

# Diagnostic Accuracy of ACR/EULAR 2010 Criteria for Rheumatoid Arthritis in a 2-Year Cohort

SOPHIE VARACHE, DIVI CORNEC, JOHANNE MORVAN, VALÉRIE DEVAUCHELLE-PENSEC, JEAN-MARIE BERTHELOT, CATHERINE LE HENAFF-BOURHIS, SYLVIE HOANG, JEAN-BAPTISTE THOREL, ANTOINE MARTIN, GÉRARD CHALÈS, EMMANUEL NOWAK, SANDRINE JOUSSE-JOULIN, PIERRE YOUINOU, and ALAIN SARAUX

**ABSTRACT. Objective.** To evaluate the diagnostic accuracy of the 2010 American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) and 1987 ACR criteria for rheumatoid arthritis (RA), and the respective role of the algorithm and scoring of the ACR/EULAR.

**Methods.** In total, 270 patients with recent-onset arthritis of < 1 year duration were included prospectively between 1995 and 1997 and followed for 2 years. RA was defined as the combination, at completion of followup, of RA diagnosed by an office-based rheumatologist and treatment with a disease-modifying antirheumatic drug or glucocorticoid. We compared the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the criteria sets in the overall population, in the subgroup meeting the tree condition for ACR/EULAR scoring, and in the overall population classified according the full tree.

**Results.** At baseline, 111 of the 270 patients had better alternative diagnoses and 16 had erosions typical for RA; of the 143 remaining patients, 52 had more than 6 ACR/EULAR 2010 points (indicating definite RA) and 91 had fewer than 6 points. After 2 years, 11/16 patients with erosions and 40/52 with more than 6 points had RA. 100 of the 270 patients met the reference standard for RA. Sensitivity, specificity, PPV, and NPV of the ACR/EULAR (full tree) were 51/100 (51%), 153/170 (90%), 51/68 (75.4%), and 153/202 (75.7%), respectively. Diagnostic accuracies of the ACR/EULAR score and ACR 1987 criteria were not statistically different.

**Conclusion.** Much of the improvement of the ACR/EULAR criteria was ascribable to the use of exclusion criteria in the algorithm. (J Rheumatol First Release May 15 2011; doi:10.3899/jrheum.101227)

## Key Indexing Terms:

RHEUMATOID ARTHRITIS      DIAGNOSIS      CRITERIA      EARLY ARTHRITIS

In 1958, the American Rheumatism Association developed a set of classification criteria for rheumatoid arthritis (RA)<sup>1</sup>, which were revised by the American College of Rheumatism (ACR) in 1987<sup>2</sup>. These criteria were designed

*From the Rheumatology Unit, University Hospital, Brest; Rheumatology Unit, University Hospital, Nantes; Rheumatology Unit, Morlaix Hospital, Morlaix; Rheumatology Unit, Vannes Hospital, Vannes; Rheumatology Unit, Lorient Hospital, Lorient; Rheumatology Unit, Saint-Brieuc Hospital, Saint-Brieuc; Rheumatology Unit, University Hospital, Rennes; and the Center of Investigation and Immunology Unit, University Hospital, Brest, France.*

*Supported by the Brest Hospital Center and the 1995 Clinical Research Hospital Program (PHRC 1995).*

*S. Varache, MD, Fellow; D. Cornec, MD, Fellow; J. Morvan, MD; V. Devauchelle-Pensec, MD, PhD, University Hospital, Brest; J.-M. Berthelot, MD, University Hospital, Nantes; C. Le Henaff-Bourhis, MD, Morlaix Hospital; S. Hoang, MD, Vannes Hospital; J.-B. Thorel, MD, Lorient Hospital; A. Martin, MD, Saint-Brieuc Hospital; G. Chalès, MD, University Hospital, Rennes; E. Nowak, PhD, Center of Investigation, University Hospital, Brest; S. Jousse-Joulin, MD, University Hospital, Brest; P. Youinou, MD, DSc, Department of Immunology, University Hospital, Brest; A. Saraux, MD, PhD, University Hospital, Brest.*

*Address correspondence to Prof. A. Saraux, Rheumatology Unit, Hôpital de la Cavale Blanche, BP 824, F 29609 Brest cedex, France. E-mail: Alain.Saraux@univ-brest.fr*

*Accepted for publication February 11, 2011.*

to distinguish established RA from other joint diseases. They were widely used as the reference standard for diagnosing RA in clinical practice<sup>3,4</sup>. However, although the ACR criteria are easy to use, they are not ideal for diagnosing RA, for several reasons<sup>5</sup>: they were designed for classification, not diagnosis; most of the patients used to develop the criteria had long-standing RA; the predictive value of each criterion cannot be determined because the numbers of RA patients and controls were preestablished; and no exclusion criteria are used. Since the development of the 1987 ACR criteria, both nodules<sup>6</sup> and radiographic findings<sup>7</sup> at first evaluation have been shown to be of limited diagnostic usefulness. In addition, anticyclic citrullinated peptide (CCP) antibodies (ACPA) have emerged as a valuable diagnostic marker for RA<sup>8</sup>. Changes to the 1987 ACR criteria that were made to take these new findings into account improved the sensitivity of the ACR 1987 criteria, most notably in patients with arthritis of less than 6 months' duration<sup>9</sup>. However, further improvement was needed, given the importance of early treatment for improving outcomes in patients with RA. The early diagnosis of RA is often difficult, as none of the clinical or laboratory features is diagnostic.

To construct diagnostic criteria, a longitudinal study must be performed in patients with early-stage arthritis, to determine which combination of features at baseline best predicts a diagnosis of RA at completion of the study. Defining the endpoint, i.e., diagnosis of RA at study completion, is challenging. Rheumatologists often differ on whether a given patient does or does not have RA<sup>10</sup>. One means of circumventing this difficulty may be to use an outcome other than a diagnosis of RA, for instance, persistent nonerosive or erosive arthritis<sup>11,12</sup>. However, RA can be defined unequivocally after a 2 year followup as a physician-made diagnosis of polyarthritis, no other diagnosis capable of explaining the symptoms, and treatment with disease-modifying antirheumatic drugs (DMARD). This definition was used by the ACR/EULAR to develop a new scoring system for RA<sup>13,14</sup>. This new system has a tree format: presence of synovitis is required (condition 1), followed by absence of a better alternative diagnosis (condition 2), and then by absence of erosions typical for RA (condition 3). Only patients meeting all 3 conditions are eligible for scoring. A score of at least 6 of 10 possible points from scores in 4 domains indicates RA.

