Predictive factors for outcome of rheumatoid arthritis

Linden, M.P.M. van der

CHAPTER 13

Classification of rheumatoid arthritis: comparison of the 1987 ACR and 2010 ACR/EULAR criteria

M.P.M. van der Linden
R. Knevel
T.W.J. Huizinga
A.H.M. van der Helm-van Mil

Department of Rheumatology, Leiden University Medical Center, The Netherlands

Arthritis Rheum 2011; 63 (1): 37-42

ABSTRACT

Objective

New criteria to classify RA have been derived in order to increase the specificity and sensitivity for early RA compared to the 1987 ACR-criteria. This study evaluated differences in classification between the 1987 ACR-criteria and 2010 ACR/EULAR-criteria in an early arthritis cohort and determined the test characteristics of the 2010 ACR/EULAR-criteria.

Methods

2258 early arthritis patients included in the Leiden EAC cohort were studied. Fulfilment of the 1987- and 2010-criteria for RA was determined at baseline. The diagnosis at 1 year was assessed. The sensitivity and specificity of the 2010-criteria were determined using the following outcome measures: initiation of methotrexate-therapy or any DMARD-therapy during the first year of follow-up and having persistent arthritis during 5 years of follow-up.

Results

At first presentation, 1099 patients fulfilled the 2010-criteria and 726 patients the 1987-criteria for RA. 82 of the 726 patients fulfilling the 1987-criteria did not fulfill the 2010-criteria. 68% of the patients that fulfilled the 1987-criteria during the first year of the disease but not at baseline, did fulfill the 2010-criteria at baseline. The 2010 classification also led in 18% to a revoked classification at year 1. The sensitivity and the specificity were 0.84 and 0.60 with methotrexate therapy as outcome and 0.74 and 0.74 with DMARD therapy as outcome. With arthritis persistency as outcome, the sensitivity and specificity were 0.71 and 0.65.

Conclusion

Compared to the 1987-criteria, the 2010-criteria classify more patients with RA and at an earlier phase. The discriminative ability of the 2010 criteria is acceptable.

INTRODUCTION

During the last decades the focus of the management of RA has shifted to the early phase of the disease. This change was fuelled by studies showing that early achievement of low disease activity states is beneficial for the further course of RA. These studies raised the awareness on the importance of treating early and pointed to the relevance of early recognition of RA. From this perspective, the 1987 ACR criteria for RA[1] have been criticized as they are not equipped to diagnose RA early. This is not surprising as they have been developed in order to define homogeneous patient groups for research purposes and therefore were based on patients with average disease duration of 7 years.

In order to be able to identify early RA patients for clinical trials and other studies new classification criteria for RA have been derived by a task force of experts from both the EULAR and ACR.[2] The main purpose of these 2010 criteria is to achieve an increased sensitivity and specificity for RA in an early phase.

At present the diagnostic and discriminative abilities of these new criteria are not known. In addition it is unclear how the 2010 criteria behave in relation to the 1987 ACR criteria. This is especially relevant because the Working group that developed the 2010 criteria stressed in their discussion that the patients fulfilling the 2010 criteria are probably less homogeneous and that therefore in clinical trials researchers should document both the proportions of study subjects that fulfill the previous (1987) and new RA classification criteria to enable comparisons. Moreover, the working group warned that the 2010 criteria may probably increase heterogeneity by including different phenotypes, thereby making basic science studies more difficult. Therefore, the present study evaluated the following questions: What proportion of early arthritis patients that do not fulfill the 1987 criteria can according to the 2010 criteria be classified as RA? Do all early arthritis patients that fulfill the 1987 criteria fulfill the 2010 criteria as well? Do RA patients indeed fulfill the 2010 criteria at an earlier point in time than the 1987 criteria? In addition, the sensitivity and specificity of the 2010 criteria for RA were assessed; for this analysis three outcome measures were studied: initiation of methotrexate (MTX), initiation of any disease-modifying-antirheumatic-drug (DMARD) and having persistent arthritis over a 5 years follow-up period.

