

Postprint Version	1.0
Journal website	<a href="http://ard.bmj.com/content/70/8/1468.long">http://ard.bmj.com/content/70/8/1468.long</a>
Pubmed link	<a href="http://www.ncbi.nlm.nih.gov/pubmed/21586440">http://www.ncbi.nlm.nih.gov/pubmed/21586440</a>
DOI	10.1136/ard.2010.148619

This is a NIVEL certified Post Print, more info at <http://www.nivel.eu>

# Validation of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: slight improvement over the 1987 ACR criteria

KARIN BRITSEMMER<sup>1</sup>, JENNIE URSUM<sup>1</sup>, MARTIJN GERRITSEN<sup>1</sup>, LILIAN VAN TUYL<sup>2</sup>, DIRKJAN VAN SCHAAARDENBURG<sup>1</sup>

<sup>1</sup>Department of Rheumatology, Jan van Breemen Research Institute I Reade, Amsterdam, The Netherlands

<sup>2</sup>Department of Rheumatology, VU University Medical Center, Amsterdam, The Netherlands

## ABSTRACT

**Background** Recently, an American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) collaboration developed new classification criteria for rheumatoid arthritis (RA).

**Objective** To evaluate the diagnostic and discriminative ability of these new criteria compared with the 1987 ACR criteria and the Visser decision rule.

**Methods** 455 patients with early arthritis were studied. The diagnostic performance of the criteria was evaluated using methotrexate treatment within 1 year, expert opinion RA and erosive disease as 'gold standards'. Erosive disease was defined as a 0–3 year change in radiographic score of  $\geq 5$ .

**Results** The discriminative ability of the three criteria sets (2010 ACR/EULAR, 1987 ACR criteria and Visser algorithm) was similar with areas under the curve of 0.71–0.78 ('gold standard' methotrexate), 0.74–0.80 (gold standard expert opinion RA) and 0.63–0.67 (gold standard erosive disease after 3 years). The sensitivity of the 2010 ACR/EULAR criteria was highest with 0.85 (gold standard methotrexate). 86% of patients with RA and 51% of 'non-RA' patients according to the new criteria used methotrexate.

**Conclusion** The 2010 ACR/EULAR criteria were slightly more sensitive, but otherwise performed similarly to the older criteria. A high percentage of 'non-RA' patients used methotrexate, the gold standard for RA. The ability of the new criteria to identify patients with erosive disease was low, possibly owing to the effect of intensive treatment.

## INTRODUCTION

Early diagnosis and treatment of rheumatoid arthritis (RA) has become a major goal in rheumatology. Recently, an American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) collaboration developed new classification criteria for RA in order to facilitate the identification of patients who would benefit from early intervention.<sup>1 2</sup> The former 1987 ACR criteria for RA were considered not sensitive enough in early disease.<sup>3</sup> Consequently, other prediction rules or algorithms have been developed to predict which patients are at high risk of persistent/erosive disease or of progression from undifferentiated arthritis to RA.<sup>4 5</sup> The aim of this study was (1) to validate the diagnostic performance of the 2010 ACR/EULAR classification criteria for RA, compared with the 1987 ACR criteria for RA and the

prediction algorithm of Visser,<sup>5</sup> in an independent early arthritis cohort (EAC) and (2) to examine the capacity of the new criteria to identify future erosive disease.

## METHODS

### Study cohort

The study population comprised a cohort of patients with early arthritis from the EAC at the Jan van Breemen Institute in Amsterdam, The Netherlands. This ongoing cohort includes patients aged  $\geq 18$  years, with at least two swollen joints, symptom duration of  $< 2$  years and no prior disease-modifying antirheumatic drug (DMARD) treatment. Patients with osteoarthritis, crystal arthropathy, spondyloarthritis, systemic lupus erythematosus, Sjögren's syndrome and infectious arthritis are excluded. The study population consists mainly of patients with oligoarthritis or polyarthritis who would formerly be classified as having RA or undifferentiated arthritis. Data were used of patients who were included from 2000 onwards (to conform with the criteria development cohorts), had been followed up for at least 12 months and from whom baseline radiographs without typical RA erosions were available. Because of the absence of agreement on the definition of typical erosions we regarded erosion scores of  $> 1$ , assessed with the Sharp–van der Heijde scoring (SHS) method, as typical RA erosions.<sup>6</sup> Next, patients were scored according to the new classification criteria.<sup>2</sup>

The analysis of erosive disease is based on patients who had available radiographs at 0 and 3 years' follow-up. Radiographs of the hands, wrists and feet were scored with SHS by either one of two independent experienced observers who were aware of the chronological sequence of the radiographs. Interclass correlation between these two observers was 0.90 (95% CI 0.86 to 0.92).