We established a cohort of 270 patients from Brittany, France, with arthritis of less than 1 year duration, included between 1995 and 1997 and followed for 2 years. These patients did not contribute to the development of the ACR/EULAR criteria. Therefore, the cohort provides an opportunity to evaluate the diagnostic accuracy of the ACR/EULAR criteria. The main problem raised by such an evaluation is the choice of the reference standard. The opinion of rheumatologists regarding a diagnosis of RA is informed by the criteria that are tested. Consequently, there is circularity in the reasoning. For development of the ACR/EULAR criteria, RA was considered in patients taking DMARD (methotrexate) therapy and confirmed by expert opinion. In our study, the reference standard was a combination of having a physician-made diagnosis of RA, no other diagnosis of joint disease, and treatment with DMARD and/or glucocorticoid after 2 years' followup.

The aim of our study was to evaluate the diagnostic accuracy of the ACR/EULAR criteria comparatively with the ACR 1987 criteria, using this reference standard, in our cohort of patients with early arthritis. In addition, we sought to determine the combination of baseline features that best predicted RA after 2 years.

## MATERIALS AND METHODS

**Study population.** This prospective observational cohort was composed of all patients seen with early arthritis from 1995 to 1997 in 7 hospitals in Brittany, France. All patients in the cohort had synovitis in at least one joint (condition 1 in the ACR/EULAR tree).

All the patients were referred to us by general practitioners and rheumatologists who had been informed of the study. Inclusion criteria were as follows: age 18 years or older, synovitis in at least one joint, absence of a previous diagnosis of joint disease, and disease duration no more than 1 year. Patients were excluded if the medical history and the physical examination suggested septic arthritis or crystal-induced arthritis. Synovitis was diag-

nosed clinically by a rheumatologist based on the presence of joint swelling with tenderness or decreased range of motion.

The appropriate ethics committee approved the study, and all patients gave their written informed consent before inclusion in the study.

**Baseline assessment.** As described<sup>6</sup>, all patients had a standardized interview, a general physical examination, and laboratory tests during which over 100 variables were measured.

**Study variables.** All items of the unmodified and modified ACR 1987 criteria sets [criteria: 1. morning stiffness  $\geq$  1 hour, 2. arthritis of  $\geq$  3 joint areas, 3. arthritis of hand joints, 4. symmetric arthritis, 5. rheumatoid nodules, 6. rheumatoid factor (positive), 7. radiographic changes, and 8. anti-CCP antibodies (positive)]; and the ACR/EULAR criteria set were evaluated in each patient. Unmodified ACR 1987 criteria were considered positive in patients with at least 4 of criteria 1 through 7<sup>2</sup>, and modified ACR 1987 criteria were considered positive in patients with at least 4 of the 8 criteria (Liao 1) or at least 3 of the 6 criteria left after eliminating criteria 5 and 7 (Liao 2)<sup>9</sup>.

For the ACR/EULAR 2010 criteria, we evaluated the presence of synovitis (condition 1), absence of a better alternative diagnosis (condition 2), and then absence of erosions typical for RA (condition 3). Only patients meeting all 3 conditions were eligible for scoring.

Patients were deemed to have RA according to the ACR/EULAR 2010 criteria<sup>14</sup> if they had erosions typical for RA or a score of at least 6 of 10 possible points from scores in 4 domains as follows:

Joint involvement (1 medium-large joint: 0 point; 2–10 medium-large joints: 1 point; 1–3 small joints: 2 points; 4–10 small joints: 3 points; 10 joints with at least one small joint: 5 points);

Serology [no rheumatoid factor (RF) or ACPA: 0 point; low-positive ( $<$  3 times the upper limit of normal [ULN] for the laboratory and assay) RF and/or ACPA: 1 point; high-positive ( $>$  3 times ULN for the laboratory and assay) RF and/or ACPA: 3 points];

Duration of synovitis ( $<$  6 weeks: 0 point;  $\geq$  6 weeks: 1 point); and

Acute-phase reactants [C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) normal: 0 point; CRP and/or ESR elevated: 1 point].

**Radiographic evaluation.** Baseline hand radiographs ( $n = 258$ ) were read by one author (VD), who was blinded to information about the patients. A standardized evaluation procedure was used to record typical erosions of RA, as described<sup>15</sup>. The intraobserver variability was determined by evaluation of the same radiographs twice, 3 months apart; the intraobserver kappa coefficient was 0.88.

**Followup.** A rheumatologist followed each patient at 6-month intervals until a clinical diagnosis of a specific joint disease was made and the patient met published criteria for the same joint disease. The patients were asked to attend a final visit between June and November 1999.

**Outcome after 2 years.** The diagnoses made in each patient after 2 years by the hospital-based rheumatologists were recorded.

We had previously evaluated the concordance between the office-based rheumatologists and the diagnosis of a panel of 5 rheumatologists<sup>6</sup>, and we observed a good concordance (kappa 0.81, 95% CI 0.77–0.85). On this basis, we chose the diagnosis of the office-based rheumatologist; but, as it is not important to predict RA benign enough not to be treated, in this study we added a treatment with DMARD or glucocorticoid 2 years later as the reference. A total of 16 patients considered to have RA by the rheumatologist did not receive either DMARD or glucocorticoid therapy (the number of “diagnosis of RA and DMARD or glucocorticoid therapy” was 100 in our study, instead of 116 using only “the diagnosis of RA”). In contrast, 50 patients considered non-RA received DMARD or glucocorticoid therapy. Thus, the concordance between “diagnosis of RA” and “DMARD or glucocorticoid therapy” was fair (kappa = 0.52).

