Prognostic Factors for Joint Destruction in Rheumatoid Arthritis: A Prospective Longitudinal Study of 318 Patients

FRANCIS GUILLEMIN, NATHALIE GÉRARD, MIEK van LEEUWEN, LIV MARIT SMEDSTAD, TORE KRISTIAN KVIE, WIM van den HEUVEL, and the EURIDISS Group

ABSTRACT. Objective. To quantify articular damage and to investigate prognostic factors for joint damage progression in rheumatoid arthritis (RA).

Methods. RA patients satisfying the 1987 American College of Rheumatology criteria and with disease duration under 5 years were sampled from the EURIDISS longitudinal cohort study in Norway, The Netherlands, and France. Hand radiographs were assessed at baseline and at 2 to 3 years followup using Sharp score modified by van der Heijde. Assessment of erosion and joint space narrowing, performed in sequential order by a single reader blinded to patients' characteristics, had high intraobserver reproducibility (intraclass correlation coefficient 0.98–0.99). Baseline prognostic factors were analyzed in a multivariate linear regression model.

Results. A total of 318 patients with RA aged 52.4 years (70.4% were female) and with a mean 2 years' disease duration at baseline were followed over 30 months. Median (quartiles) baseline and followup modified Sharp scores were 3 (0–11) and 9 (1–27), respectively, with 35.8% and then 22.3% of patients with no radiological damage. Controlling for age, sex, and country, the final joint damage was predicted by baseline modified Sharp score, rheumatoid factor positivity, time from disease diagnosis, patient global health assessment, and erythrocyte sedimentation rate, and by followup duration, explaining 76.8% of the outcome variance.

Conclusion. This multinational study confirmed the prognostic role in RA of a set of features previously identified in smaller cohorts. It indicates which disease characteristics should be focused on in the early years of RA to identify patients at higher risk of developing severe disease and who are candidates for aggressive therapy. (J Rheumatol 2003;30:2585–9)
MATERIALS AND METHODS

Patients: Patients were sampled from defined regions in 3 European countries as part of the EURIDISS (European Research on Incapacitating Diseases and Social Support) study. The EURIDISS investigation is a multicenter, longitudinal cohort study focusing on the relationships between clinical, biological, therapeutic, and social characteristics and the development of impairment in RA. The sampling procedure and sample characteristics are described in detail. Briefly, the inclusion criteria were as follows: residency in the study area, age 20–70 years, diagnosis of RA according the 1987 American College of Rheumatology criteria, and disease duration of ≤ 4 years at entry into the cohort study started in 1991. Exclusion criteria were presence of another severe incapacitating disease, stage IV in the Steinbrocker functional classification, or possibility of potential loss to followup by moving residence from the study area.

A total of 516 patients in France (n = 51), Norway (n = 205), and The Netherlands (n=260) included in the EURIDISS study had hand radiographs at the time of inclusion. Patients underwent annual standard clinical examination. Hand radiographs were obtained at the 2-year followup visit in Norway and at 3-year followup in France and The Netherlands in a total of 318 patients. Among the 516 patients, those without a radiograph at followup (n = 198) did not show any statistically significant difference from the 318 patients regarding age, sex, or clinical and baseline radiographic characteristics, except for a lower proportion of extraarticular manifestations and slightly worse overall evaluation of health (Table 1). Among the 318 patients there were no baseline differences by followup duration or country.

Assessment measures. At onset and at the yearly followup visits, clinical, biological, and therapeutic data were recorded by a rheumatologist or a rheumatology research nurse. These data included age, sex, disease duration since diagnosis and since first symptoms, Ritchie index, presence of subcutaneous nodules, and any other extraarticular manifestations (muco-cutaneous, hematological, neurological, renal, other), erythrocyte sedimentation rate (ESR), rheumatoid factor (RF; IgM-RF was measured at one center for all samples with an ELISA technique and World Health Organization international standard reference RF preparations), Steinbrocker functional class, Health Assessment Questionnaire (HAQ), patient’s overall estimation of health on 100 mm visual analog scale, and investigator’s overall estimation of health by the Karnofsky performance status index.

