

## EXTENDED REPORT

# Variations in performance characteristics of commercial enzyme immunoassay kits for detection of antineutrophil cytoplasmic antibodies: what is the optimal cut off?

J U Holle, B Hellmich, M Backes, W L Gross, E Csernok



*Ann Rheum Dis* 2005;64:1773–1779. doi: 10.1136/ard.2005.035279

See end of article for authors' affiliations

Correspondence to: Dr Elena Csernok, Rheumaklinik Bad Bramstedt, Oskar-Alexander-Str 26, 24576 Bad Bramstedt, Germany; csernok@rheuma-zentrum.de

Accepted 14 April 2005  
Published Online First  
20 April 2005

**Background:** Previous studies have shown considerable variation in diagnostic performance of enzyme linked immunosorbent assays (ELISAs) for measuring antineutrophil cytoplasmic antibodies (ANCA) specific for proteinase 3 (PR3) and myeloperoxidase (MPO).

**Objective:** To analyse the performance characteristics of different commercially available direct ANCA ELISA kits.

**Methods:** ELISA kits for detecting PR3-ANCA and MPO-ANCA from 11 manufacturers were evaluated. Serum samples were taken from patients with Wegener's granulomatosis (15), microscopic polyangiitis (15), other vasculitides (10), and controls (40). Results were compared with data obtained by indirect immunofluorescence (IFT). The diagnostic performance of the tests was analysed and compared by receiver operating characteristic (ROC) curve analysis.

**Results:** Applying the manufacturers' cut off resulted in great variation in sensitivity of the commercial PR3-ANCA kits for diagnosing Wegener's granulomatosis (ranging from 13.3% to 66.7%), and of the MPO-ANCA kits for diagnosing microscopic polyangiitis (ranging from 26.7% to 66.7%). Specificities were relatively constant (from 96.0% to 100%). IFT was superior to all ELISAs (C-ANCA for Wegener's granulomatosis: sensitivity 73.3%, specificity 98%; P-ANCA for microscopic polyangiitis: sensitivity 86.7%, specificity 98%). The sensitivities of PR3-ANCA and MPO-ANCA ELISA kits were increased by lowering the cut off values. This reduced specificity but increased overall diagnostic performance.

**Conclusions:** The low sensitivity of some commercial kits reflects the high cut off levels recommended rather than methodological problems with the assays. Comparative analyses using sera from well characterised patients may help identify optimum cut off levels of commercial ANCA ELISA tests, resulting in better comparability of results among assays from different manufacturers.

The detection of proteinase 3 and myeloperoxidase antineutrophil cytoplasmic antibodies (PR3-ANCA, MPO-ANCA) is a helpful tool for establishing the diagnosis of Wegener's granulomatosis and microscopic polyangiitis (MPA), respectively. Moreover, follow up measurement of both PR3-ANCA and MPO-ANCA may be important, as there is evidence that increases in titre may precede a relapse of the disease.<sup>1</sup> For routine detection of ANCA, a consensus statement on ANCA testing has been published by an international group of ANCA researchers. According to this statement, a further test (for example, a direct enzyme linked immunosorbent assay (ELISA)) is required if the indirect immunofluorescence technique (IFT, used as a screening method) successfully detects ANCA. This is because the value of IFT can be greatly enhanced by additional direct ELISA testing.<sup>2</sup>

There are various commercially available ELISA kits for detecting ANCA directed against proteinase 3 and myeloperoxidase. The manufacturing companies apply their own arbitrary units to express ANCA concentrations in a patient's serum, which therefore makes it impossible to compare the concentrations measured by different kits. Moreover, the various kits do not have equal performance.<sup>3–6</sup> Consequently, a clinician can only interpret the evolution of ANCA concentrations in a patient's serum by sticking to one assay.

Our aim in this study was to compare the performance of 11 commercially available direct ELISA kits, an in-house

direct ELISA, and IFT for detecting PR3-ANCA and MPO-ANCA in well characterised patients with ANCA associated vasculitides in order to identify factors associated with poor diagnostic performance.

## METHODS

### Patients

We studied 40 consecutive patients diagnosed with Wegener's granulomatosis (n = 15), microscopic polyangiitis (n = 15), Churg-Strauss syndrome (n = 3), giant cell arteritis (n = 1), Takayasu syndrome (n = 1), cryoglobulinaemic vasculitis (n = 1), unclassified small vessel vasculitis (n = 2), and connective tissue diseases with secondary vasculitis (n = 2) in the Department of Rheumatology in Lübeck (University Schleswig Holstein, Campus Lübeck) and Rheumaklinik Bad Bramstedt. Serum samples were obtained from well defined patients with a known clinical diagnosis of Wegener's granulomatosis or microscopic polyangiitis, where the diagnosis was made irrespective of serology (the presence of ANCA was not included as a criterion for the diagnosis in these patients). Classification of patients with primary systemic vasculitides was based on American College of

