ABSTRACT. **Objective.** We assessed levels of agreement between a diagnosis of rheumatoid arthritis (RA) at inclusion in a recent-onset arthritis cohort, then 2 and 10 years later. Performance of American College of Rheumatology (ACR) criteria alone or combined with rheumatologist diagnosis, and of recent new criteria adding antibodies to cyclic citrullinated peptides (“anti-CCP-revised criteria”) to existing ACR criteria, was evaluated.

**Methods.** In total, 270 patients with recent-onset arthritis of less than 1 year duration were included between 1995 and 1997 and followed for 2 years. A diagnosis was recorded by an office-based rheumatologist (OBR) at inclusion, then 2 years later. In 2007, a questionnaire was sent to each rheumatologist to collect the final diagnosis, which was considered the reference.

**Results.** Final diagnosis was available for 164 patients: 57 had RA. Agreement was low (κ = 0.27) between the baseline and final diagnoses, and substantial (κ = 0.69) between the 2-year and final diagnoses. Anti-CCP-revised criteria had sensitivity of 65% to 81% and specificity of 55% to 75%. Sensitivity and specificity of ACR criteria were 57.9% (44.1%–70.9%) and 74.8% (65.5%–82.7%) at inclusion, 80.7% (70.5%–90.0%) and 63.6% (54.5%–72.7%) at 2 years. The combination OBR diagnosis/ACR criteria after 2 years showed considerably increased specificity (87% vs 64%) and slightly decreased sensitivity (77% vs 81%).

**Conclusion.** ACR criteria for RA showed poor performance even at 2 years. The absence of exclusion criteria may explain the lack of specificity, which improved when combined with the OBR diagnosis. Adding anti-CCP criteria to the existing criteria could help in diagnosing RA.

(J Rheumatol First Release Oct 15 2009; doi:10.3899/jrheum.090072)

Key Indexing Terms:

EARLY ARTHRITIS RHEUMATOID ARTHRITIS DIAGNOSIS COHORT
tematic review of the diagnostic value of the ACR criteria for RA showed pooled sensitivity from 77% to 80% and pooled specificity from 33% to 76% in early arthritis. Nevertheless, these criteria are widely used as the reference standard for diagnosing RA in clinical practice.

Recently, consideration of antibodies to cyclic citrullinated peptides (anti-CCP) in new revised classification criteria for RA (anti-CCP-revised criteria) was suggested by some authors. Some combinations of criteria including anti-CCP were tested by Liao, et al. We investigated the extent to which a clinical diagnosis of RA made by office-based rheumatologists (OBR) within 1 year of symptom onset, then 2 years later, and/or fulfillment of ACR criteria at inclusion and at the 2-year timepoint agreed with a diagnosis of RA reported at least 10 years after the first OBR diagnosis. Recent anti-CCP-revised criteria proposed by Liao, et al were evaluated at time of inclusion.

MATERIALS AND METHODS

Study population. In total, 270 patients were recruited prospectively between 1995 and 1997 in 7 hospitals in Brittany, France. They were referred by general practitioners (GP) or rheumatologists. Inclusion criteria were age ≥ 16 years, swelling of at least one joint, absence of previously diagnosed joint disease, and symptom duration < 1 year. The study was approved by the appropriate ethics committee, and written informed consent was obtained from each patient before inclusion.

Study design. Patients included in the cohort were followed for 2 years in the centers. Their treating rheumatologists were contacted 10 years later to determine their outcomes (Figure 1).

Baseline assessment. Assessment included a standardized interview, a general physical examination, and a standardized rheumatologic examination. Over 100 measures were evaluated, including the medical history; family history of arthritis (RA, spondyloarthropathy); joint pain on motion, tenderness, and soft tissue swelling; and ACR criteria for RA. The following investigations were performed: blood cell counts; creatinine level and proteinuria; C-reactive protein level; latex test and ELISA for IgM, IgG, and IgA rheumatoid factors (RF); tests for antiperoxidase factor, antikeratin antibody, and antinuclear antibody; HLA-DR phenotype determination; retrospectively performed anti-CCP measurements; and radiographs of the chest, hands, feet, and pelvis. Joint aspiration was performed when it was required.