PATIENTS AND METHODS

Patients

The early arthritis patients studied are from the Leiden Early Arthritis Clinic (EAC) cohort, a population-based prospective cohort that started in 1993.[3] Inclusion took place when arthritis was confirmed at physical examination and symptom duration was <2 years. The inclusion criteria had not changed over time. At the first visit, patients and rheumatologists completed questionnaires, physical examination was performed, radiographs were taken and blood was taken for determination of amongst other C-reactive protein (CRP), erythrocyte sedimenta-

tion rate (ESR), IgM-rheumatoid factor (RF) and ACPA (anti-CCP2, Immunoscan RA Mark 2; Euro-Diagnostica, Arnhem, The Netherlands). Follow-up visits were performed yearly. Written informed consent was obtained from all participants. The study was approved by the local Medical Ethical Committee.

2258 early arthritis patients with at least one year of follow-up were included between 1993 and February 2009. Of these, 1632 were formerly classified as RA (1987 ACR criteria) or undifferentiated arthritis. The remaining 626 early arthritis patients were classified with other diagnoses. The treatment of patients with RA differed; hydroxychloroquine, penicillamine or suphasalazine were the initial DMARDs in the '90s, methotrexate was the initial DMARD since 1999.[4] Patients that were classified for other diagnoses than RA were treated accordingly. The treatment of patients that were undifferentiated was not protocollized.

Application of the 2010 criteria

The 2010 ACR/EULAR criteria were applied as described by Aletaha et al.[2] We used the 66-swollen joint count and 68-tender joint count. According to the guideline, the distal phalangeal joints, 1st carpo-metacarpal joint and 1st metatarso-phalangeal joints were excluded from assessment. Involvement of interphalangeal joints of the feet was considered as small joint involvement. The reference value for RF positivity in our cohort is 5, therefore a level ≥15 was considered high positive. Similarly, the reference value for anti-CCP2 positivity is 25 AU in our cohort and a level of ≥75 AU was considered high positive. An abnormal CRP was according to the reference value defined as ≥10 mg/l, and an abnormal ESR was ≥25 mm/hr for females and ≥15 mm/hr for males. In the new criteria it is stated that presence of a significant erosion is *prima facie* evidence of RA which precludes the need for applying other criteria. However, it is not yet agreed on what size, number or site of erosions is necessary to define erosiveness. Because of this uncertainty, we initially did not consider radiological information when applying the 2010 ACR/EULAR criteria. Afterwards analyses were repeated when patients with a SHS erosionscore ≥2 were classified as having RA as well, irrespective of fulfilling any of the other criteria. In addition, the effect of evaluating 44 or 28 instead of 66/68 joints was assessed.

Analysis

The following analyses were done. First, baseline characteristics of all 2258 early arthritis patients were studied to define the proportions of early arthritis patients that were classified as RA according to the 2010 criteria and the 1987 ACR-criteria. It was assessed whether patients that were diagnosed with RA using the 1987 ACR criteria fulfilled the 2010 criteria as well.

In order to determine whether the 2010 criteria are indeed fulfilled in an earlier stage than the 1987 criteria, patients that did fulfill the 1987 criteria during the first year of disease but not at first presentation were studied (n=297). It was determined how many of these patients already fulfilled the 2010 ACR criteria at baseline, and thus were indeed recognized in a more early phase by the 2010 criteria.

It was also evaluated whether the 2010 criteria would yield "false-positive" classifications. To this end, patients that at baseline fulfilled the 2010 ACR criteria were studied for their diagnosis after one year to determine whether they were classified differently at that time-point.