**Abbreviations:** ANCA, antineutrophil cytoplasmic antibodies; AUC, area under the curve; BVAS, Birmingham vasculitis activity score; ELISA, enzyme linked immunosorbent assay; IFT, indirect immunofluorescence technique; ROC, receiver operating characteristic

**Table 1** Clinical findings and organ involvement in 15 patients with Wegener's granulomatosis and 15 patients with microscopic polyangiitis

	WG (n = 15)			MPA (n = 15)		
	Total	Active disease	Inactive disease	Total	Active disease	Inactive disease
Women	5	4	1	11	7	4
Men	10	9	1	4	3	1
Age (years)						
Range	28 to 77			31 to 73		
Median	55			62		
<i>Organ involvement</i>						
ENT	7			0		
Lung	2			1		
Kidney	4			5		
Eye	4			0		
Joints (A)	2			6		
Skin	1			0		
PNS	3			1		
CNS	0			0		
Heart	0			1		

CNS, central nervous system; ENT, ear, nose and throat; joints (A), arthritis; MPA, microscopic polyangiitis; PNS, peripheral nervous system; WG, Wegener's granulomatosis.

Rheumatology (ACR) classification criteria<sup>7</sup> and Chapel Hill consensus conference definitions.<sup>8</sup> All the sera were taken at the time of diagnosis.

A prospective cohort of 30 disease controls was created, consisting of patients who were admitted consecutively to the medical clinic of the University of Schleswig Holstein, Campus Luebeck, over a period of one month with severe symptoms of pulmonary disease (unilateral or bilateral nodes or infiltration), renal disease (acute renal failure), or ENT disease (recurrent sinusitis) but without vasculitis. Sera from 10 healthy volunteers were also tested in all assays.

The diagnosis was biopsy proven in each patient. Biopsies had been examined by two different pathologists in the German reference centre for vasculitis (Department of Pathology, University of Schleswig-Holstein, Campus Lübeck). The disease activity was assessed by the Birmingham vasculitis activity score (BVAS)<sup>9</sup> at the time of serum collection.

The study was carried out according to the 1997 Declaration of Helsinki of the World Medical Association. The design of the work was approved by the ethics committee of the University of Schleswig-Holstein, Campus Lübeck, and each patient gave informed consent before participation in the study.

## Methods of ANCA detection

### Indirect immunofluorescence

ANCA detection by IFT was carried out on ethanol and formaldehyde fixed leucocytes, as described previously.<sup>10</sup> In our laboratory, a positive ANCA is defined as a titre of antibodies higher than 1:16.

### ELISA for PR3- and MPO-ANCA

In-house direct ELISAs were undertaken as described earlier.<sup>5</sup>

### Commercial ELISA kits to detect PR3-ANCA and MPO-ANCA

For testing, each company was designated by a letter (A to M) and their kits are identified in this way in the results section and the tables. Participating companies were: AESCU Diagnostics, Axis Shield, Binding Site, Biorad, Euro Diagnostica, Euroimmun, IBL, Inova, Pharmacia, Trinity Biotech, and Wieslab. Serum levels of PR3-ANCA and MPO-ANCA were assessed according to the manufacturers' instructions. The results were regarded as positive or negative according to the scale provided by each manufacturer.

To confirm the test results, all sera were analysed by capture ELISA and immunoblotting (data not shown).

**Table 2** Diagnostic performance of commercially available enzyme linked immunosorbent assay kits for detection of PR3-ANCA compared with direct immunofluorescence using cut off levels provided by the manufacturer

Test	Provided cut off	Sensitivity	Specificity
IFT (CA)	1:16	73.3	98
IHELI	<2.5 neg, >5 pos	66.7	100
A	>2.0	60.0	100
B	<10 neg, >15 pos	13.3	100
E	<7 neg, >10 pos	46.7	100
F	>15	26.7	98
G	>3.5	60.0	100
H	<0.9 neg, >1.1 pos	60.0	98
I	>20	60.0	100
J	>2	60.0	100
K	<20 neg, <30 pos	60.0	100
L	<10 neg, >20 pos	40.0	100
M	>15	26.7	96

CA, C-ANCA; IFT, indirect immunofluorescence technique; IHELI, in-house enzyme linked immunosorbent assay; neg, negative; pos, positive; PR3-ANCA, antineutrophil cytoplasmic antibodies to proteinase 3.