Patients included in the cohort were followed for 2 years in the centers. Their treating rheumatologists were contacted 10 years later to determine their outcomes (Figure 1). Baseline assessment. Assessment included a standardized interview, a general physical examination, and a standardized rheumatologic examination. Over 100 measures were evaluated, including the medical history; family history of arthritis (RA, spondyloarthropathy); joint pain on motion, tenderness, and soft tissue swelling; and ACR criteria for RA. The following investigations were performed: blood cell counts; creatinine level and proteinuria; C-reactive protein level; latex test and ELISA for IgM, IgG, and IgA rheumatoid factors (RF); tests for antiperoxidase factor, antikeratin antibody, and antinuclear antibody; HLA-DR phenotype determination; retrospectively performed anti-CCP measurements; and radiographs of the chest, hands, feet, and pelvis. Joint aspiration was performed when it was required.

Performances of the combination ACR criteria/OBR diagnosis at inclusion and after 2 years were computed. The ACR-OBR combination was considered to indicate a RA diagnosis when the OBR-diagnosed RA and the patient’s diagnosis. At the 2-year visit, the OBR assigned the patient to the RA or non-RA group. In January 2007, 10 to 12 years after study start, a questionnaire was sent to each patient’s rheumatologist (or, when this was not possible, to the GP, who was asked to report the diagnosis of the last rheumatologist visited) to collect the final diagnosis and its degree of certainty (from 0 = null, to 10 = absolute), which served as the reference standard.

Statistical analysis. Analyses were performed using SAS 9.1 software (SAS Institute, Cary, NC, USA). The diagnosis reported by the treating rheumatologist in 2007 was considered the reference standard. Agreement between the final diagnosis and the diagnoses at baseline and after 2 years was assessed using Cohen’s kappa coefficient.

Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) with 95% confidence intervals were computed for the OBR diagnosis at baseline and after 2 years.

Performances of the ACR criteria were calculated at baseline and after 2 years. Patients were required to satisfy cumulatively at least 4 of the criteria set to be considered as having RA. Three anti-CCP-revised criteria sets were tested at baseline: (1) adding anti-CCP to the ACR criteria (≥ 4 criteria required/8); (2) replacement of rheumatoid nodules as criteria with anti-CCP (≥ 3 criteria required/7); (3) replacement of rheumatoid nodules and erosions as criteria with anti-CCP (≥ 3 criteria/6 required).

Performances of the combination ACR criteria/OBR diagnosis at inclusion and after 2 years were computed. The ACR-OBR combination was considered to indicate a RA diagnosis when the OBR-diagnosed RA and the patient met at least 4 ACR criteria. Patient groups were compared using the chi-square test or the Mann-Whitney test, as appropriate. P values < 0.05 were considered significant.

RESULTS

In 2007, information on the diagnosis was available for 164 (61%) of the patients initially included. Their baseline characteristics are reported in Table 1. Among them, 108 (65.9%) were regularly still followed by a rheumatologist on
the date of the last report. Final diagnoses are reported in Table 2; 57 (34.8%) patients were diagnosed as having RA, with a mean diagnostic certainty of 9.4/10. Most of them (84.6%) were treated with DMARD, in particular with methotrexate (57.9%). In the no-RA group, 22 subjects (20.6%) were treated with DMARD, mainly sulfasalazine and hydroxychloroquine. Four patients, with psoriatic arthritis or undifferentiated arthritis (UA), were treated with methotrexate. Forty subjects of the RA group were reported as having joint erosions, whereas there were only 6 in the no-RA group.

At baseline, the 164 patients still followed in 2007 showed no significant differences compared to the other 106 patients regarding age at inclusion, sex, swollen joint count, or laboratory data.