Finally, the sensitivity and specificity of the 2010 criteria were determined and the area under the receiver operator characteristic curve (AUC) assessed in the patients that were formerly classified as RA or undifferentiated arthritis (n=1632). Patients that received DMARD treatment in a randomized trial were not studied, leaving 1404 regularly treated patients for evaluation. Three outcome measures were used. First, initiation of methotrexate therapy within the first year of follow-up, the same outcome measure as used for the derivation of the 2010 criteria. Since methotrexate was not the anchor drug before 1999, initiation of any DMARD within the first year was studied as well. Thirdly, in the subgroup of patients that achieved 5 years of follow-up (n=790), arthritis persistency was assessed and defined by the absence of a sustained DMARD-free remission. Patients were defined as being in remission if DMARD therapy could be discontinued and no synovitis was detected for at least one year.[5] Analyses were done using SPSS (version 17.0).

RESULTS

The baseline characteristics of all early arthritis patients are presented in Table 1. The characteristics of the subset of patients that at baseline were classified as RA according to the 1987 and 2010 criteria are presented as well.

Agreement in classification

At baseline, 1099 out of 2258 early arthritis patients fulfilled the 2010 criteria for RA. 726 patients fulfilled the 1987 ACR criteria for RA. From these 726 patients, 644 (88.7%) also fulfilled the 2010 criteria whereas 82 (11.3%) patients did not fulfill the 2010 criteria. From the 1099 patients that fulfilled the 2010 criteria, 455 patients did not fulfill the 1987 criteria (Table 2A). From the 1099 patients that fulfilled the 2010 criteria, 455 patients did not fulfill the 1987 criteria (Table 2A). The agreement in classification criteria was not different when patients included before of after 1999 were studied separately (data not shown). Characteristics of the patients that fulfilled both the 1987 and 2010 criteria and that fulfilled the 1987 but not the 2010 criteria are presented in Supplementary Table 1.

Baseline classification in relation to the disease course

297 patients fulfilled the 1987 ACR criteria during the first year, but not at baseline. From these, 202 (68.0%) did fulfill the 2010 criteria at baseline, indicating that the 2010 criteria indeed classify RA patients in an earlier phase of the disease.

The 1099 early arthritis patients that fulfilled the 2010 ACR criteria at baseline were studied for their diagnosis at year 1. In 194 cases patients were classified differently at that time-point. Study-

Table 1. Baseline characteristics of all early arthritis patients, the subset of early arthritis patients that fulfilled the 1987 ACR criteria and the subset that fulfilled the 2010 ACR/EULAR criteria at first presentation

Characteristics	All early arthritis pts (n=2258)	1987 RA (n=726)	2010 RA[$] (n=1099)
Age at inclusion (yrs), mean (SD)	51.9 (17.2)	57.4 (16.3)	56.1 (16.4)
Female, N (%)	1340 (59.3)	470 (64.7)	718(65.3)
Symptom duration at first presentation, weeks, mean (SD)[‡]	25.9 (41.6)	31.6 (36.3)	29.8 (43.4)
< 6 weeks, N (%)[$]	*436 (21.4)*	*0 (0)[†]*	*98 (9.6)*
≥ 6 weeks, N (%)[$]	*1602 (78.6)*	*726 (100)[†]*	*925 (90.4)*
66 Swollen joint count, mean (SD)	6.5 (6.8)	11.6 (7.3)	10.3 (7.7)
1 medium-large joint, N (%)[$]	*253 (11.2)*	*0 (0)*	*0 (0)*
2-10 medium-large joints, N (%)[$]	*142 (6.3)*	*0 (0)*	*3 (0.3)*
1-3 small joints, N (%)[$]	*532 (23.6)*	*48 (6.6)*	*106 (9.6)*
4-10 small joints, N (%)[$]	*560 (24.8)*	*206 (28.4)*	*236 (21.5)*
> 10 joints, N (%)[$]	*771 (34.1)*	*472 (65.0)*	*754 (68.6)*
ESR (mm/hr), mean (SD)[‡]	33.2 (28.1)	40.3 (28.2)	38.3 (28.0)
CRP (mg/l), mean (SD)[‡]	27.1 (28.1)	28.2 (35.2)	28.3 (35.1)
Normal CRP and ESR, N (%)[$]	*747 (33.1)*	*147 (20.2)*	*251 (22.8)*
Abnormal CRP or ESR, N (%)[$]	*1511 (66.9)*	*579 (79.8)*	*848 (77.2)*
RF positive, N (%)[‡]	671 (30.1)	399 (55.0)	601 (55.3)
Anti-CCP2-positive, N (%)[‡]	506 (29.7)	323 (51.4)	472 (52.2)
Negative RF and ACPA, N (%)[$]	*1484 (65.7)*	*285 (39.3)*	*412 (37.5)*
Low positive RF or ACPA, N (%)[$]	*203 (9.0)*	*78 (10.7)*	*145 (13.2)*
High positive RF or ACPA, N (%)[$]	*571 (25.3)*	*363 (50.0)*	*542 (49.3)*
Erosive, N (%)	590 (26.1)	392 (54.0)	467 (42.5)