Thus, for evaluation of diagnostic accuracy of the ACR/EULAR and ACR 1987 criteria sets, the definition of RA (reference standard) was a diagnosis of RA by the office-based rheumatologist after 2 years combined with treatment with DMARD or glucocorticoid. There was no protocol for deciding to start or stop DMARD therapy.

**Statistical analysis.** First, in the subgroups that did and did not meet the reference standard definition of RA after 2 years, we evaluated the proportion of patients having each item of the unmodified ACR 1987 criteria set<sup>2</sup>, the modified ACR 1987 criteria set<sup>9</sup>, and the ACR/EULAR criteria set.

The only subgroup adequate for evaluating the criteria sets was the population of patients meeting all 3 ACR/EULAR scoring criteria (synovitis, no radiographic evidence of RA, and no other diagnosis more likely than RA). We therefore compared the criteria sets in this sample. To determine which item or items best separated patients with and those without RA in this sample and then to check whether these items matched those considered by the experts in the ACR/EULAR criteria set, we performed multiple logistic regression with backward selection using the likelihood ratio test.

As the ACR 1987 criteria are used frequently in cohorts of patients with early arthritis to separate patients with and without RA after some time, irrespective of whether the patients have radiographic evidence of RA and/or another more plausible diagnosis, we also compared the diagnostic accuracy of the ACR/EULAR scoring system with the diagnostic value of the ACR 1987 criteria used without patient selection based on conditions 1–3 in the overall population.

Statistical tests were performed using the Statistical Package for the Social Sciences (SPSS 13.0, 2005; SPSS Inc., Chicago, IL, USA) except for comparisons of areas under the curve (AUC), which were done using SAS 9.0 (SAS Institute, Cary, NC, USA). Quantitative variables are

described as means ± standard deviation and qualitative variables as number (percentage). To compare the distributions of laboratory test results between patients with and without RA after 2 years, we used the chi-square test (or Fisher exact test, if appropriate) for qualitative variables and the Mann-Whitney test for quantitative variables. For each criterion, sensitivity was plotted against 1 – specificity to obtain the receiver operating characteristic (ROC) curve<sup>16</sup>. A ROC curve was plotted for each criteria set by varying the cutoffs (values lower than the cutoff considered negative and other values positive), and AUC were compared. For each criteria set used at baseline, we computed sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), with 95% confidence intervals, for predicting RA after 2 years.

## RESULTS

**Study population.** The study cohort comprised 270 patients with arthritis of < 1 year duration. Median followup was 30 months [< 1 year in 16 cases (6%), 1 to 2 years in 21 (8%), and more than 2 years in 233 cases (86%)]. After 2 years, 100/270 (37%) patients met our reference standard definition of RA. Table 1 shows the criteria present at baseline in patients with and without RA after 2 years.

*Table 1.* Prevalence of each item of the unmodified ACR 1987 criteria, modified ACR 1987, and ACR/EULAR criteria in patients with and without a final diagnosis of RA after 2 years, defined as RA diagnosed by the office-based rheumatologist and receiving disease-modifying antirheumatic drug and/or glucocorticoid therapy.

| Item                                   | Final Diagnosis of RA    |                           | p        |
|--|--------------------------|---------------------------|----------|
|  | No,<br>N = 170,<br>n (%) | Yes,<br>N = 100,<br>n (%) |          |
| 1. Morning stiffness ≥ 1 hour*         | 59 (34.7)                | 68                        | < 0.0001 |
| 2. Arthritis of 3 or more joint areas* | 74 (43.5)                | 78                        | < 0.0001 |
| 3. Arthritis of hand joints*           | 76 (44.7)                | 82                        | < 0.0001 |
| 4. Symmetric arthritis*                | 61 (35.9)                | 77                        | < 0.0001 |
| 5. Rheumatoid nodules                  | 0 (0)                    | 3                         | < 0.0001 |
| 6. Rheumatoid factor (positive)        | 24 (14.1)                | 57                        | < 0.0001 |
| Low                                    | 14 (8.2)                 | 24                        |          |
| High                                   | 10 (5.9)                 | 33                        |          |
| 7. Radiographic changes                | 10 (5.9)                 | 11                        | 0.15     |
| 8. ACPA (positive)                     | 11 (6.4)                 | 48                        | < 0.0001 |
| Low                                    | 2 (1.2)                  | 5                         |          |
| High                                   | 9 (5.3)                  | 43                        |          |
| ACR 1987 criteria ≥ 4                  | 34 (20)                  | 64                        | < 0.0001 |
| ACR 1987 Liao 2 modified criteria ≥ 3  | 58 (34.1)                | 84                        | < 0.0001 |
| ACR 1987 Liao 1 modified criteria ≥ 4  | 34 (20)                  | 69                        | < 0.0001 |
| Joint involvement, points              |                          |                           |          |
| 0                                      | 78 (45.9)                | 17                        |          |
| 1                                      | 14 (8.2)                 | 3                         |          |
| 2                                      | 42 (24.7)                | 26                        | < 0.0001 |
| 3                                      | 22 (12.9)                | 31                        |          |
| 5                                      | 14 (8.2)                 | 23                        |          |
| More than 6 weeks (1 point)            | 120 (70.6)               | 83                        | 0.028    |
| Elevated ESR or CRP (1 point)          | 121 (71.2)               | 81                        | 0.08     |
| RF or ACPA, points                     |                          |                           |          |
| 0                                      | 132 (77.6)               | 29                        |          |
| 2                                      | 13 (7.6)                 | 11                        | < 0.0001 |
| 3                                      | 17 (10)                  | 49                        |          |
| ACR/EULAR ≥ 6                          | 24 (14.1)                | 58                        | < 0.0001 |

\* More than 6 weeks. RF: rheumatoid factor; ACPA: anticyclic citrullinated peptide antibodies; ESR: erythrocyte sedimentation rate; CRP: serum C-reactive protein; ACPA: anticyclic citrullinated peptide antibodies.

*Diagnostic value of the RA classification criteria sets.* At baseline, 111 of the 270 patients had alternative diagnoses that better explained the arthritis. Sixteen patients had erosions typical for RA and 11 of them had RA at 2 years. Among the patients having erosions at inclusion but no diagnosis of RA at completion, one had algodystrophy, 3 had RA but were taking no DMARD, and 1 was considered by the rheumatologist to have undifferentiated arthritis.