Agreement between the rheumatologist’s diagnosis and the final diagnosis. At baseline, the OBR classified 104 of the 164 patients as possibly having RA. Compared to the 60 other patients (the non-RA group), the RA group had significantly higher prevalence of a positive test for IgM RF (36.5% vs 14.0%, respectively; \( p = 0.003 \)), arthritis in more than one joint (94.2% vs 71.7%; \( p < 0.001 \)), symmetric arthritis (59.8% vs 43.6%; \( p = 0.02 \)), wrist involvement (42.3% vs 18.6%; \( p = 0.002 \)), and hand involvement [metacarpophalangeal (41.4% vs 11.7%) and proximal interphalangeal (41.4% vs 15.3%) joints; \( p < 0.001 \)].

For RA diagnosis, agreement between the baseline OBR diagnosis and the final diagnosis was low. Of the 164 patients, 53 were assigned to the RA group and 56 to the non-RA group at both timepoints. Only 4 patients with a final diagnosis of RA were not diagnosed with RA at baseline. Fifty-one patients were assigned to the RA group at baseline but in the non-RA group at 10 years (\( \kappa = 0.27 \) for RA vs non-RA). Changes in the diagnosis (RA vs non-RA) are reported in Figure 2. A baseline OBR diagnosis of RA had high sensitivity but low specificity for a diagnosis of RA 10 years later. The PPV was therefore low and the NPV was high (Table 3).

After 2 years, OBR gave a diagnosis of RA for 64 patients. The final diagnosis was RA in 48 of these patients and another condition in 13 patients (osteoarthritis, \( n = 2 \); Sjögren’s syndrome, \( n = 1 \); UA, \( n = 5 \); psoriatic arthritis, \( n = 1 \); polymyalgia rheumatica, \( n = 2 \); and recovery with no diagnosis, \( n = 2 \)). Of the 100 patients assigned to the non-RA group at the 2-year OBR visit, 9 had a final diagnosis of RA. Among these 9 patients, 7 were diagnosed with UA at the 2-year visit, including 5 in whom RA was consid-
Figure 2. Changes in the diagnosis of RA or non-RA. * Cured without data on final diagnosis.

Table 3. Sensitivities and specificities of office-based rheumatologist (OBR) diagnosis and ACR criteria at inclusion and at 2 years.