**Table 3** Diagnostic performance of commercially available enzyme linked immunosorbent assay kits for detection of MPO-ANCA compared with direct immunofluorescence using cut off levels provided by the manufacturer

Test	Provided cut off	Sensitivity		Specificity	
		Mean	95% CI	Mean	95% CI
IFT (PA)	1:16	86.7		98	
IHELI	>20	60.0		100	
A	>6	26.7		100	
B	<10 neg, >15 pos	66.7		98	
E	<7 neg, >10 pos	60.0		100	
F	>15	53.3		100	
G	>9	46.7		100	
H	<0.9 neg, >1.1 pos	66.7		100	
I	>20	46.7		100	
J	>6	53.3		100	
K	<20 neg, >30 pos	46.7		100	
L	<20 neg, >25 pos	60.0		100	
M	>15	53.3		100	

IFT, indirect immunofluorescence technique; IHELI, in-house enzyme linked immunosorbent assay; MPO-ANCA, antineutrophil cytoplasmic antibodies to myeloperoxidase; neg, negative; PA, P-ANCA; pos, positive.

Samples from the patients with Wegener's granulomatosis and microscopic polyangiitis were tested in both assays (PR3-ANCA and MPO-ANCA ELISAs) and were not used as controls for PR3 or MPO assays.

### Statistical analysis

The diagnostic sensitivity, specificity, and receiver operating characteristics (ROC) were analysed for the different ELISAs and the IFT to estimate the diagnostic performance of the respective tests. The optimal cut off levels for calculation of sensitivity and specificity were determined by ROC curves. Sensitivity and specificity were calculated by 2x2 tables. ROC curves were analysed as described previously.<sup>11</sup> In order to assess overall diagnostic performance, the area under the ROC curve (AUC) was calculated and AUC values of all different assays were compared. A difference in AUC between two assays was considered significant at a probability (p) value of <0.05. The MedCalc<sup>®</sup> software package (Medcalc<sup>®</sup> Software, Mariakerke, Belgium) was used for data analysis.

### RESULTS

Patient characteristics are outlined in table 1. Patients were categorised into two groups according to disease activity at the time of sampling. Thirteen of the Wegener's patients were classified as having active disease and 10 of the microscopic polyangiitis patients had active disease. Their median BVAS values were 17 and 14, respectively.

Fifty patients served as controls. Ten of these were diagnosed with vasculitides other than Wegener's granulomatosis and microscopic polyangiitis (three with Churg-Strauss syndrome, one with giant cell arteritis, one with Takayasu Syndrome, one with cryoglobulinaemic vasculitis, two with unclassified small vessel vasculitis and two with connective tissue diseases with secondary vasculitis). In this group, there was one false positive result among the PR3-ANCA tests (kit H) and no false positive results among the MPO-ANCA tests. Ten other subjects served as healthy controls. In all tests, these controls gave a negative result when tested for PR3-ANCA and MPO-ANCA, respectively.

Thirty patients suffering from non-vasculitic disorders were also tested (disease controls). Seven of these had renal insufficiency or glomerulonephritis, 10 had sinusitis, six had bronchial carcinoma and seven had various other disorders. In the disease control group, there were three false positive results (kits F and M) and one false positive IFT result in the PR3-ANCA tests; in the MPO-ANCA tests there were two false positive IFT results and one false positive result with kit B.

### Sensitivity and specificity

Sensitivity and specificity varied among the different kits and was dependent on the cut off that had been provided by the manufacturer. Concerning the tests for the detection of C-ANCA and PR3-ANCA, the sensitivity varied greatly if the

**Table 4** Diagnostic performance of commercially available enzyme linked immunosorbent assay kits for detection of PR3-ANCA compared with direct immunofluorescence using optimum cut off levels calculated by ROC curve analysis

Test	Optimum cut off	Sensitivity		Specificity		PPV	NPV	AUC	
		Mean	95% CI	Mean	95% CI			Mean	95% CI
IFT (CA)	>0	73.3	44.9 to 92.0	98	91.7 to 99.7	84.62	96.15	0.863	0.76 to 0.92
IHELI	>2.55	66.7	38.4 to 88.1	100	91.7 to 99.7	24.56	93.75	0.821	0.72 to 0.89
A	>0.53	80.0	51.9 to 95.4	90	79.1 to 95.5	70.59	100	0.900	0.81 to 0.96
B	>3.46	66.7	38.4 to 95.4	96	89.3 to 99.5	83.33	94.34	0.749	0.64 to 0.84
E	>1	66.7	38.4 to 88.1	96	85.0 to 98.3	83.33	94.34	0.791	0.69 to 0.87
F	>3.27	60.0	32.3 to 83.6	90	81.0 to 96.5	50.0	87.23	0.724	0.61 to 0.82
G	>1.29	66.7	38.4 to 88.1	96	91.7 to 99.7	100	89.29	0.820	0.72 to 0.89
H	>0.64	66.7	38.4 to 88.1	94	89.3 to 99.5	90.90	92.59	0.834	0.73 to 0.91
I	>2.74	86.7	59.5 to 98.0	94	82.9 to 97.4	72.72	100	0.900	0.81 to 0.96
J	>0.64	73.3	44.9 to 92.0	90	81.0 to 96.5	70.58	93.61	0.896	0.81 to 0.96
K	>11.29	66.7	38.4 to 88.1	100	91.7 to 99.7	100	90.90	0.875	0.78 to 0.93
L	>0	53.3	26.6 to 78.7	100	91.7 to 99.7	100	87.72	0.756	0.65 to 0.84
M	>5.77	53.3	26.6 to 78.7	92	85.0 to 98.3	72.72	92.59	0.729	0.62 to 0.82