[$]applied without considering data on erosiveness at baseline. [‡]Data missing for analysis (n): symptom duration (220); ESR (15); CRP (219); RF (32); anti-CCP2 (553). [†]According to the 1987 ACR criteria, 4/7 criteria have to exist for >6 weeks to be a valid criterion. [$]Subdivision of criteria according to the score based algorithm from the 2010 ACR/EULAR criteria

ing the medical records of these patients confirmed that these patients clinically had evidently another diagnosis than RA. The diagnoses of these patients were: psoriatic arthritis (n=46), inflammatory osteoarthritis (n=28), reactive arthritis (n=20), RS3PE (n=17), sarcoidosis (n=15), (pseudo)gout (n=15), para-malignant arthritis (n=6), spondylarthropathy (n=6), SLE (n=10), MCTD (3), other systemic disorders (n=21) and other diagnoses (n=7). These patients concern 17.7% of the total population of patients fulfilling the 2010 ACR criteria and 27.7% of the patients that at baseline did fulfill the 2010 criteria but not the 1987 criteria.

In the description of the 2010 criteria is stated that these should be applied only in case no other diagnosis can be established. Thus this means that in case a patient can be classified with two diagnoses, the other diagnosis prevails. Therefore we repeated the analysis presented above in the patients that were formerly classified as RA or undifferentiated arthritis (n=1632). Of

Table 2. Classification according to the 1987 and 2010 criteria for RA without A) and with B) including radiologic information on erosiveness when applying the 2010 ACR/EULAR criteria

		(A) 2010 ACR/EULAR Classification Criteria[§]		(B) 2010 ACR/EULAR Classification Criteria[†]		
		RA at baseline	no RA baseline	RA at baseline	no RA baseline	*Total*
1987 ACR Classification Criteria	RA at baseline	644	82	678	48	*726*
	no RA at baseline	455	1077	544	988	*1532*
	Total	*1099*	*1159*	*1222*	*1036*	*2258*

[§]application of the 2010 ACR/EULAR criteria without the use of radiologic information on erosiveness. [†]application of the 2010 ACR/EULAR criteria with the use of radiologic information on erosiveness and defining erosiveness as a total SHS erosion score ≥2. Criteria were applied using data on 66/68 joints for swelling/tenderness

these, 939 patients fulfilled the 2010 ACR criteria at baseline and were studied for their diagnosis at year 1. In 88 cases patients were classified differently at that time-point.

Studying the medical records of these patients confirmed that these patients clinically had evidently another diagnosis than RA. The diagnoses of these patients then were: psoriatic arthritis (n=20), inflammatory osteoarthritis (n=13), reactive arthritis (n=7), RS3PE (n=7), (pseudo)gout (n=7), SLE (n=6), para-malignant arthritis (n=4), spondylarthropathy (n=4), sarcoidosis (n=3), MCTD (2), other systemic disorders (n=12) and other diagnoses (n=3). These patients concern 9.4% of the 939 patients that fulfilled the 2010 ACR criteria and 14.1% of the patients that at baseline did fulfill the 2010 criteria but not the 1987 criteria.