### Validation comparisons

We compared the 2010 ACR/EULAR criteria with the 1987 ACR criteria and the Visser algorithm. We were not able to include the van der Helm algorithm in the analysis because this algorithm was designed to be applied in patients who do not fulfil the 1987 ACR criteria at baseline, and of the remaining patients only 10 fulfilled the van der Helm criteria.<sup>4</sup> Methotrexate treatment within 1 year after inclusion was used as 'gold standard' in accordance with the methods of the ACR/EULAR task force.<sup>1</sup> In addition the performance of the 2010 ACR/EULAR criteria was calculated with expert opinion RA (defined as the diagnosis RA in the hospital diagnosis registry at 1 year after EAC inclusion) and erosive disease after 3-years' follow-up as alternative gold standards. Patients were considered to have erosive disease if the 3-year interval change in total SHS was  $\geq 5$ .<sup>7</sup>

### Statistical analysis

Two by two classification tables were generated to estimate sensitivity, specificity, positive and negative predictive value (PPV and NPV, respectively), and positive and negative likelihood ratio (LR+ and LR–, respectively) at the proposed cut-off values for the respective algorithms (2010 ACR/EULAR and the Visser algorithm  $\geq 6$ , and 1987 ACR  $\geq 4$ ). To evaluate the capacity of the new criteria to discriminate between patients with and without methotrexate treatment within 1 year and the ability of the new criteria to identify erosive disease, receiver operating characteristic (ROC) curves with corresponding areas under the curve (AUC) were calculated.

## RESULTS

### Patients

The patients had a mean age of 52 years, a median disease duration of 5.6 months and active disease as judged by the median Disease Activity Score (28-joint count) of 4.8 (table 1). Thirty-one patients (6.8%) had an erosion score of 1 at baseline.

### [TABLE 1]

Baseline characteristics of 455 patients with early arthritis, and radiographic progression data of 175 patients

### Performance of the criteria

Three hundred and fifty-four (78%) patients had RA according to the gold standard (methotrexate use). Methotrexate treatment was given to 302 (86%) patients with RA and to 52 (51%) patients without RA according to the 2010 ACR/EULAR criteria. At baseline 78% fulfilled the 2010 ACR/EULAR criteria while 68% fulfilled the 1987 ACR criteria. Of 455 patients, 277 (61%) fulfilled both criteria sets, 76 (17%) patients fulfilled only the 2010 ACR/EULAR criteria and 22 (5%) fulfilled only the 1987 ACR criteria.

The AUC of the ROC curves show good discrimination for all three criteria sets between patients who did and did not receive methotrexate (table 2), with the highest value for the 2010 ACR/EULAR criteria. The sensitivity, specificity, PPV, NPV, LR+ and LR- of the three criteria sets at the proposed cut-off points for the three different gold standards are shown in table 2. The sensitivity of the 2010 ACR/EULAR criteria was higher than of the 1987 ACR criteria: 0.85 versus 0.76, respectively. The Visser algorithm cut-off points resulted in a lower sensitivity but higher specificity in our cohort. When we performed sensitivity analyses, and varied the cut-off point for typical RA erosions from >1 to >4, the results were similar (data not shown).

### [TABLE 2]

Performance of three sets of classification criteria for rheumatoid arthritis (RA) or for prediction of erosive disease

Of 455 patients, 310 (68%) had RA at 1 year according to expert opinion. Again the 2010 ACR/EULAR criteria had the highest score for sensitivity. The most common diagnoses after RA were oligoarthritis (n=44), psoriatic arthritis (n=31) and polyarthritis (n=27).

### PREDICTION OF EROSIIVE DISEASE

Radiographs at 3 years' follow-up were available from 175 patients. In these patients median total SHS was 0 (IQR 0–0) at baseline and 0.0 (IQR 0–4) at 3 years (table 1). Twenty of 33 'non-RA' patients according to the 2010 ACR/EULAR criteria (61%) in this subgroup used methotrexate within the first year.