AUC, area under the curve; CI, confidence interval; IFT, indirect immunofluorescence technique; IHELI, in-house enzyme linked immunosorbent assay; NPV, negative predictive value; PPV, positive predictive value; PR3-ANCA, antineutrophil cytoplasmic antibodies to proteinase 3; ROC, receiver operating characteristic.

**Table 5** Diagnostic performance of commercially available enzyme linked immunosorbent assay kits for detection of MPO-ANCA compared with direct immunofluorescence using optimum cut off levels calculated by ROC curve analysis

Test	Optimum cut off	Sensitivity		Specificity		PPV	NPV	AUC	
		Mean	95% CI	Mean	95% CI			Mean	95% CI
IFT (PA)	>0	86.7	59.9 to 98.0	98	91.7 to 99.7	86.67	96.15	0.931	0.852 to 0.976
IHELI	>0	60.0	32.2 to 83.6	100	94.4 to 100.0	100	89.29	0.800	0.696 to 0.881
A	>1.85	66.7	38.4 to 88.1	94	87.1 to 99.0	92.31	94.34	0.866	0.772 to 0.932
B	>12.06	80.0	51.9 to 95.4	92	85.0 to 98.3	83.34	94.34	0.885	0.794 to 0.945
E	>3	73.7	44.9 to 92.0	90	81.0 to 96.5	57.89	91.31	0.905	0.819 to 0.959
F	>2.03	86.7	59.5 to 98.0	94	87.1 to 99.0	100	94.34	0.933	0.854 to 0.976
G	>1.57	80.0	51.9 to 95.4	90	81.0 to 96.5	70.58	95.74	0.933	0.855 to 0.977
H	>0.77	80.0	51.9 to 95.4	98	94.4 to 100	86.67	100	0.907	0.821 to 0.960
I	>4.33	80.0	51.9 to 95.4	98	91.7 to 99.7	85.71	98.04	0.883	0.792 to 0.944
J	>2.56	66.7	38.4 to 88.1	96	89.3 to 99.5	77.78	100	0.846	0.748 to 0.917
K	>5.34	93.3	68.0 to 98.9	92	81.0 to 96.5	56.00	100	0.956	0.886 to 0.989
L	>0	73.3	44.9 to 92.0	100	94.4 to 100	92.31	96.15	0.867	0.772 to 0.932
M	>4.43	80.0	51.9 to 95.4	98	94.4 to 100	100	94.34	0.878	0.786 to 0.940

AUC, area under the curve; CI, confidence interval; IFT, indirect immunofluorescence technique; IHELI, in-house enzyme linked immunosorbent assay; MPO-ANCA, antineutrophil cytoplasmic antibodies to myeloperoxidase; NPV, negative predictive value; PA, P-ANCA; PPV, positive predictive value; ROC, receiver operating characteristic.

given cut off values were applied, ranging from 13.3% (kit B) to 73.3% (IFT) (table 2). The in-house ELISA performed second best for sensitivity (66.7%), followed by kits A, G, H, I, J, and K (60%) and kit E (46.7%). Specificity did not have such a large range when the given cut off levels were used, ranging from 96% (kit M) to 98% (kits F and H and the IFT). Most of the assays (in-house ELISA, A, B, E, G, I, J, K, and L) reached a specificity of 100%.

The test for the detection of P-ANCA and MPO-ANCA gave similar results for sensitivity and specificity. There was a wide range of sensitivity, from 26.7% (kit A) to 86.7% (IFT) (table 3). Most of the kits showed a sensitivity in the range of 46% to 67% (kits G, I, and K, 46.7%; kits F, J, and M, 53.3%; kits E and L, and in-house ELISA, 60%; kits B and H, 66.7%). Again specificity was more uniform. All assays except IFT and kit B reached 100% specificity (IFT and kit B, 98%).