Test characteristics of the 2010 criteria

When using initiation of methotrexate-therapy within the first year as outcome the sensitivity and specificity of the 2010 criteria were 0.84 and 0.60. With initiation of any DMARD-therapy within the first year as outcome the sensitivity and specificity were 0.74 and 0.74. The AUCs when using these two outcomes were 0.72 and 0.74 respectively. The third outcome measure was arthritis persistency over 5 years. Here the sensitivity of the 2010 criteria was 0.71, the specificity 0.65 and the AUC 0.65 (Table 3A).

Value of baseline erosiveness

It is unclear what number of erosions in early arthritis patients is specific for early RA. In the evaluation of the consequences of considering erosiveness, here defined as total SHS erosion score ≥2, we observed that 1222 patients were at baseline classified as RA according to the 2010 criteria. Thus when including erosiveness in the evaluation, 123 (5.4%) early arthritis patients were additionally classified as RA. The analyses on the agreement in classification and on the test

Table 3. Test characteristics of the 1987 and 2010 criteria using three outcome measures: methotrexate initiation during the first year, initiation of any DMARD during the first year and having disease persistency over 5 years. Analyses were done without (A) and with (B) including radiologic information on erosiveness when applying the 2010 ACR/EULAR criteria

	Outcome Measure								
	MTX-initiation			DMARD-initiation			5-year Persistency		
Criteria Set	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
1987 ACR Classification Criteria	0.61	0.74	0.67	0.54	0.87	0.71	0.53	0.75	0.61
(A) 2010 ACR/EULAR Classification Criteria[§]	0.84	0.60	0.72	0.74	0.74	0.74	0.71	0.65	0.65
(B) 2010 ACR/EULAR Classification Criteria[†]	0.88	0.54	0.71	0.79	0.68	0.79	0.77	0.56	0.65

[§]application of the 2010 ACR/EULAR criteria without the use of radiologic information on erosiveness. [†]application of the 2010 ACR/EULAR criteria with the use of radiologic information on erosiveness and defining erosiveness as a total SHS erosion score ≥2

characteristics were repeated (Table 2B and 3B), but the results were not substantially different compared to the results when radiological information was disregarded.

Effect of number of assessed joints

We used a 66/68 count for swollen and tender joints. In clinical practice 44 or 28 joints may be evaluated more frequently. In order to determine whether this would results in different classification, the 2010 criteria were applied with the 44 and 28 joint counts. The numbers of patients classified as RA were then 1082 and 940 respectively, instead of 1099 (Supplementary Table 2). The test characteristics were fairly comparable when a lower number of joints was considered (Supplementary Table 3).

DISCUSSION

The present study compared classification of RA using the 1987 and 2010 criteria in a large early arthritis cohort. It was observed that the 2010 criteria classified more patients with RA than the 1987 criteria and that 11.3% of the patients with RA according to the 1987 criteria were not classified as RA according to the 2010 criteria. A large proportion of the early arthritis patients that developed RA according to the 1987 criteria during the disease course could at first presentation be classified as RA according to the 2010 criteria. This denotes that the 2010 criteria have come up to the demand of classifying RA in an earlier phase of the disease than the 1987 criteria. The 2010 classification also led in 18% (or 9.4% dependent on the studied population) to a revoked diagnosis in a later phase, substantiating the concerns with regard to increase in heterogeneity by use of the 2010 criteria. Compared to the 1987 criteria, the sensitivity of the 2010 criteria was higher but the specificity lower.