The AUC of the ROC curves show a similar performance for the three criteria sets in the discrimination of patients with and without erosive disease (table 2). For all three criteria sets sensitivity was high (0.85–0.91) but specificity was low (0.17–0.41). When we performed sensitivity analyses, and varied the cut-off point for erosive disease from  $\geq 3$  to  $\geq 8$  increase in SHS points, the results remained similar (data not shown).

### DISCUSSION

In an independent EAC, we tested the performance of the 2010 ACR/EULAR and 1987 ACR classification criteria for RA and of the Visser decision rule for prediction of persistent/erosive disease. Apart from a higher frequency of anti-citrullinated protein antibody (52% vs 30%) in the presently studied EAC, there were no major differences with the EACs used for the derivation of the 2010 ACR/EULAR criteria.<sup>1</sup> The three criteria sets had a similar ability to discriminate between RA or persistent arthritis and self-limiting arthritis. Sensitivity was highest with the 2010 ACR/EULAR criteria, and specificity was highest with the Visser rule. Recently, a study by van der Linden *et al*<sup>8</sup> reported similar diagnostic and discriminative abilities of the 2010 ACR/EULAR criteria compared with the results of this study; however, their validation cohort was also used to develop the criteria. In this cohort the recommended cut-off points result in a high proportion of 'false-positive' patients—that is, patients classified as RA who did not receive methotrexate. If the new criteria were to be used as diagnostic instead of classification criteria, this could lead to unnecessary treatment with potentially toxic DMARDs.

One of the goals of the ACR/EULAR task force was to develop a criteria set that identifies patients at high risk of erosive disease.<sup>2</sup> The discriminative ability of the 2010 ACR/EULAR criteria in this respect was similar to the Visser algorithm but slightly lower than the 1987 ACR criteria. The trend towards a lower radiographic progression that was observed in the group of 'non-RA' patients may be more the consequence of a high rate of methotrexate use in these patients than of a good prognosis itself in this group. The high rate of methotrexate treatment in the group of 'non-RA' patients suggests that the rheumatologists in our clinic had a more aggressive approach to early arthritis during the same period than

the rheumatologists treating the cohorts that were used to derive the criteria. Obviously, this gold standard is not a static feature as rheumatologists have a tendency to treat earlier and more aggressively.

In conclusion the results suggest that, when applied to patients of this specific EAC with a high prevalence of RA, the new ACR/EULAR RA classification criteria are slightly more sensitive, and otherwise similar to, the older criteria. Since the 1987 ACR criteria set contains criteria that are highly uncommon in early disease, such as rheumatoid nodules, use of the new criteria is preferred.

#### ACKNOWLEDGMENTS

The authors thank E de Wit-Taen, V van de Lugt en J Sterk for collecting patient data.

#### FOOTNOTES

**Funding** Partly funded by the Dutch Arthritis Association.

**Competing interests** None.

**Ethics approval** This study was conducted with the approval of the local medical ethical committee.

**Provenance and peer review** Not commissioned; externally peer reviewed.

#### REFERENCES

1. Funovits J, Aletaha D, Ytterberg K, et al. The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: methodological report phase I. *Ann Rheum Dis* 2010;69:1589–95.
2. Aletaha D, Neogi T, Silman AJ, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69:1580–8.
3. Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
4. van der Helm-van Mil AH, le Cessie S, van Dongen H, et al. A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. *Arthritis Rheum* 2007;56:433–40.
5. Visser H, le Cessie S, Vos K, et al. How to diagnose rheumatoid arthritis early: a prediction model for persistent (erosive) arthritis. *Arthritis Rheum* 2002;46:357–65.
6. van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. *J Rheumatol* 2000;27:261–3.
7. Bruynesteyn K, van der Heijde D, Boers M, et al. Determination of the minimal clinically important difference in rheumatoid arthritis joint damage of the Sharp/van der Heijde and Larsen/Scott scoring methods by clinical experts and comparison with the smallest detectable difference. *Arthritis Rheum* 2002;46:913–20.
8. van der Linden MP, Knevel R, Huizinga TW, et al. Classification of rheumatoid arthritis: comparison of the 1987 American College of Rheumatology criteria and the 2010 American College of Rheumatology/European League Against Rheumatism criteria. *Arthritis Rheum* 2011;63:37–42.