Of the 143 remaining patients, 91 had scores lower than 6, including 38 (42.2%) with RA after 2 years; and 52 had scores  $\geq 6$  (indicating definite RA), including 40 (75.6%) with RA after 2 years (Figure 1). Among no RA patients having a score  $\geq 6$ , diagnoses were as follows: chondrocalcinosis (n = 1), mixed connective tissue disease (1), hydroxyapatitis (1), polymyositis (1), polymyalgia rheumatica (1), spondyloarthropathy (2), lymphoma (1), and 4 with RA but taking no DMARD.

The sensitivity, specificity, PPV, and NPV of the ACR/EULAR (full tree) were 51/100 (51%), 153/170 (90%), 51/68 (75.4%), and 153/202 (75.7%), respectively (Table 2). For all criteria, specificity was better in the subgroup meeting the ACR/EULAR scoring conditions 1–3 (use of the full tree) than in the overall population (use of the scoring criteria regardless of whether conditions were met).

The ACR/EULAR criteria performed slightly better than the unmodified or modified ACR 1987 criteria in the subgroup meeting ACR/EULAR scoring conditions (Figure 2A), whereas no significant difference was found in the

overall population (Figure 2B). However, we failed to demonstrate any statistical differences between the AUC of the ACR/EULAR scoring system versus the ACR 1987 criteria (unmodified or modified by Liao 1 and 2) in the overall population ( $p = 0.29$ ,  $p = 0.61$ , and  $p = 0.70$ , respectively) or in the subgroup meeting the 3 ACR/EULAR conditions ( $p = 0.49$ ,  $p = 0.89$ , and  $p = 0.6$ , respectively).

*Best combination of items.* Logistic regression analysis was done to identify the best combinations of items for diagnosing RA in the subgroup meeting all 3 conditions for ACR/EULAR scoring. Analysis identified the following variables: symmetric arthritis (ACR criterion 4), rheumatoid factor (positive or negative), ACPA (negative, low, high), ESR or CRP, and joints (0 to 5 points), respectively (Table 3). Diagnostic performance was best when all ACR/EULAR items were combined with the “symmetric arthritis” item (item 4) of the ACR 1987 criteria, as follows: [(joint 0–5)/5 + (ACPA 0–3) + (RF 0–1) x 2 + (ACR4 0–1) x 2 + (ESR or CRP 0–1) x 2], where joint 0–5 indicates the ACR/EULAR joint score, ACPA 0–3 (negative 0, low 2, high 3), RF 0–1 the absence or presence of RF, ACR4 0–1 the absence or presence of item 4 in the ACR 1987 criteria set, ESR 0–1 the absence or presence of ESR elevation, and CRP 0–1 the absence or presence of CRP elevation. This equation is designated “the best combination of criteria” hereafter. A score of 5 or more had the best diagnostic value. Figure 3A shows the diagnostic value of each item in the subgroup meeting all 3 conditions for ACR/EULAR scoring. Figure

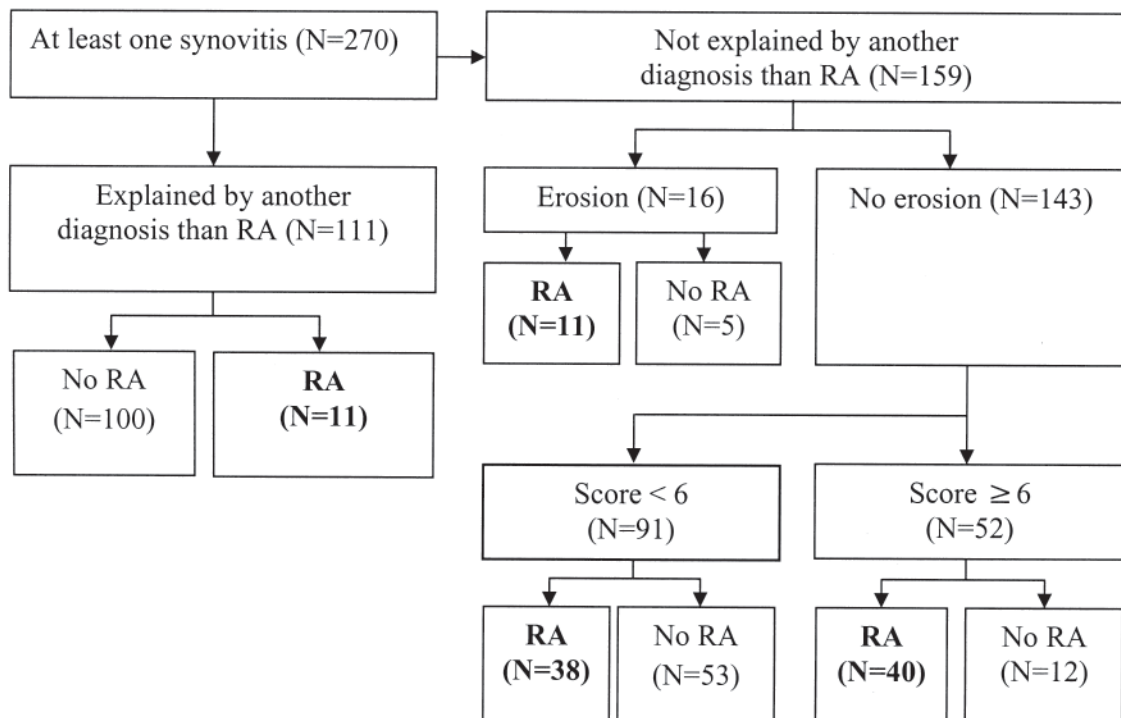
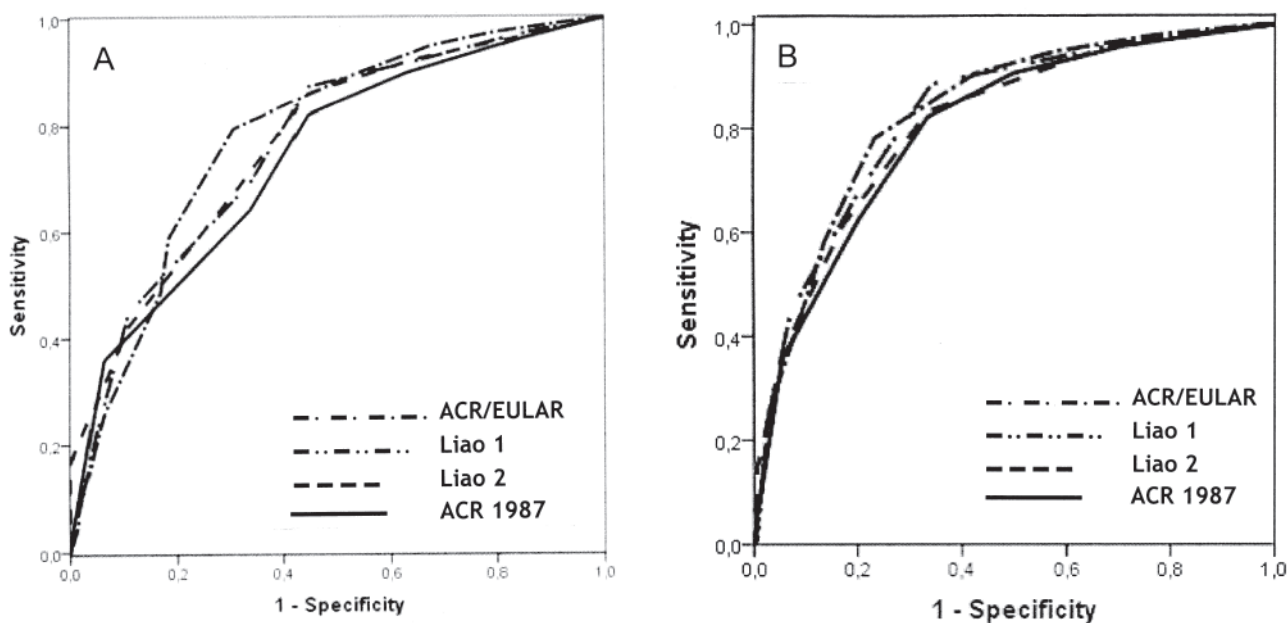