<table>
<thead>
<tr>
<th>At inclusion</th>
<th>Number Among RA+ (n = 57)</th>
<th>Sensitivity %</th>
<th>95% CI</th>
<th>Number Among RA– (n = 107)</th>
<th>Specificity %</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBR diagnosis*</td>
<td>53</td>
<td>93.0</td>
<td>86.4–99.6</td>
<td>56</td>
<td>52.3</td>
<td>42.5–62.1</td>
</tr>
<tr>
<td>ACR criteria</td>
<td>33</td>
<td>57.9</td>
<td>44.1–70.9</td>
<td>80</td>
<td>74.8</td>
<td>65.5–82.7</td>
</tr>
<tr>
<td>Anti-CCP-revised criteria 1**</td>
<td>37</td>
<td>64.9</td>
<td>51.1–77.1</td>
<td>80</td>
<td>74.8</td>
<td>65.5–82.7</td>
</tr>
<tr>
<td>Anti-CCP-revised criteria 2†</td>
<td>48</td>
<td>84.2</td>
<td>72.1–92.5</td>
<td>56</td>
<td>52.3</td>
<td>42.5–62.1</td>
</tr>
<tr>
<td>Anti-CCP-revised criteria 3††</td>
<td>46</td>
<td>80.7</td>
<td>68.1–90.0</td>
<td>59</td>
<td>55.1</td>
<td>45.2–64.8</td>
</tr>
<tr>
<td>OBR diagnosis/ACR criteria</td>
<td>32</td>
<td>56.1</td>
<td>42.4–69.3</td>
<td>86</td>
<td>80.4</td>
<td>71.6–87.4</td>
</tr>
<tr>
<td>OBR diagnosis/Anti-CCP-revised criteria 1</td>
<td>36</td>
<td>63.2</td>
<td>49.3–75.6</td>
<td>86</td>
<td>80.4</td>
<td>71.6–87.4</td>
</tr>
<tr>
<td>OBR diagnosis/Anti-CCP-revised criteria 2</td>
<td>46</td>
<td>80.7</td>
<td>68.1–90.0</td>
<td>77</td>
<td>72.0</td>
<td>62.5–80.2</td>
</tr>
<tr>
<td>OBR diagnosis/Anti-CCP-revised criteria 3</td>
<td>44</td>
<td>77.2</td>
<td>64.2–87.3</td>
<td>79</td>
<td>73.8</td>
<td>64.5–81.9</td>
</tr>
<tr>
<td>At 2 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OBR diagnosis*</td>
<td>48</td>
<td>84.2</td>
<td>74.7–93.7</td>
<td>91</td>
<td>85.0</td>
<td>78.3–91.8</td>
</tr>
<tr>
<td>ACR criteria</td>
<td>46</td>
<td>80.7</td>
<td>70.5–90.0</td>
<td>68</td>
<td>63.6</td>
<td>54.5–72.7</td>
</tr>
<tr>
<td>Combination OBR diagnosis/ACR criteria</td>
<td>44</td>
<td>77.2</td>
<td>66.3–88.1</td>
<td>93</td>
<td>86.9</td>
<td>80.5–93.3</td>
</tr>
</tbody>
</table>

* For OBR diagnosis, positive and negative predictive values were estimated. At inclusion, PPV was 51.0% (95% CI 41.4%–60.6%) and NPV was 93.3% (95% CI 87.0%–99.6%). At 2 years, PPV was 75.0% (95% CI 64.4%–85.6%) and NPV was 91.0% (95% CI 85.4%–96.6%). ** ACR criteria + anti-CCP; † ACR criteria – nodules + anti-CCP; †† ACR criteria – nodules – erosions + anti-CCP.
cluded possible. Cohen’s kappa coefficient was 0.69 for agreement between the 2-year OBR diagnosis and the final diagnosis of RA or absence of RA.

Compared to a baseline OBR diagnosis of RA, a 2-year OBR diagnosis of RA was slightly less sensitive but considerably more specific for a final diagnosis of RA (Table 3).

Performance of the ACR criteria for RA (Table 3). At the inclusion visit, 60 subjects fulfilled at least 4 ACR criteria for RA; 33 of them had a final diagnosis of RA. The ACR criteria had a lower sensitivity and a higher specificity than the initial OBR diagnosis for the diagnosis of RA.

At the 2-year visit, 85 patients fulfilled at least 4 ACR criteria for RA; only 46 had a final diagnosis of RA. On the other hand, 11 patients who did not fulfill the ACR criteria at the 2-year visit had a final diagnosis of RA. Both sensitivity and specificity of the ACR criteria were lower than the OBR diagnosis at the 2-year visit for the diagnosis of RA.

Performances of the combination OBR diagnosis/ACR criteria and of the combination OBR diagnosis/anti-CCP-revised criteria at inclusion (Table 3). At inclusion, 41 subjects (25%) were anti-CCP-positive, among whom 28 were diagnosed as having RA in 2007. Adding anti-CCP to the ACR criteria increased the sensitivity from 57.9% to 64.9%, with no change in the specificity. Removing rheumatoid nodules and then rheumatoid nodules and erosions and adding anti-CCP increased sensitivity (84% and 80.7% vs 57.9%) but decreased specificity (52.3% and 55.1% vs 74%).