Outcome measures. All patients had posteroanterior radiographs for both hands on the same film at onset and at followup. Patients were excluded when radiographs were not technically correct due, for example, to inaccurate joint positioning or incorrect film exposure.

The Sharp method modified by van der Heijde (modified Sharp) was used to score hand radiographs. This scoring method is sensitive to detect radiologic changes in RA and, our collaborative group has a member with broad experience of its use. The validity of its use for hand radiographs has only recently been documented; it was found to have high intra/interobserver reliability. Erosion was scored for articular areas in both hands: 5 metacarpophalangeal (MCP) joints, 5 proximal interphalangeal (PIP) joints, first metacarpal base, multangular, lunate, ulna, radius, and navicular. Erosions were given a score of 1 if there was a discrete interruption of the cortical surface. If there was a greater defect, a score was assigned according to the amount of surface area involved. The maximum erosion score in a joint of the hand was 5.

Joint space narrowing (JSN) was scored for 15 articular joints in both hands (5 MCP joints; the 2nd, 3rd, 4th and 5th PIP joints, first metacarpal base, multangular, lunate, ulna, radius, and navicular. JSN were given a score of 0 [none], 1 [light], 2 [moderate], and 3 [severe] according the 1987 American College of Rheumatology criteria, and disease duration or country.