Figure 1. Final diagnosis of rheumatoid arthritis (RA) in the overall population according to the ACR/EULAR criteria.

**Table 2.** Diagnostic value of the 3 criteria sets (unmodified ACR 1987, modified ACR 1987 criteria, and more than 6 points on the ACR/EULAR score), and of the best combination of items in the Brittany cohort in the overall population, in the subgroup meeting the 3 conditions for ACR/EULAR scoring (synovitis, no better alternative diagnosis, and no radiographic erosion typical for RA), and in the overall population classified according to the full ACR/EULAR tree.

| Criteria   | Subgroups      | Sensitivity, %<br>(95% CI) | Specificity, %<br>(95% CI) | Positive Predictive Value,<br>(95% CI) | Negative Predictive Value,<br>(95% CI) | AUC*   |
|--|----------------|----------------------------|----------------------------|--|--|--------|
| ACR 1987<br>criteria, $\geq 4$                                     | Overall        | 64/100, 64 (53.7–73.2)     | 136/170, 80 (73.0–85.6)    | 64/98, 65.3 (54.9–74.4)                | 136/172, 79.1 (72.1–84.71)             | 0.81   |
|  | Conditions met | 50/78, 64.1 (52.4–74.4)    | 43/65, 66.1 (53.2–77.1)    | 50/72, 69.4 (58.4–79.5)                | 43/71, 60.6 (48.2–71.7)                | 0.74   |
|  | Full tree      | 61/100, 61 (50.7–70.4)     | 151/170, 88.8 (82.9–92.9)  | 61/88, 69.3 (58.4–78.5)                | 143/182, 78.6 (71.8–84.1)              |        |
| ACR 1987<br>criteria with Liao<br>2 modification,<br>$\geq 3$      | Overall        | 69/100, 69 (58.9–77.7)     | 136/170, 80 (73.3–85.6)    | 69/98, 70.4 (60.2–78.9)                | 136/172, 79.1 (72.1–84.7)              | 0.81   |
|  | Conditions met | 67/78, 85.9 (75.7–92.4)    | 36/65, 55.4 (42.6–67.5)    | 67/96, 69.8 (59.5–78.5)                | 36/47, 76.6 (61.6–87.2)                | 0.77   |
|  | Full tree      | 72/100, 72 (61.9–80.3)     | 144/170, 84.7 (78.2–89.6)  | 78/112, 69.6 (60.1–77.8)               | 136/158, 86.1 (79.5–90.9)              |        |
| ACR 1987<br>criteria with Liao<br>1 modification,<br>$\geq 4$      | Overall        | 84/100, 84 (75.0–90.3)     | 112/170, 65.9 (58.2–72.9)  | 84/142, 59.1 (50.6–67.2)               | 112/128, 87.5 (80.2–92.5)              | 0.83   |
|  | Conditions met | 54/78, 69.2 (57.6–78.9)    | 43/65, 66.1 (53.2–77.1)    | 54/76, 71.0 (59.3–80.6)                | 43/67, 64.2 (51.5–75.3)                | 0.77   |
|  | Full tree      | 65/100, 65 (54.7–74.1)     | 143/170, 84.1 (77.5–89.1)  | 65/92, 70.6 (60.1–79.4)                | 143/178, 80.3 (73.6–85.8)              |        |
| ACR/EULAR<br>score $\geq 6$  | Overall        | 58/100, 58 (47.7–67.7)     | 143/170, 85.9 (79.5–90.6)  | 58/82, 70.7 (59.5–79.9)                | 146/188, 77.6 (70.9–83.3)              | 0.83   |
|  | Conditions met | 40/78, 51.3 (39.8–62.7)    | 53/65, 81.5 (69.6–89.7)    | 40/52, 76.9 (62.8–87)                  | 53/91, 58.2 (47.4–68.3)                | 0.78   |
|  | Full tree      | 51/100, 51 (40.9–61.1)     | 153/170, 90 (84.2–93.9)    | 51/68, 75 (64.2–83.9)                  | 153/202, 75.7 (69.1–81.4)              |        |
| Best combination<br>of items in the<br>Brittany cohort<br>$\geq 5$ | Overall        | 62/100, 62 (51.7–71.4)     | 145/170, 85.3 (78.9–90.1)  | 62/87, 71.3 (60.4–80.2)                | 145/183, 79.2 (72.5–84.7)              | 0.84   |
|  | Conditions met | 50/78, 64.1 (52.4–74.4)    | 53/65, 81.8 (69.6–89.7)    | 50/62, 80.6 (68.2–89.2)                | 53/81, 65.4 (53.9–75.4)                | 0.83** |
|  | Full tree      | 61/100, 61 (50.7–70.4)     | 153/170, 90 (84.2–93.9)    | 61/88, 69.3 (58.4–78.5)                | 153/192, 79.7 (73.2–84.9)              |        |