Performances of the combination OBR diagnosis/ACR criteria after 2 years. Of the 85 patients who met ACR criteria for RA after 2 years, only 58 (68.2%) were given a diagnosis of RA by the OBR at the same visit. The OBR diagnoses in the other 27 patients are reported in Table 4. Only 2 of these 27 patients had a final diagnosis of RA; both had a 2-year OBR diagnosis of UA. Of the 6 patients given a diagnosis of RA at the 2-year OBR visit despite failure to meet ACR criteria, 4 had a final diagnosis of RA, 1 a diagnosis of either RA or Sjögren’s syndrome, and 1 a diagnosis of polymyalgia rheumatica.

Compared to the ACR criteria alone, combining a 2-year OBR diagnosis of RA with presence of ACR criteria was slightly less sensitive but considerably more specific (Table 3).

DISCUSSION

Consecutive patients with swelling in at least one joint were included in our cohort, irrespective of symptom severity. Consequently, we obtained a heterogeneous population in terms of clinical presentation and expected outcomes, reflecting the patient population seen by rheumatologists in everyday practice. Patients with monoarticular disease at inclusion had previously been studied, showing their favorable outcome.

In our study, meeting the ACR criteria showed limited effectiveness in early arthritis for identifying patients who would have a 10-year diagnosis of RA. This result is consistent with the recent systematic review. As suggested recently, in our cohort we tested revised classification criteria including anti-CCP (anti-CCP-revised criteria). Our results are comparable to those of Liao, et al. Adding anti-CCP improved the sensitivity without modifying specificity. Removing rheumatoid nodules alone or rheumatoid nodules and erosions and adding anti-CCP increased sensitivity but decreased the specificity, probably because rheumatoid nodules and erosions are rare in new-onset RA but are highly specific for the diagnosis.

In this study, an OBR baseline diagnosis of RA had low PPV but high NPV. RA was suspected in 64% of patients at the baseline visit. This diagnosis was associated with factors known to predict persistent and erosive arthritis. Only 51% of these patients had a final diagnosis of RA. This considerable difference between the baseline and final diagnoses has 2 implications: (1) it emphasizes the difficulty of diagnosing RA, or predicting progression to RA in patients who present with early arthritis; and (2) it illustrates the key goal of the OBR at the baseline visit: to identify patients who might have RA, even when EULAR recommendations for the management of early arthritis were not available at the beginning of this study. In contrast, only 4 patients in whom RA was not considered at baseline had a final diagnosis of RA. Thus, the likelihood of having RA after 10 years is probably very low in patients for whom this diagnosis is not considered initially by the rheumatologist.

Agreement was substantial between the 2-year OBR diagnosis and the final diagnosis. However, the clinical and radiological course varies widely across patients with RA.

### Table 4. Diagnoses by office-based rheumatologists at the 2-year visit in 27 patients who met ACR criteria for RA but were not given a diagnosis of RA.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>No. of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spondyloarthropathy</td>
<td>10</td>
</tr>
<tr>
<td>Undifferentiated arthritis</td>
<td>7</td>
</tr>
<tr>
<td>Systemic lupus erythematosus</td>
<td>2</td>
</tr>
<tr>
<td>Polymyositis</td>
<td>2</td>
</tr>
<tr>
<td>Calcium pyrophosphate deposition disease</td>
<td>2</td>
</tr>
<tr>
<td>Hydroxyapatite crystal-induced arthritis</td>
<td>1</td>
</tr>
<tr>
<td>Sarcoidosis</td>
<td>1</td>
</tr>
<tr>
<td>Polymyalgia rheumatica</td>
<td>1</td>
</tr>
<tr>
<td>Sjögren’s syndrome</td>
<td>1</td>
</tr>
</tbody>
</table>

Morvan, et al: Diagnosis of RA
Patients at one end of the spectrum have low disease activity and mild symptoms, whereas others experience severe symptoms and rapid joint destruction. The diagnosis of RA is extremely difficult to establish in the milder forms, and the rheumatologist may hesitate for several years between RA and UA, which may require DMARD regardless of the diagnosis. Moreover, several diseases may simulate RA during the first few years, such as psoriatic arthritis or polymyalgia rheumatica.