**Change of cut off points and its influence on specificity**

Sensitivity and specificity could be altered to improve the diagnostic power of the test by changing the cut off levels. Using the optimum cut off for the individual tests, which can

be calculated by ROC analysis, sensitivity of the PR3-ANCA tests increased to a range from 53.3% (kits L and M) to 86.7% (kit I). Kit A gave the second highest sensitivity (80%). Most of the kits reached a sensitivity of 66.7% (kits B, E, F, G, H, and K). The gain in sensitivity was achieved at the cost of a minor loss in specificity, ranging now from 90% (kits F and J) to 100% (kits K and L)). Kits A and I gave the best diagnostic performance (AUC 0.900).

In all kits testing for the presence of PR3-ANCA, except for IFT and the in-house ELISA, a change in cut off values was able to increase their overall performance, as shown in tables 4 and 5.

By changing cut off levels, sensitivity of MPO-ANCA assays increased from 66.7% (kits A and J) to 93.3% (kits G and K), and specificities were therefore reduced slightly (from 90% (kits E and G) to 100% (kit L)). Kit K has the highest diagnostic power (AUC 0.956). In all of the MPO-ANCA assays, performance (measured by AUC) could be increased by changing the cut off levels.

In a few cases optimum performance (measured by AUC) was accompanied by a loss in specificity to below 90% (kits F

**Table 6** Comparison of false positive and true positive results applying the provided and optimum cut off points in enzyme linked immunosorbent assay kits for detection of PR3-ANCA

Test	Controls (n = 50)			WG (n = 15)		
	False positive results		Additional false positive results	True positive results		Additional positive results
	Provided cut off	Optimum cut off		Provided cut off	Optimum cut off	
IFT	1	1	0 (0%)	11	11	0 (0%)
IHELI	0	0	0 (0%)	10	10	0 (0%)
A	0	5	5 (10%)	9	12	3 (20.0%)
B	0	2	2 (4%)	2	10	8 (53.4%)
E	0	2	2 (4%)	7	10	3 (20.0%)
F	1	9	8 (16%)	4	9	5 (33.4%)
G	0	2	2 (4%)	9	10	1 (6.7%)
H	1	1	0 (0%)	9	10	1 (6.7%)
I	0	3	3 (6%)	9	13	4 (20.7%)
J	0	6	6 (12%)	9	11	2 (13.3%)
K	0	0	0 (0%)	9	10	1 (6.7%)
L	0	0	0 (0%)	6	8	2 (13.3%)
M	2	4	2 (4%)	4	8	4 (26.7%)

IFT, indirect immunofluorescence technique; IHELI, in-house enzyme linked immunosorbent assay; PR3-ANCA, antineutrophil cytoplasmic antibodies to proteinase 3; WG, Wegener's granulomatosis.

**Table 7** Comparison of false positive and true positive results applying the provided and optimum cut off points in enzyme linked immunosorbent assay kits for detection of MPO-ANCA

Test	Controls (n = 50)			MPA (n = 15)		
	False positive results		Additional false positive results Optimum cut off	True positive results		Additional positive results Optimum cut off
	Provided cut off	Optimum cut off		Provided cut off	Optimum cut off	
IFT	2	2	0 (0%)	13	13	0 (0%)
IHELI	0	0	0 (0%)	9	9	0 (0%)
A	0	3	3 (6%)	4	10	6 (14.0%)
B	1	4	3 (6%)	10	12	2 (13.3%)
E	0	8	8 (16%)	9	11	2 (13.3%)
F	0	3	3 (6%)	8	13	5 (33.3%)
G	0	5	5 (10%)	7	12	5 (33.3%)
H	0	1	1 (2%)	10	12	2 (13.3%)
I	0	1	1 (2%)	7	12	5 (33.3%)
J	0	2	2 (4%)	8	10	2 (13.3%)
K	0	4	4 (8%)	7	14	7 (46.7%)
L	0	0	0 (0%)	9	11	2 (13.3%)
M	0	1	1 (2%)	8	12	4 (26.6%)

IFT, indirect immunofluorescence technique; IHELI, in-house enzyme linked immunosorbent assay; MPA, microscopic polyangiitis; MPO-ANCA, antineutrophil cytoplasmic antibodies to myeloperoxidase.

and J with PR3-ANCA tests, and kits E and G with MPO-ANCA tests). As specificity should not be lower than 90%, AUC values were calculated on the basis of a minimum specificity of 90%.

As stated above, lowering cut off points to give a better overall diagnostic performance gives a higher sensitivity but means a reduction in specificity. The effect of lowering cut off points is shown in tables 6 and 7. Applying the cut off points provided by the manufacturer, there were three tests among the PR3-ANCA detecting kits that gave false positive test results in the control groups (n = 50), ranging from one to two false positive results per test; healthy controls were all tested as being negative for PR3-ANCA (by all tests). Using optimum cut off levels, nine of the tests gave false positive results in the disease control group, ranging from one to nine false positive results per test. False positive disease controls were distributed relatively equally among the different disease groups (renal, pulmonary, or ENT disease), showing one or two false positive results in one or more of the disease groups. In one case there were four false positive results in the ENT disease control group. Moreover, this kit showed two false positive results in both the renal and the pulmonary disease control groups (PR3-ANCA kit F).