\* Area under the curve of the receiver operating characteristic curves when conditions were met. \*\*  $p$ , 0.05 versus ACR/EULAR score.



**Figure 2.** Receiver operating characteristic curve of the criteria sets in the subgroup meeting the 3 conditions for ACR/EULAR scoring (synovitis, no better alternative diagnosis, and no typical erosions; panel A;  $n = 143$ ) and in the overall population (panel B;  $n = 270$ ).

3B and Table 2 show that the Brittany criteria performed slightly better than other criteria sets in this subgroup ( $p = 0.006$  vs ACR criteria;  $p = 0.02$  vs Liao 1 and  $p = 0.04$  vs Liao 2 modified criteria; and  $p = 0.03$  vs the ACR/EULAR criteria).

## DISCUSSION

We report the first data on the diagnostic accuracy of the ACR/EULAR criteria for RA in a prospective cohort of patients with recent-onset arthritis that was not among the cohorts used to develop the criteria.

Table 3. Logistic regression in the group that was not better explained by another diagnosis and without erosion typical for RA.

| Criteria   | Exp β (95% CI)   | p        |
|--|------------------|----------|
| Symmetric arthritis (ACR criteria 4)               | 2.96 (1.19–7.37) | 0.02     |
| Rheumatoid factor (positive or negative)           | 3.23 (1.27–8.22) | 0.014    |
| Cyclic citrullinated peptide (negative, low, high) | 2.90 (1.6–5.23)  | < 0.0001 |
| ESR or CRP   | 2.67 (1.01–7.05) | 0.03     |
| Joint (0 to 5 points)                              | 1.33 (1.02–1.73) | 0.032    |

ESR: erythrocyte sedimentation rate; CRP: C-reactive protein.

Application of the ACR/EULAR criteria is appropriate in the population of patients with synovitis, no radiographic evidence of RA, and no other more plausible diagnosis. We therefore evaluated the ACR/EULAR criteria in this population. The ACR 1987 criteria have been used both in similar patients (for example, by clinicians to confirm a doubtful diagnosis) and in unselected patients with early arthritis (to separate RA from other conditions). Consequently, we also evaluated the scoring component of the ACR/EULAR tree in our overall population, compared to the ACR 1987 criteria (the original version and modifications by Liao).

In our cohort, we found no statistically significant differences in diagnostic accuracy between the ACR/EULAR 2010 score and the ACR 1987 criteria (unmodified and modified by Liao, *et al*) in the overall population or in the subgroup meeting the 3 conditions for ACR/EULAR scor-

ing. However, specificity was clearly improved by the tree format, which excluded patients having a better alternative diagnosis at baseline. Of the 270 patients, 111 (41%) had a better alternative diagnosis at baseline. Of these 111 patients, 11 were finally diagnosed as having RA. The most common diagnoses in these patients were spondyloarthropathy (21%), crystal-induced arthritis (5%), and connective tissue disease (7%; including Sjögren's syndrome, systemic lupus erythematosus, scleroderma, and others). These results are consistent with those in the 936 patients of the Leiden cohort<sup>17</sup>.

Patients were included in our cohort if they had synovitis in one or more joints of less than 1 year duration. They were followed prospectively for 2 years. This is the ideal situation for testing new diagnostic criteria. RA was defined according to a practical viewpoint as RA diagnosed by the office-based rheumatologist combined with use of DMARD and/or steroid therapy.

The third condition for ACR/EULAR scoring is absence of erosions typical for RA on plain radiographs. Such erosions are rare in patients with recent-onset arthritis. Of the 159 patients with no better diagnosis than RA at baseline, 16 had typical erosions and were therefore classified by the ACR/EULAR criteria as having RA. However, only 11 of these patients were finally diagnosed as having RA, and thus 5 patients would have received a DMARD for a condition other than RA. In an earlier study of the same cohort, we found that presence of typical erosions on the baseline