In this study several choices were made when applying the 2010 ACR/EULAR criteria. Initially we left out information on hand and feet radiographs as a clear description of erosive disease resembling RA was not provided by Aletaha et al.[2] Afterwards we repeated analyses defining total SHS erosionscore ≥2 as erosiveness. Fairly similar observations were done. This may suggest that evaluating radiographs in this early phase is not highly relevant for classification of RA. A moderate predictive ability of the number of erosive joints in early arthritis patients for RA development has been described recently.[6]

We used a 66/68 count for swollen and tender joints. In clinical practice 44 or 28 joint counts are frequently used. To evaluate the effect of assessing different numbers of joints we repeated the analyses with these joint counts. The number of patients classified with RA decreased but the test characteristics were only marginally affected.

This study has some limitations. First of all it is based on one inception cohort and more studies are needed to establish the sensitivity and specificity of the new criteria. A complicating factor is that it is somewhat ambiguous what outcome measure to take as the gold standard for RA. This has been subject to discussion within the working group that derived the 2010 criteria

and usage of methotrexate was chosen. This outcome may not be appropriate when studying older cohorts. For example in the 1990's it was not common practice to start methotrexate early in a patient with arthritis of recent onset that did not fulfill the 1987 criteria. For this reason we chose any DMARD-therapy instead of methotrexate-therapy as outcome. Additionally we chose a second outcome (arthritis persistency over 5 years) for verification. However, differences in the outcome measure may yield variations in observed test characteristics.

The three outcome measures used here (initiation of methotrexate or any DMARD or arthritis persistency) all contain risk of misclassification as these can also be fulfilled in case of other diagnoses, for example in psoriatic arthritis. Psoriatic arthritis was also the most frequent cause of a revoked classification at year 1.

Another consideration is that 213 early arthritis patients included in the EAC after 2000 were used in the derivation phase of the 2010 criteria. In the present study we evaluated a considerable larger number of patients, as well as two outcome measures additional to methotrexate usage. In order to see whether this subset of patients affected the results, analyses were repeated excluding the 213 patients. This did not yield substantially different results (data not shown). Nonetheless, validation of the 2010 criteria in other cohorts is required as well.

Given the emerging evidence on a "window of opportunity",[7] pointing to the need to treat as early as possible, the question is what method serves best to identify individual patients in an early phase of RA. The authors of the 2010 criteria underline that the new classification criteria were not developed as a diagnostic tool and that a separate body of work is needed to develop such tools.[2] Prediction rules aiming at early diagnosis have been developed and validated.[8,9] The question what method is best to identify early RA on the individual patient level is still open and a subject for future studies.

In conclusion, the 2010 criteria for RA classify more patients with RA and do so in an earlier phase. The discriminative ability of the 2010 ACR/EULAR criteria is reasonable, indicating that these criteria perform well to classify early RA.

REFERENCES

1. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988; 31(3):315-24.
2. Aletaha D, Neogi T, Silman A, Funovits J, Felson D, Bingham III CO et al. 2010 Rheumatoid arthritis classification criteria: An American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010; 62(9):2569-81.
3. van Aken J, van Bilsen JH, Allaart CF, Huizinga TW, Breedveld FC. The Leiden Early Arthritis Clinic. Clin Exp Rheumatol 2003; 21(5 Suppl 31):S100-S105.
4. van Nies JAB, de Jong Z, van der Helm-van Mil A.H.M., Knevel R, le Cessie S, Huizinga TWJ. Improved treatment strategies reduce the increased mortality risk in early RA patients. Rheumatology 2010; In Press.
5. van der Woude D, Young A, Jayakumar K, Mertens BJ, Toes REM, van der Heijde D et al. Prevalence of and predictive factors for sustained disease-modifying antirheumatic drug-free remission in rheumatoid arthritis: results from two large early arthritis cohorts. Arthritis Rheum 2009; 60(8):2262-71.
6. Thabet MM, Huizinga TW, van der Heijde DM, van der Helm-van Mil AH. The prognostic value of baseline erosions in undifferentiated arthritis. Arthritis Res Ther 2009; 11(5):R155.
7. Finckh A, Liang MH, van Herckenrode CM, de PP. Long-term impact of early treatment on radiographic progression in rheumatoid arthritis: A meta-analysis. Arthritis Rheum 2006; 55(6):864-72.
8. van der Helm-van Mil AHM, le Cessie S, van Dongen H, Breedveld FC, Toes REM, Huizinga TWJ. A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. Arthritis Rheum 2007; 56(2):433-40.
9. van der Helm-van Mil AHM, Detert J, le Cessie S, Filer A, Bastian H, Burmester GR et al. Validation of a prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: moving toward individualized treatment decision-making. Arthritis Rheum 2008; 58(8):2241-7.