Moreover, among the healthy controls, there were now five tests giving false positive results (ranging from one to two per test). Kits K and L did not give false positive results in the control groups irrespective of the cut off level and IFT had one false positive result.

However, the reduction in specificity was outweighed by the increase in sensitivity, as shown in table 6. This shows the increase in false positive results among the controls (absolute value and percentage) and the increase of correct results among the disease group (Wegener's granulomatosis). In most cases, the increase in correct results exceeded the increase in false negative results (for example, kit A of the PR3-ANCA ELISAs, with 10% additional false positive results and 20% additional true positive results), which indicates an increased performance. The improvement in the tests by changing cut off values was variable: in a few cases, changing cut offs did not change false or true positive tests results at all (in-house ELISA, IFT), which means that these assays could not be improved by changing cut off points, whereas in other

cases there was a substantial increase in positive results and only a moderate increase in false positive results (kit B additional false positive result 4%, additional true positive results 53.4%)

In some cases, companies provided a "borderline positive" scale and these results were counted as negative. After lowering cut off points (by applying the optimum cut off), these borderline results then gave a definite positive test result. In the disease group (Wegener's granulomatosis), kits E, H, and L all gave positive "borderline" results (kits E = 1, H = 1, L = 2), which became positive after lowering the cut off values. These results did not appear as additional false positive or true positive results among the control or disease groups. Similar results were obtained for the MPO-ANCA ELISA tests (table 7).

To summarise, lowering cut off points is in most cases effective in improving diagnostic performance. However, one should be aware of the reduction of specificity (increased false positive results).

### ROC analysis

In order to determine the diagnostic performance of the different ELISAs and the IFT, ROC curves were analysed for the respective diagnostic tests. According to a method described by Hanley and McNeil,<sup>12</sup> the 95% confidence interval for the area under the ROC curve was used to test the hypothesis that the theoretical area is 0.5. So long as the confidence interval does not include the 0.5 values, there is evidence that the test under investigation has the ability to distinguish between disease and controls.

The value of the AUC depends on two factors: first, sensitivity/specificity should be high; second, differences in specificity and sensitivity should be slight. The higher the AUC value, the greater the diagnostic power.

We calculated AUC values for all the tests. In general, the results showed that each of the tests had a high diagnostic power for detecting their respective target (PR3-ANCA or MPO-ANCA). Concerning kits testing for C-ANCA/PR3-ANCA, kits A and I had the highest diagnostic power (AUC 0.900), followed by kit J (AUC 0.896), kit K (AUC 0.875), and IFT (AUC 0.863); however, these differences in AUC among the five best performing kits were not statistically significant



( $p > 0.05$ ). IFT did not show any significant differences in overall diagnostic performance (measured by AUC) compared with any of the ELISA kits. However, the four best performing kits were significantly different from the four worst performing kits (kits F, M, B, and L;  $p < 0.05$ ).

With respect to assays detecting P-ANCA/MPO-ANCA, kit K gave the highest diagnostic power (AUC 0.956), followed by kits F and G (AUC 0.933) and IFT (AUC 0.931). There was no statistically significant difference between the performances of the tests except for kit K (best performance, AUC 0.956) and in-house ELISA (AUC 0.800;  $p = 0.025$ ). It was remarkable that there was no statistically significant difference between IFT and any of the ELISA kits, as shown for the PR3-ANCA ELISAs.

### Inter-test agreement

The inter-test agreement among the different PR3-ANCA kits was between 55% and 100% and the concordance among the MPO-ANCA tests was between 18% and 100%.

## DISCUSSION

Along with the clinical presentation and the histological picture, the detection of ANCA is one of the cornerstones of the diagnosis of ANCA associated vasculitis. Current guidelines on ANCA testing recommend dual testing—first by IFT and then by ELISA or other tests.<sup>2</sup> Many previous studies have shown that the sensitivity of most direct PR3-ANCA and MPO-ANCA ELISAs is equivalent to standard IFT for ANCA detection using neutrophil cytospin preparations. However, in these studies the specificity of the assays was lower than that of IFT.<sup>12–14</sup>

The current problem in ANCA testing is the lack of international standardisation of ANCA assays to allow comparison of test results from different laboratories or hospitals. Commercially available kits, as well as in-house tests from different hospitals, have different performance characteristics. This may affect the diagnosis and management of patients with ANCA associated vasculitides.

