.0 IU/ml
None[a]	0.00003	<0.001	<0.001	<0.001	0.002	1.00
Sinus[a]	0.01	0.00	0.01	0.03	0.40	1.00
Sinus and pulmonary[a]	0.10	0.01	0.12	0.27	0.88	1.00
Sinus and glomerulonephritis[a]	0.30	0.05	0.35	0.59	0.97	1.00
Sinus, pulmonary and glomerulonephritis[a]	0.85	0.42	0.87	0.95	1.00	1.00
Rapidly progressive glomerulonephritis[b]	0.47	0.10	0.52	0.75	0.98	1.00
Haematuria, proteinuria, creatinine >3 mg/dl[b]	0.21	0.03	0.25	0.48	0.95	1.00
Haematuria, proteinuria, creatinine 1.5–3 mg/dl[b]	0.07	0.01	0.08	0.20	0.83	1.00
Haematuria, proteinuria, creatinine <1.5 mg/dl[b]	0.02	0.003	0.02	0.07	0.57	1.00

Test results obtained with assays from Euro-Diagnostica are shown. [a]Pre-test probabilities from Langford [10]; sinus: radiographic evidence of mucosal thickening involving one or more sinuses; pulmonary: radiographic presence of pulmonary infiltrates or nodules, or both; glomerulonephritis: urinalysis demonstrating haematuria and red blood cell casts. [b]Pre-test probabilities from Jennette *et al.* [8]; the conditions are for an adult patient (>18 years old).

establish the cut-off and the criteria applied differ between the manufacturers. In order to overcome these limitations, we defined test result intervals based on predefined specificities (95, 97.5, 99 and 100%) calculated for each assay from our cohort using the same large disease control population. This has the advantage that the same criteria were used across the different assays. In that way, we maximally harmonized clinical interpretation between the different assays and companies. For most assays, the cut-off value proposed by the manufacturer had a specificity between 95 and 97.5%. Test result intervals were based on predefined specificities rather than on predefined sensitivities in order to secure a minimal level of specificity (95%).

The LR for AAV increased with increasing antibody levels. For all assays, at least 75% of the patients with AAV had a test result with a corresponding LR >~10 (which indicates a clinically significant difference in

Table 4 Comparison of diagnostic strategies

Test diagnostic approach		1	2	3	4	5	6	7	8	IIF-BB	IIF-C
IA	AUC	0.942 (0.927, 0.955)	0.943 (0.928, 0.955)	0.938 (0.923, 0.951)	0.949 (0.935, 0.961)	0.932 (0.916, 0.945)	0.931 (0.915, 0.945)	0.947 (0.933, 0.960)	0.908 (0.890, 0.924)		
IIF	AUC									0.923 (0.906, 0.938)	0,842 (0,820, 0,863)
Comparison IIF-BB *vs* IIF-C	P-value										<0.0001
Comparison IA *vs* IIF-BB	P-value	0.14	0.11	0.20	0.04	0.51	0.55	0.04	0.40		
Comparison CytoBead *vs* other IA	P-value	0.03	0.02	0.06	0.006	0.16	0.17	0.008			
IIF-BB, if IIF(+) then IA	AUC	0.929 (0.913, 0.943)	0.928 (0.911, 0.942)	0.927 (0.911, 0.941)	0.928 (0.911, 0.942)	0.928 (0.912, 0.942)	0.928 (0.911, 0.942)	0.928 (0.912, 0.942)	0.928 (0.912, 0.942)		
Comparison *vs* IA	P-value	0.28	0.20	0.34	0.07	0.78	0.82	0.08	0.20		
IIF-C, if IIF(+) then IA	AUC	0.860 (0.839, 0.879)	0.859 (0.838, 0.879)	0.859 (0.838, 0.878)	0.859 (0.838, 0.879)	0.859 (0.838, 0.879)	0.859 (0.838, 0.879)	0.859 (0.838, 0.879)	0.859 (0.838, 0.878)		
Comparison *vs* IA	P-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0004	0.004		
IA, if IA(+) then IIF-BB (for all conditions)	AUC	0.942 (0.927, 0.955)	0.942 (0.928, 0.955)	0.938 (0.923, 0.952)	0.950 (0.936, 0.962)	0.931 (0.915, 0.945)	0.931 (0.914, 0.944)	0.948 (0.934, 0.960)	0.909 (0.892, 0.925)		
Comparison *vs* IA	P-value	0.99	0.98	0.91	0.83	0.96	0.95	0.83	0.63		
IA, if IA(+) then other IA (Euroimmun)	AUC	0.942 (0.927, 0.955)	0.943 (0.929, 0.956)	0.940 (0.925, 0.953)	0.950 (0.936, 0.962)	0.932 (0.916, 0.946)	0.931 (0.915, 0.945)		0.912 (0.892, 0.926)		
Comparison *vs* IA	P-value	0.99	0.87	0.65	0.76	0.94	0.95		0.52		
IA, if IA(+) then other IA (QuantaFlash)	AUC	0.942 (0.928, 0.955)		0.940 (0.925, 0.953)	0.951 (0.937, 0.962)	0.932 (0.916, 0.946)	0.932 (0.916, 0.946)	0.949 (0.935, 0.961)	0.912 (0.893, 0.926)		
Comparison *vs* IA	P-value	0.84		0.56	0.62	0.91	0.85	0.46	0.34		
IA and IIF-BB on all samples	AUC	0.944 (0.929, 0.956)	0.950 (0.936, 0.962)	0.946 (0.932, 0.959)	0.956 (0.942, 0.967)	0.948 (0.934, 0.960)	0.938 (0.922, 0.951)	0.950 (0.936, 0.962)	0.940 (0.925, 0.953)		
Comparison *vs* IA	P-value	0.81	0.31	0.25	0.32	0.07	0.45	0.69	0.0004		
IA and other IA (Euroimmun) on all samples	AUC	0.952 (0.938, 0.964)	0.952 (0.938, 0.964)	0.949 (0.935, 0.961)	0.953 (0.940, 0.965)	0.954 (0.940, 0.965)	0.951 (0.937, 0.962)		0.949 (0.935, 0.961)		
Comparison *vs* IA	P-value	0.30	0.18	0.13	0.45	0.01	0.02		<0.0001		
IA and other IA (QuantaFlash) on all samples	AUC	0.949 (0.935, 0.961)		0.948 (0.934, 0.960)	0.954 (0.940, 0.965)	0.950 (0.936, 0.962)	0.948 (0.933, 0.960)	0.952 (0.938, 0.964)	0.949 (0.935, 0.961)		
Comparison *vs* IA	P-value	0.27		0.09	0.39	0.02	0.04	0.41	<0.0001		