At 2 years, the performance of the ACR criteria remained limited. In particular, specificity was low, but increased substantially when a 2-year OBR diagnosis of RA was added, in accord with the fact that the OBR routinely looked for evidence of other diagnoses such as crystal deposition disease, connective tissue disease (e.g. systemic lupus erythematosus), or spondyloarthropathy. In a case-control study, specificity improved from 73% with the ACR criteria alone to 89% after adding routine evaluation for other diagnoses characterized by distinctive manifestations, such as psoriasis or presence of monosodium urate crystals. Thus, adding exclusion criteria would probably improve the performance of the ACR criteria for RA.

Our study has some limitations. First, patients included in the cohort were free of standardized evaluations after 2 years. No standardized evaluation was therefore performed to establish the final diagnosis. The reference for the final diagnosis was the report of the rheumatologist at 10 years, based on the course of the symptoms, laboratory tests, and radiological findings. We chose this reference strategy because we hypothesized that a long-term followup allowed a relatively reliable diagnosis. Moreover, it was the result of everyday practice, but we cannot exclude that results would be slightly different if the reference diagnosis had been based on a standardized evaluation performed by an OBR. Second, because the reference evaluation was partly based on OBR diagnosis and partly on ACR criteria, it could lead to an overestimation of the diagnostic values of these tests. Third, because of the study design, 39% of the subjects initially included were lost to followup. At inclusion, these patients showed no significant differences compared to the 164 patients followed for 10 years. However, their diagnoses at 2 years and 10 years could have differed. We can in particular suppose that patients free of symptoms were followed less closely than symptomatic patients. A sensitivity analysis (data not shown) showed that, considering all patients lost to followup, specificity was slightly modified (from 2% to 4%), whereas the sensitivity remained the same. When considering all patients lost to followup as having RA, the sensitivity was greatly decreased whereas specificity remained the same. When considering all patients lost to followup as not having RA, specificity was slightly modified (from 2% to 4%), whereas the sensitivity remained the same. Thus our results must be confirmed by other research having a more regular followup to limit the missing subjects. Fourth, anti-CCP measurement was performed retrospectively and was not yet available in 1995 when the study started.

Anti-CCP measures are highly specific and 67% sensitive for RA according to a recent metaanalysis. The anti-keratin antibody titer, with a diagnostic value close to that of anti-CCP, was available, however. The baseline OBR diagnosis would perhaps be more accurate if anti-CCP titers were taken into account. Anti-CCP measurements were not performed at 2 years. The new anti-CCP-revised criteria were therefore not evaluated at this time.

Our study confirms that the heterogeneity of early RA raises major diagnostic challenges. The initial OBR diagnosis had high NPV but low PPV for a final diagnosis of RA; thus RA was often suspected initially in patients who finally turned out to have another cause of their symptoms. The term “rheumatoid arthritis” is probably inappropriate during the first few months after onset of symptoms. On the other hand, DMARD therapy is usually warranted in patients with UA (i.e., with possible RA).

Even after 2 years of followup, the diagnosis of arthritis is difficult to establish. The diagnosis should be reappraised frequently, to ensure that the treatment is adjusted to changes over time in the nosology and severity of arthritis.

ACR criteria for RA in early arthritis and after 2 years of followup have limited diagnostic value. The absence of exclusion criteria may explain this poor performance, as suggested by the improvement in specificity that was obtained by combining the OBR diagnosis with the ACR criteria. Adding anti-CCP measures and exclusion criteria (such as psoriasis, presence of monosodium urate crystals in joint effusion, or calcium pyrophosphate deposition in radiographs) to the existing criteria for RA could be helpful for classifying patients with early arthritis into RA and non-RA groups or, probably more important than the diagnosis, for identifying patients at risk of developing a persistent arthritis or erosion.

REFERENCES