We have previously undertaken three different studies to evaluate the performance characteristics of commercially available direct and capture ELISA kits for detecting PR3-ANCA and MPO-ANCA in patients with ANCA associated vasculitides.<sup>4–6</sup> Direct PR3-ANCA and MPO-ANCA ELISAs have not shown an improvement in performance over the years, although the numbers of test kits have increased. In 1997, we tested eight direct PR3-ANCA and MPO-ANCA ELISA kits, finding sensitivities between 44% and 84% and specificities of 90% to 100%.<sup>3</sup> In 2002, we tested 11 direct PR3 and MPO ELISA kits. Sensitivities ranged from 22% to 77% and 45% to 67%, respectively.<sup>4</sup> Specificity was 93–100% and 97–100%. The most recent evaluation of six in-house direct PR3-ANCA ELISAs from academic laboratories in 2003 showed sensitivities between 53% and 80% and specificities from 95% to 100%.<sup>6</sup>

In the present study, we analysed the performance of commercially available direct ELISA kits to detect autoantibodies to PR3 and MPO in patients with ANCA associated vasculitides. In a relatively large number of clinically well defined patients ( $n = 70$ ) we analysed their serum samples to compare the results of in-house IFT and direct ELISA. In general, if the cut off levels provided by the manufacturer were applied, the tests for direct PR3-ANCA and MPO-ANCA ELISA both showed a wide range of sensitivities, whereas specificity was high among all the tests.

Comparing the test results with previous studies, the performance of direct PR3-ANCA ELISAs was not improved by applying the cut off points provided (sensitivity ranging from 13.3% to 66.7%, specificity from 96% to 100%). Moreover, both sensitivity and specificity decreased.

Applying lower cut off points, the sensitivity was similar to the 1997 results (53.3% to 86.7%), but the specificity was lower (82% to 100%). With respect to direct MPO-ANCA ELISAs, sensitivity could not be improved (26.7% to 66.7%); however, specificity has increased over recent years (now, 98% to 100%; in 1997, 90% to 100%). By applying lower cut off points, sensitivity could be improved (60% to 93.3%), but specificity was reduced substantially (90–100%). Specificity was kept at a minimum of 90%. Theoretically, in some cases it would have been necessary to lower specificity below 90% to achieve an optimum AUC. The overall diagnostic performance of the tests could be improved in all tests except for two (PR3-ANCA in-house ELISA and IFT) by lowering the cut off levels. This is generally achieved by an improvement in sensitivity and diagnostic performance, but may sometimes be accompanied by a loss of specificity.

The AUC value served to assess the overall diagnostic performance of a test independent of the manufacturer-provided cut off levels. As stated above, among the PR3-ANCA kits, kits A and I had the best diagnostic performance (AUC 0.900) followed by kit J (AUC 0.896), kit K (AUC 0.875), and IFT (AUC 0.863); however, these differences were not statistically significant. A significant difference in performance was shown by comparing the four best kits with the four weakest kits.

With respect to MPO-ELISA, after lowering cut off levels, there was a relatively uniform performance, with a significant difference occurring only between the best and worst performing kit. Kits F, G, and K give a marginally better performance than IFT; this, however, was not statistically significant. None of the ELISA kits showed a significant difference in the AUC values, which means that that they had equivalent performance after lowering the cut off points.

Previous studies have shown that IFT is superior to direct ELISA. In this study, we found that by changing cut off levels to optimise overall diagnostic performance, all ELISAs (for PR3-ANCA and MPO-ANCA) had a performance equal to IFT, and a few even had a minimally better overall diagnostic performance (measured by AUC) than IFT (kits A, I, J, and K among the PR3-ANCA ELISA kits, and kits F, G, and K among the MPO ANCA kits); however, these differences were not different. Comparing the different ELISA kits, there was a significant difference among the four best and the four worst performing PR3-ANCA ELISA kits, while among MPO-ANCA ELISAs there was a more uniform performance, with a significant difference only between kit K and the in-house ELISA.

It is however, necessary to mention that by lowering cut off points specificity is in some cases substantially reduced. Although by optimising cut off points the percentage of additional correctly positive results usually exceeded the percentage of additional negative results, it is important to be aware that the increased in false positive results is greater when optimum cut off points are applied. For example, regarding the direct PR3-ANCA ELISAs, the best performing kits A and I had no false positive results when the manufacturer-provided cut off points were applied, while the optimum cut off resulted in five and three false positive results, respectively. Consequently, companies need to be critical when planning to lower cut off points to increase overall diagnostic performance. If a test has a sufficiently high sensitivity, a negative result would rule out the diagnosis of ANCA associated vasculitides. If the test has high specificity, a positive result would support the diagnosis. Thus the cut off level can be set to make the test highly sensitive or highly specific, depending on the target disease. It must be emphasised that the rank orderings of sensitivity and specificity do not imply a definitive overall qualitative ranking of the test kits relative to one another.