TABLES

**Table 1** Baseline characteristics of 455 patients with early arthritis, and radiographic progression data of 175 patients

Characteristics	All patients (n=455)	2010 ACR/EULAR+ (n=353)	1987 ACR+ (n=310)	Visser+ (n=254)
<b>Demographic variables</b>				
Age (years, mean (SD))	52 (13)	52 (13)	53 (12)	51 (12)
Women (% of patients)	69	71	70	75
Disease duration (months, mean (SD))	5.6 (5.5)	5.5 (5.0)	5.6 (5.0)	5.8 (4.9)
SJC	5 (3–10)	6 (3–11)	7 (4–11)	6 (3–10)
TJC	5 (2–9)	6 (3–10)	6 (3–10)	5 (3–8)
DAS28 score (mean (SD))	4.8 (1.3)	5.0 (1.2)	5.2 (1.2)	4.9 (1.1)
HAQ score (range 0–3)	1.13 (0.63–1.63)	1.13 (0.75–1.75)	1.25 (0.75–1.75)	1.13 (0.75–1.63)
RA according to ACR/EULAR (%)	353 (77.6)	100.0	89.4	97.6
RA according to 1987 ACR (%)	310 (68.1)	78.5	100.0	84.3
Persistent arthritis according to Visser (%)	254 (55.8)	70.3	69.7	100.0
<b>Laboratory variables</b>				
ESR (mm/first hour)	25 (11–38)	26 (13–41)	27 (15–41)	26 (13–41)
CRP (mg/l)	11 (4–31)	13 (4–35)	15 (5–35)	13 (4–30)
IgM-RF positive (%)	42.6	53.3	54.2	72.0
ACPA positive (%)	52.0	66.2	59.4	87.4
<b>Radiographic scores</b>				
Erosion score 0 (% of patients)	93.4	92.9	91.0	91.3
Erosion score 1 (% of patients)	6.8	7.1	9.0	8.7
Total Sharp score at baseline	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.0 (0.0–0.0)
Total Sharp score at 3 years (n=175)	0.0 (0.0–4.0)	1.0 (0.0–5.0)	1.0 (0.0–4.5)	1.0 (0.0–6.0)

Results are shown as median (IQR) unless stated otherwise.

ACPA, anti-citrullinated protein antibody; ACR, American College of Rheumatology; CRP, C-reactive protein; DAS28, Disease Activity Score (28-joint count); ESR, erythrocyte sedimentation rate; EULAR, European League Against Rheumatism; HAQ, Health Assessment Questionnaire; RA, rheumatoid arthritis; RF, rheumatoid factor; SJC, swollen joint count; TJC, tender joint count.

**Table 2** Performance of three sets of classification criteria for rheumatoid arthritis (RA) or for prediction of erosive disease

	Gold standard								
	Methotrexate within 1 year			Expert opinion RA			Erosive disease at 3 years		
	2010 ACR/ EULAR	1987 ACR	Visser algorithm	2010 ACR/ EULAR	1987 ACR	Visser algorithm	2010 ACR/ EULAR	1987 ACR	Visser algorithm
Sensitivity	0.85	0.76	0.64	0.90	0.81	0.70	0.91	0.97	0.85
Specificity	0.50	0.59	0.71	0.48	0.59	0.73	0.21	0.17	0.41
PPV	0.86	0.87	0.86	0.79	0.81	0.84	0.22	0.22	0.26
NPV	0.49	0.41	0.37	0.69	0.59	0.54	0.91	0.96	0.92
LR+	1.70	1.85	2.21	1.73	1.98	2.59	1.16	1.17	1.45
LR-	0.30	0.41	0.51	0.21	0.32	0.41	0.42	0.17	0.36
AUC (95% CI)	0.78 (0.73 to 0.83)	0.71 (0.64 to 0.77)	0.75 (0.70 to 0.81)	0.80 (0.76 to 0.84)	0.74 (0.69 to 0.79)	0.77 (0.72 to 0.82)	0.63 (0.53 to 0.73)	0.67 (0.57 to 0.77)	0.64 (0.54 to 0.74)

ACR, American College of Rheumatology; AUC, area under the curve; EULAR, European League Against Rheumatism; LR, likelihood ratio; NPV, negative predictive value; PPV, positive predictive value.