Area under the curve (AUC; with 95% CI) of eight different immunoassays (IA), IIF performed at Bad Bramstedt (IIF-BB) and Copenhagen (IIF-C), and various combinations of tests for diagnosis of ANCA-associated vasculitis. The combinations included screening with IIF-BB and immunoassay if IIF-BB was positive, screening with immunoassay and if immunoassay was positive (+) (using the threshold that corresponds to a specificity of 95%) then IIF-BB or another immunoassay (either Euroimmun or QuantaFlash) was performed or immunoassay and IIF-BB or another immunoassay (either Euroimmun or QuantaFlash) was performed on all samples. The various combinations were compared with immunoassay. Statistical analysis of comparisons between AUCs was performed with the method of Hanley and McNeil (Medcalc). The immunoassays included were from Inova [QuantaLite (1) and QuantaFlash (2)], Thermo-Fisher (EliA) (3), Bio-Rad (BioPlex 2200) (4), Euro-Diagnostica (5), Orgentec (6), Euroimmun (7) and Medipan (CytoBead assay) (8).

pre-test to post-test probability for disease). For all assays, except for CytoBead, at least 55% of the patients had a test result with a corresponding LR >55. This indicates that for a substantial fraction of patients PR3-ANCA/MPO-ANCA test results substantially affect the post-test probability of disease.

For almost all assays, the test result interval delimited by thresholds that define specificities of 95 and 97.5% had a LR of ±1, indicating no difference between pre-test and post-test probability. Values lower than the threshold that defined a specificity of 95% had a low LR (±0.10), except for the CytoBead assay (0.17).

In order to verify that the LRs deducted from the study population can be widely applied, we calculated the LRs for each of the four laboratories (Belgium, Denmark, Germany and The Netherlands) that participated in this study. Although there were differences between the sites (most probably because of the smaller subgroups), there was an overall agreement between the test result interval-specific LRs across the four sites (supplementary Table S1, available at *Rheumatology* Online). This confirmed that the proposed LRs can be widely applied. Moreover, our findings confirm a previous smaller study [5]. After local validation, laboratories could apply the LRs proposed in this manuscript.

Interpretation of laboratory results must be done in the clinical context and as a function of the pre-test probability [6]. Jennette *et al.* [11] determined the pre-test probability for pauci-immune crescentic glomerulonephritis in patients with haematuria, proteinuria and elevated levels of creatinine, and Langford [10] described estimates of pre-test probability for AAV in patients with mucosal thickening of (one or more) sinuses, radiographic presence of pulmonary infiltrates or nodules and/or haematuria and red blood cell casts. Based on the pre-test probability and on test result interval-specific LRs, post-test probability can be estimated. We provided a visual representation of how the post-test probability depends on the pre-test probability and the antibody level. Such graphical representation has been shown to be the best way to convey diagnostic information [13]. When added to a test result, it can easily be understood and used by clinicians. Thus, we showed that in comparison to the classical way of reporting PR3- or MPO-ANCA test results as either positive or negative (using a single cut-off), reporting test result-specific LRs and/or post-test probability as a function of pre-test probability allows value to be added to a test result because it gives a more precise indication of the likelihood for disease.

We previously reported marked differences in performance between manual [4] and automated IIF methods [14]. Here, we report that a strategy based on immunoassay is equivalent or superior to a strategy based on screening with IIF and confirming positive results by immunoassay.