## Conclusions

Our data show that the diagnostic performance of commercially available direct PR3-ANCA and MPO-ANCA ELISA kits has not improved over the years. Thus the clinician needs to be aware of the different performance characteristics of the various ELISA kits. We propose that by optimising cut off levels using ROC curve analysis, the overall diagnostic performance of many ELISA tests can be improved. Manufacturers of ANCA ELISA kits need to consider the correct cut off level to optimise a test. More studies will be needed to optimise cut off levels and to compare the performance of ELISAs using optimum cut offs.

## ACKNOWLEDGEMENTS

This study was supported by BMBF, grant No 01 G1 9951, Competence Network systemic inflammatory rheumatic diseases.

## Authors' affiliations

J U Holle, B Hellmich, M Backes, W L Gross, E Csernok, Department of Rheumatology, University of Schleswig-Holstein Campus Lübeck and Rheumaklinik Bad Bramstedt, Germany

## REFERENCES

- Boomsma MM, Stegeman CA, van der Leij MJ, Oost W, Hermans J, Kallenberg CG, *et al.* Prediction of relapses in Wegener's granulomatosis by measurement of antineutrophil cytoplasmic antibody levels. *Arthritis Rheum* 2000;**43**:2025–33.
- Savage J, Gillis D, Benson E, Davies D, Esnault V, Falk RJ, *et al.* International consensus statement on testing and reporting of antineutrophil cytoplasmic antibodies (ANCA). *Am J Clin Pathol* 1999;**111**:507–13.
- Lim LC, Taylor JG, Schmitz JL, Folds JD, Wilkman AS, Falk RJ, *et al.* Diagnostic usefulness of antineutrophil cytoplasmic antibody serology. Comparative evaluation of commercial indirect fluorescent antibody kits and enzyme immunoassay kits. *Am J Clin Pathol* 1999;**111**:363–9.
- Csernok E, Ahlquist D, Ullrich S, Gross WL. A critical evaluation of commercial immunoassays for antineutrophil cytoplasmic antibodies directed against proteinase 3 and myeloperoxidase in Wegener's granulomatosis and microscopic polyangiitis. *Rheumatology* 2002;**41**:1313–17.
- Wang G, Csernok E, da Groot K, Gross WL. Comparison of eight commercial kits for quantitation of antineutrophil cytoplasmic antibodies (ANCA). *J Immunol Methods* 1997;**208**:203–11.
- Csernok E, Holle J, Hellmich B, Willem J, Tervaert C, Kallenberg CG, *et al.* Evaluation of Capture ELISA for the detection of antineutrophil cytoplasmic antibodies directed against proteinase 3 in Wegener's granulomatosis: first results from a multicenter study. *Rheumatology* 2004;**43**:174–80.
- Leavitt RY, Fauci AS, Bloch DA, Michel BA, Hunder GG, Arend WP, *et al.* The American College of Rheumatology 1990 criteria for the classification of Wegener's granulomatosis. *Arthritis Rheum* 1990;**33**:1101–7.
- Jennette JC, Falk RJ, Andrassy K, Bacon PA, Churg J, Gross WL, *et al.* Nomenclature of systemic vasculitides: The proposal of an International Consensus Conference. *Arthritis Rheum* 1994;**37**:187–9.
- Luqmani RA, Bacon PA, Moots RJ, Janssen BA, Pall A, Emery P, *et al.* Birmingham Vasculitis Activity Score (BVAS) in systemic necrotizing vasculitis. *Q J Med* 1994;**87**:671–3.
- Schonermark U, Lamprecht P, Csernok E, Gross WL. Prevalence and spectrum of rheumatic diseases associated with proteinase 3-antineutrophil cytoplasmic antibodies (ANCA) and myeloperoxidase-ANCA. *Rheumatology (Oxford)* 2001;**40**:178–84.
- Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;**143**:29–36.
- Hagen EC, Daha MR, Hermans J, Andrassy K, Csernok E, Gaskin G, *et al.* Diagnostic value of standardized assays for anti-neutrophil cytoplasmic antibodies in idiopathic systemic vasculitis. EC/BCR Project for ANCA Assay Standardization. *Kidney Int* 1998;**53**:743–53.
- Merkel PA, Polisson RP, Chang Y, Skates SJ, Niles JL. Prevalence of antineutrophil cytoplasmic antibodies in a large inception cohort of patients with connective tissue disease. *Ann Intern Med* 1997;**126**:866–73.
- Specks U, Wiegert EM, Homburger HA. Human mast cells expressing recombinant proteinase 3 (PR3) as substrate testing for anti-neutrophil cytoplasmic antibodies (ANCA). *Clin Exp Immunol* 1997;**109**:286–95.