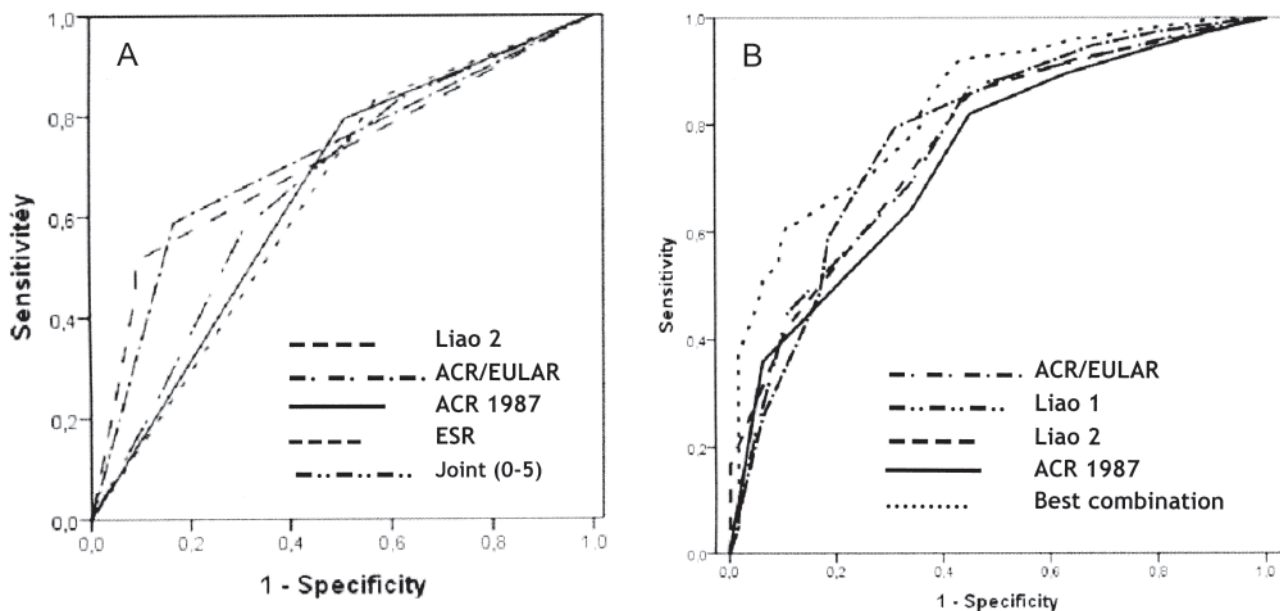


Figure 3. Receiver operating characteristic curve of each item identified by the logistic regression in the subgroup meeting the 3 conditions for ACR/EULAR scoring (synovitis, no better alternative diagnosis, and no typical erosions; panel A); and in panel B, the diagnostic value of the best combination of items in the Brittany cohort [(joint 0–5)/5 + (ACPA 0–3) + (RF 0–1) × 2 + (ACR4 0–1) × 2 + (ESR or CRP 0–1) × 2], where joint 0–5 indicates the ACR/EULAR joint score, ACPA 0–3 (negative, low, high), RF absence or presence of rheumatoid factors, ACR4 0–1 the absence or presence of item 4 in the ACR 1987 criteria set, ESR 0–1 the absence or presence of ESR elevation, and CRP 0–1 the absence or presence of CRP elevation, compared to the other criteria sets.

radiographs was 22% sensitive and 98% specific for RA<sup>6</sup>. However, radiographs of the feet were not performed and would be expected to increase the number of patients with erosions<sup>6</sup>. The value of radiographic erosions for the diagnosis of RA has been evaluated in 2 other prospective cohorts of patients with undifferentiated arthritis<sup>12,18</sup>. Sensitivity was low (16%<sup>12</sup> and 24%<sup>17</sup>) and specificity was high (93%<sup>12</sup> and 92%<sup>18</sup>).

To compare the performance of the ACR/EULAR scoring system to other criteria sets, we used the unmodified ACR 1987 criteria and the ACR 1987 criteria modified by Liao in the patient subgroups that were defined based on whether the second and third conditions for ACR/EULAR scoring were met (all patients met the first condition). In the subgroup of patients meeting both the second and third conditions, the ACR/EULAR scoring system performed slightly better than did the other criteria sets (Table 2). To determine the best combination of items in patients meeting the second condition (no better alternative diagnosis), we performed a logistic regression analysis. We found that all items in the ACR/EULAR score were associated with a final diagnosis of RA except the duration of synovitis. This item proved difficult to study in the first phase of ACR/EULAR criteria development, and the 6-week cutoff was chosen arbitrarily to encourage early referral and diagnosis. In our study, the RF titer did not provide diagnostic information in addition to that supplied by the presence or absence of rheumatoid factors. In a cohort of patients with early arthritis from The Netherlands, baseline IgM RF levels showed only a weak correlation with structural disease progression and had no association at all with disease activity after 2 years<sup>19</sup>. In our study, presence of RF and the ACPA titer were independent factors predicting a final diagnosis of RA after 2 years, suggesting that both should be assayed when RA is suspected, and separated in the scoring system. No consensus exists concerning the methods to be used for autoantibody detection or the definitions of negative, low-positive, and high-positive values. Therefore, arbitrary cutoffs were used in the ACR/EULAR criteria. Local laboratory ranges should be used until international units are developed, and a high titer should be defined as a titer > 3 times the upper limit of the normal range.

In our study, acute-phase reactants (ESR or CRP) were strong predictors of RA in the subgroup meeting the second condition for ACR/EULAR scoring although they were not associated with a final diagnosis of RA in the overall population. Similarly, in earlier studies acute-phase reactants were associated with RA in cohorts of patients with undifferentiated arthritis<sup>12,20</sup>, but not in cohorts established without excluding patients who had better alternative diagnoses<sup>21,22</sup>.

In the ACR/EULAR criteria, joint involvement is defined as tender and/or swollen joints at the time of assessment by an expert. Small joints are defined as the proximal interpha-

langeal joints, metacarpophalangeal joints, wrists excluding the first carpometacarpal joints, and metatarsophalangeal joints of toes 2 through 5. The distal interphalangeal joints are not considered. Large joints are defined as the elbows, shoulders, hips, knees, and ankles. In our regression analysis, the pattern of joint involvement predicted RA, but to a lesser degree than the other items. The ACR/EULAR system for scoring synovitis between 0 and 5 could probably be simplified for easier use in daily practice. In our study, the fourth item of the ACR 1987 classification criteria, namely symmetric joint involvement, improved the diagnostic performance when used in combination with the ACR/EULAR criteria.

Thus, the best scoring system involved slight modifications to the ACR/EULAR criteria, consisting of removal of synovitis duration, addition of symmetric joint involvement, and handling of the RF criterion as a binary variable (present/absent). The calculated weights favored items that strongly predicted RA. We developed a scoring system on a 0–9 scale, with 2 points assigned to symmetric joint involvement, presence of rheumatoid factors, low ACPA titer, and elevated ESR or CRP; 3 points assigned to high anti-CCP titer; and the joint involvement pattern between 0 and 5 divided by 5. A score greater than 4 was the optimal cutoff to discriminate RA from non-RA patients. Our goal was not to develop new criteria but to determine whether logistic regression performed in a cohort with early arthritis identified an effective scoring system using items of both the ACR 1987 criteria set and the ACR/EULAR 2010 criteria set. We identified such a system, whereas we found no significant difference in diagnostic accuracy between the ACR 1987 criteria (modified or unmodified) and the ACR/EULAR scoring after patient selection by the tree format. Thus, these 2 criteria sets may be similar, and a combination of items from both sets may slightly improve diagnostic performance.