Supplementary Table 1. Baseline characteristics of all early arthritis patients, the subsets of early arthritis patients that fulfilled the 1987 ACR criteria but not the 2010 criteria and that fulfilled both the 1987 and the 2010 ACR/EULAR criteria at first presentation. Data are presented without and with assessing radiologic information on erosiveness when applying the 2010 ACR/EULAR criteria

Characteristics	All early arthritis pts (n=2258)	1987+/2010- (n=82)[§]	1987+/2010+ (n=644)[§]	1987+/2010- (n=48)[†]	1987+/2010+ (n=678)[†]
Age at inclusion (yrs), mean (SD)	51.9 (17.2)	55.8 (17.4)	57.6 (16.2)	53.3 (18.8)	57.7 (16.1)
Female, N (%)	1340 (59.3)	51 (62.2)	419 (65.1)	29 (60.4)	441 (65.0)
Symptom duration at first presentation in weeks, mean (SD)	25.9 (41.6)	25.4 (23.1)	32.4 (37.6)	26.3 (24.6)	32.0 (37.0)
< 6 weeks, N (%)	*436 (21.4)*	*0 (0)[†]*	*0 (0)*	*0 (0)*	*0 (0)*
≥ 6 weeks, N (%)	*1602 (78.6)*	*79 (100)[†]*	*626 (100)*	*47 (100)*	*658 (100)*
66 Swollen joint count, mean (SD)	6.5 (6.8)	5.9 (2.2)	12.3 (7.4)	5.6 (2.1)	12.3 (7.4)
1 medium-large joint, N (%)	*253 (11.2)*	*0 (0)*	*0 (0)*	*0 (0)*	*0 (0)*
2-10 medium-large joints, N (%)	*142 (6.3)*	*0 (0)*	*0 (0)*	*0 (0)*	*0 (0)*
1-3 small joints, N (%)	*532 (23.6)*	*19 (23.2)*	*29 (4.5)*	*10 (20.8)*	*38 (5.6)*
4-10 small joints, N (%)	*560 (24.8)*	*63 (76.8)*	*143 (22.2)*	*38 (79.2)*	*168 (24.8)*
> 10 joints, N (%)	*771 (34.1)*	*0 (0)*	*472 (73.3)*	*0 (0)*	*472 (69.9)*
ESR (mm/hr), mean (SD)	33.2 (28.1)	32.9 (23.9)	41.3 (28.5)	32.3 (26.7)	40.8 (28.2)
CRP (mg/l), mean (SD)	27.1 (28.1)	27.5 (29.5)	31.6 (35.9)	23.3 (29.5)	31.7 (35.5)
Normal CRP and ESR, N (%)	*747 (33.1)*	*23 (28.0)*	*124 (19.3)*	*16 (33.3)*	*275 (40.9)*
Abnormal CRP or ESR, N (%)	*1511 (66.9)*	*72.0 (72.0)*	*520 (80.7)*	*32 (66.7)*	*398 (59.1)*
RF positive, N (%)	671 (30.1)	1 (1.2)	399 (62.4)	1 (2.1)	399 (62.4)
Anti-CCP2-positive, N (%)	506 (29.7)	2 (3.0)	321 (57.1)	2 (5.7)	321 (57.1)
Negative RF and ACPA, N (%)	*1484 (65.7)*	*80 (97.6)*	*205 (31.8)*	*45 (93.8)*	*240 (35.4)*
Low positive RF or ACPA, N (%)	*203 (9.0)*	*0 (0.0)*	*76 (11.8)*	*0 (0)*	*75(11.1)*
High positive RF or ACPA, N (%)	*571 (25.3)*	*3 (3.7)*	*363 (56.4)*	*3 (6.3)*	*363 (53.5)*
Erosive, N (%)	590 (26.1)	34 (41.5)	356 (55.3)	0 (0)	390 (57.5)