We also addressed the question of whether a positive PR3- and MPO-ANCA test result should be confirmed by IIF or by another immunoassay. Our results indicate that confirming positive PR3- and MPO-ANCA test results by IIF or another immunoassay does not significantly improve overall diagnostic performance. It does not increase sensitivity and, depending on the assay, it only increases specificity marginally. The best overall performance was found for an approach in which all samples were assayed by two different immunoassays (or by the combination of immunoassay and IIF). The added value of systematically performing a second test (either immunoassay or IIF) was assay dependent and consisted in a slight increase of the sensitivity and of the specificity (not for high antibody levels). This indicates that in patients with a high clinical suspicion of AAV but with negative or low PR3- or MPO-ANCA, additional testing with another immunoassay (or with IIF) is warranted.

In conclusion, we illustrated that defining thresholds of antibody levels and assigning test result interval-specific LRs may help with the interpretation of PR3- and MPO-ANCA. A graphical representation of the post-test probability as a function of the pre-test probability and the antibody level can assist clinicians to interpret PR3- and MPO-ANCA results. In patients with a high clinical suspicion and a negative test result or with a low antibody level, additional testing by another immunoassay or by IIF may be warranted.

Acknowledgements

We thank BioRad Laboratories Inc., Euro-Diagnostica AB, Euroimmun, Medipan/Generic assays GmbH, Inova Diagnostics, Orgentec and Thermo-Fisher Scientific for providing the reagents and performing the analyses.

Funding: BioRad Laboratories Inc., Euro-Diagnostica AB, Euroimmun, Medipan/Generic assays GmbH, Inova Diagnostics, Orgentec and Thermo-Fisher Scientific provided financial support for the logistics of this study.

Disclosure statement: B.H. has received a speaker fee from Orgentec. X.B. has received speaker fees from Thermo-Fisher, Menarini and Inova, has been a consultant for Inova and has received a research grant from Thermo-Fisher. E.C. has received lecture fees from BioRad Laboratories Inc., Euro-Diagnostica AB, Euroimmun, Medipan/Generic assays GmbH, Inova Diagnostics, Orgentec and Thermo-Fisher Scientific. All other authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

1 Cohen Tervaert JW, Damoiseaux J. Antineutrophil cytoplasmic autoantibodies: how are they detected and what is their use for diagnosis, classification and follow-up? Clin Rev Allergy Immunol 2012;43:211–9.

2 Radice A, Bianchi L, Sinico RA. Anti-neutrophil cytoplasmic autoantibodies: methodological aspects and clinical significance in systemic vasculitis. Autoimmun Rev 2013;12:487–95.

3 Csernok E, Moosig F. Current and emerging techniques for ANCA detection in vasculitis. Nat Rev Rheumatol 2014;10:494–501.

4 Damoiseaux J, Csernok E, Rasmussen N *et al*. Detection of anti-neutrophil cytoplasmic antibodies (ANCAs): a multicentre European Vasculitis Study Group (EUVAS) evaluation of the value of indirect immunofluorescence (IIF) versus antigenspecific immuno-assays. Ann Rheum Dis 2016;76:647–53.

5 Vermeersch P, Blockmans D, Bossuyt X. Use of likelihood ratios can improve the clinical usefulness of enzyme immunoassays for the diagnosis of small-vessel vasculitis. Clin Chem 2009;55:1886–8.

6 Bossuyt X. Clinical performance characteristics of a laboratory test. A practical approach in the autoimmune laboratory. Autoimmun Rev 2009;8:543–8.

7 Leavitt RY, Fauci AS, Bloch DA *et al*. The American College of Rheumatology 1990 Criteria for the classification of Wegener's granulomatosis. Arthritis Rheum 1990;22:1101–7.

8 Jennette JC, Falk RJ, Andrassy K *et al*. Nomenclature of systemic vasculitides. Proposal of an International Consensus Conference. Arthritis Rheum 1994;37:187–92.

9 Jennette JC, Falk RJ, Bacon PA *et al*. 2012 revised international Chapel Hill consensus conference nomenclature of vasculitides. Arthritis Rheum 2013;65:1–11.

10 Langford CA. The diagnostic utility of c-ANCA in Wegener's granulomatosis. Cleve Clin J Med 1998;65:135–40.

11 Jennette JC, Wilkman AS, Falk RJ. Diagnostic predictive value of ANCA serology. Kidney Int 1998;53:796–8.

12 Choi HK, Liu S, Merkel PA, Colditz GA, Niles JL. Diagnostic performance of antineutrophil cytoplasmic antibody tests for idiopathic vasculitides: metaanalysis with a focus on antimyeloperoxidase antibodies. J Rheumatol 2001;28:1584–90.

13 Vermeersch P, Bossuyt X. Comparative analysis of different approaches to report diagnostic accuracy. Arch Intern Med 2010;170:734–5.

14 Csernok E, Damoiseaux J, Rasmussen N *et al*. Evaluation of automated multi-parametric indirect immunofluorescence assays to detect anti-neutrophil cytoplasmic antibodies (ANCA) in granulomatosis with polyangiitis (GPA) and microscopic polyangiitis (MPA). Autoimmun Rev 2016;15:736–41.