In summary, much of the improvement supplied by using the ACR/EULAR criteria for RA is ascribable to the exclusion of patients with alternative diagnoses.

## ACKNOWLEDGMENT

We are grateful to the following rheumatologists for referring patients: E. Blat, P. Busson, A. Castagné, J.P. Caumon, P. Chicault, V. Desmas, J.P. Elie, X. Filliol, C. Gauthier, J. Glemarec, J.Y. Grolleau, M.N. Guillermit, R. Guyader, M. Hamidou, P. Herrou, G. Lavel, F. Le Jean, R. Lemaître, M.C. Lheveder, A. Martin, Y. Maugars, I. Nouy Trolle, J. Olivry, C. Paturel, N. Paugam, A. Prost, D. Rault, B. Ribeyrol, A. Rossard, D. Rodet, I. Valls, P. Vilon, and P. Voisin.

## REFERENCES

1. Ropes MW, Bennett GA, Cobb S, Jacox R, Jessar RA. 1958 Revision of diagnostic criteria for rheumatoid arthritis. *Bull Rheum Dis* 1958;9:175-6.
2. 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;31:315-24.

3. Harrison BJ, Symmons DP, Barrett EM, Silman AJ. The performance of the 1987 ARA classification criteria for rheumatoid arthritis in a population based cohort of patients with early inflammatory polyarthritis. *J Rheumatol* 1998;25:2324-30.
4. Hülsemann JL, Zeidler H. Diagnostic evaluation of classification criteria for rheumatoid arthritis and reactive arthritis in an early synovitis outpatient clinic. *Ann Rheum Dis* 1999;58:278-80.
5. Saraux A, Youinou P, Le Goff P. Classification criteria for rheumatoid arthritis. *Scand J Rheumatol* 1997;26:229.
6. Saraux A, Berthelot JM, Chalès G, Le Henaff C, Thorel JB, Hoang S, et al. Ability of the American College of Rheumatology 1987 criteria to predict rheumatoid arthritis in patients with early arthritis and classification of these patients two years later. *Arthritis Rheum* 2001;44:2485-91.
7. Devauchelle Pensec V, Saraux A, Berthelot JM, Alapetite S, Chalès G, Le Henaff C, et al. Ability of hand radiographs to predict a further diagnosis of rheumatoid arthritis in patients with early arthritis. *J Rheumatol* 2001;28:2603-7.
8. Sebbag M, Simon M, Vincent C, Masson-Bessière C, Girbal E, Durieux JJ, et al. The antiperinuclear factor and the so-called antikeratin antibodies are the same rheumatoid arthritis-specific autoantibodies. *J Clin Invest* 1995;95:2672-9.
9. Liao KP, Batra KL, Chibnik L, Schur PH, Costenbader KH. Anti-cyclic citrullinated peptide revised criteria for the classification of rheumatoid arthritis. *Ann Rheum Dis* 2008;67:1557-61.
10. Berthelot JM, Klarlund M, McGonagle D, Bernelot-Moens HJ, Calin A, Harrison B, et al. Lessons from an international survey of paper cases of 10 real patients from an early arthritis clinic. CRI (Club Rhumatismes et Inflammation) Group. *J Rheumatol* 2001;28:975-81.
11. Visser H, le Cessie S, Vos K, Breedveld FC, Hazes JMW. How to diagnose rheumatoid arthritis early: a prediction model for persistent (erosive) arthritis. *Arthritis Rheum* 2002;46:357-65.
12. 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;56:433-40.
13. Aletaha D, Neogi T, Silman A, Funovits J, Felson D, Clifton O, et al. The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for RA. *Arthritis Rheum* 2010;62:2569-81.
14. Funovits J, Aletaha D, Bykerk VP, Combe B, Dougados M, Emery P, et al. The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: methodological report phase I. *Ann Rheum Dis*; 69:1589-95.
15. Devauchelle-Pensec V, Josseume T, Samjee I, Dougados M, Combe B, Saraux A. Ability of oblique foot radiographs to detect erosions in early arthritis: results in the ESPOIR cohort. *Arthritis Rheum* 2008;59:1729-34.
16. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978;8:283-98.
17. van Gaalen FA, Linn-Rasker SP, van Venrooij WJ, de Jong BA, Breedveld FC, Verweij CL, et al. Autoantibodies to cyclic citrullinated peptides predict progression to rheumatoid arthritis in patients with undifferentiated arthritis: a prospective cohort study. *Arthritis Rheum* 2004;50:709-15.
18. Kuriya B, Cheng CK, Chen HM, Bykerk VP. Validation of a prediction rule for development of rheumatoid arthritis in patients with early undifferentiated arthritis. *Ann Rheum Dis* 2009; 68:1482-5.
19. Ursul J, Bos WH, van Dillen N, Dijkmans BA, van Schaardenburg D. Levels of anti-citrullinated protein antibodies and IgM rheumatoid factor are not associated with outcome in early arthritis patients: a cohort study. *Arthritis Res Ther* 2010;12:R8.
20. Tamai M, Kawakami A, Uetani M, Takao S, Arima K, Iwamoto N, et al. A prediction rule for disease outcome in patients with undifferentiated arthritis using magnetic resonance imaging of the wrists and finger joints and serologic autoantibodies. *Arthritis Rheum* 2009;61:772-8.
21. Mjaavatten MD, Uhlig T, Haugen AJ, Nygaard H, Sidenvall G, Helgetveit K, et al. Positive anti-citrullinated protein antibody status and small joint arthritis are consistent predictors of chronic disease in patients with very early arthritis: results from the NOR-VEAC cohort. *Arthritis Res Ther* 2009;11:R146.
22. Machold KP, Stamm TA, Eberl GJM, Nell VKP, Dunky A, Uffmann M, et al. Very recent onset arthritis — clinical, laboratory, and radiological findings during the first year of disease. *J Rheumatol* 2002;29:2278-87.