[§]application of the 2010 ACR/EULAR criteria without the use of radiologic information on erosiveness. [†]application of the 2010 ACR/EULAR criteria with the use of radiologic information on erosiveness and defining erosiveness as a total SHS erosion score ≥2

Supplementary Table 2. Classification according to the 1987 and 2010 criteria for RA without (A) and with (B) including radiologic information on erosiveness when applying the 2010 ACR/EULAR criteria

			(A) 2010 ACR/EULAR Classification Criteria[§]		(B) 2010 ACR/EULAR Classification Criteria[†]		
			RA at baseline	no RA baseline	RA at baseline	no RA baseline	*Total*
Assessing 44 joints[‡]	1987 ACR classification Criteria	RA at baseline	639	87	677	49	*726*
		no RA at baseline	443	1089	533	999	*1532*
		Total	*1082*	*1176*	*1210*	*1048*	*2258*
Assessing 28 joints[$]	1987 ACR classification Criteria	RA at baseline	603	123	653	73	*726*
		no RA at baseline	337	1195	443	1089	*1532*
		Total	*940*	*1318*	*1096*	*1162*	*2258*

[‡]Criteria were applied using the 44 joint counts data. [$]Criteria were applied using the 28-swollen joint counts. [§]application of the 2010 ACR/EULAR criteria without the use of radiologic information on erosiveness [†]application of the 2010 ACR/EULAR criteria with the use of radiologic information on erosiveness and defining erosiveness as a total SHS erosion score ≥2

Supplementary Table 3. Test characteristics of the 1987 and 2010 criteria using methotrexate initiation during the first year, initiation of any DMARD during the first year and having disease persistency over 5 years as outcome measures. Analyses were done without (A) and with (B) including radiologic information on erosiveness when applying the 2010 ACR/EULAR criteria

		Outcome Measure								
		MTX-initiation			DMARD-initiation			Persistency over 5 years		
	Criteria Set	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
Assessing 44 joints[‡]	1987 ACR Classification Criteria	0.61	0.74	0.67	0.54	0.87	0.71	0.53	0.75	0.61
	(A) 2010 ACR/EULAR Classification Criteria[§]	0.83	0.60	0.71	0.73	0.74	0.74	0.71	0.65	0.65
	(B) 2010 ACR/EULAR Classification Criteria[†]	0.87	0.54	0.71	0.79	0.68	0.74	0.77	0.56	0.65
Assessing 28 joints[§]	1987 ACR Classification Criteria	0.61	0.74	0.67	0.54	0.87	0.71	0.53	0.75	0.61
	A) 2010 ACR/EULAR Classification Criteria[§]	0.83	0.60	0.71	0.73	0.74	0.72	0.65	0.69	0.64
	B) 2010 ACR/EULAR Classification Criteria[†]	0.87	0.54	0.71	0.79	0.68	0.73	0.74	0.59	0.64

[‡]Criteria were applied using the 44-joint count. [§]Criteria were applied using the 28- joint count. [§]application of the 2010 ACR/EULAR criteria without the use of radiologic information on erosiveness. [†]application of the 2010 ACR/EULAR criteria with the use of radiologic information on erosiveness and defining erosiveness as a total SHS erosion score ≥2