Table 1. Patients’ characteristics at baseline and followup (n = 318).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Netherlands, Baseline (n = 132)</th>
<th>Norway, Baseline (n = 170)</th>
<th>France, Baseline (n = 16)</th>
<th>All Countries, Baseline (n = 318)</th>
<th>Baseline Characteristics of Patients Lost to Followup, n = 198</th>
<th>All Countries at Followup, n = 318</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yrs</td>
<td>53.4 (12.2)</td>
<td>51.2 (12.6)</td>
<td>57.8 (10.3)</td>
<td>52.4 (12.4)</td>
<td>53.8 (12.4)</td>
<td></td>
</tr>
<tr>
<td>Female, %</td>
<td>65.9</td>
<td>74.7</td>
<td>62.5</td>
<td>70.4</td>
<td>67.1</td>
<td></td>
</tr>
<tr>
<td>Time from first symptoms, yrs</td>
<td>Mean (SD)</td>
<td>4.6 (4.7)</td>
<td>2.8 (1.5)</td>
<td>5.0 (2.6)</td>
<td>3.6 (3.4)</td>
<td>4.2 (4.1)</td>
</tr>
<tr>
<td>Median [quartiles]</td>
<td>3.3 [1.9–4.8]</td>
<td>2.7 [1.6–3.9]</td>
<td>4.5 [3.4–5.0]</td>
<td>3.0 [1.8–4.3]</td>
<td>3.3 [1.9–4.7]</td>
<td></td>
</tr>
<tr>
<td>Time from disease diagnosis, yrs</td>
<td>Mean (SD)</td>
<td>2.1 (1.2)</td>
<td>1.8 (1.2)</td>
<td>3.6 (1.1)</td>
<td>2.0 (1.2)</td>
<td>2.1 (1.3)</td>
</tr>
<tr>
<td>Median [quartiles]</td>
<td>2.1 [1.0–3.2]</td>
<td>1.6 [0.8–2.8]</td>
<td>4.0 [2.5–4.4]</td>
<td>2.0 [0.9–3.1]</td>
<td>2.2 [1.0–3.1]</td>
<td></td>
</tr>
<tr>
<td>Subcutaneous nodes, %</td>
<td>13.6</td>
<td>12.9</td>
<td>18.8</td>
<td>13.5</td>
<td>14.3</td>
<td>24.2</td>
</tr>
<tr>
<td>Extraarticular manifestations, %</td>
<td>12.1</td>
<td>19.4</td>
<td>37.5</td>
<td>17.3</td>
<td>8.6*</td>
<td>27.8</td>
</tr>
<tr>
<td>Steinbrocker classification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I, %</td>
<td>8.4</td>
<td>8.3</td>
<td>6.3</td>
<td>8.2</td>
<td>10.1</td>
<td>8.8</td>
</tr>
<tr>
<td>II, %</td>
<td>77.1</td>
<td>75.7</td>
<td>68.7</td>
<td>75.9</td>
<td>69.0</td>
<td>76.5</td>
</tr>
<tr>
<td>III, %</td>
<td>14.5</td>
<td>16</td>
<td>25</td>
<td>15.8</td>
<td>20.8</td>
<td>14.7</td>
</tr>
<tr>
<td>HAQ, 0–3</td>
<td>Mean (SD)</td>
<td>1.10 (0.78)</td>
<td>0.88 (0.61)</td>
<td>1.28 (0.58)</td>
<td>0.99 (0.69)</td>
<td>1.09 (0.77)</td>
</tr>
<tr>
<td>Median [quartiles]</td>
<td>1.12 [0.50–1.62]</td>
<td>0.87 [0.37–1.25]</td>
<td>1.37 [0.87–1.62]</td>
<td>1.00 [0.37–1.50]</td>
<td>1.00 [0.50–1.62]</td>
<td>1.00 [0.37–1.50]</td>
</tr>
<tr>
<td>Karnofsky index, 0–100</td>
<td>74.2</td>
<td>77.2</td>
<td>74.0</td>
<td>75.8</td>
<td>74.3</td>
<td>77.3</td>
</tr>
<tr>
<td>Ritchie index, 0–72</td>
<td>11.2</td>
<td>9.2</td>
<td>13.8</td>
<td>10.3</td>
<td>11.0</td>
<td>9.3</td>
</tr>
<tr>
<td>OEH, 0–100</td>
<td>60.6</td>
<td>21.3</td>
<td>43.3</td>
<td>50.8</td>
<td>55.8*</td>
<td>57.2</td>
</tr>
<tr>
<td>ESR, mm/h</td>
<td>27.9</td>
<td>25.1</td>
<td>19.1</td>
<td>25.9</td>
<td>29.5</td>
<td>21.9</td>
</tr>
<tr>
<td>Rheumatoid factor positive, %</td>
<td>79.4</td>
<td>75.5</td>
<td>75.0</td>
<td>77.2</td>
<td>23.9</td>
<td></td>
</tr>
<tr>
<td>Modified Sharp score</td>
<td>Mean (SD)</td>
<td>8.9</td>
<td>9.8</td>
<td>12.7</td>
<td>9.6</td>
<td>8.3</td>
</tr>
<tr>
<td>Median [quartiles]</td>
<td>3 [0–10]</td>
<td>2 [0–11]</td>
<td>6.5 [3.5–17.5]</td>
<td>3 [0–11]</td>
<td>2.5</td>
<td>9 [1–27]</td>
</tr>
<tr>
<td>Erosive RA**, %</td>
<td>66.7</td>
<td>60.0</td>
<td>87.5</td>
<td>64.2</td>
<td>59.1</td>
<td>77.7</td>
</tr>
</tbody>
</table>

* p < 0.05 vs baseline characteristics of patients followed up. ** Modified Sharp erosion score over 0. OEH: Patient’s overall evaluation of health.
were performed using SAS software with alpha level = 0.05 for significance using stepwise forward and backward procedures. All analyses of significance in univariate analysis. The final model was checked for duration of followup, and baseline variables reaching the 0.10 alpha level model. Candidate variables for multivariate analysis were age, sex, country, formed outcome variable was then analyzed using a linear regression boxcox transformation prevented departure from normality. The transformation of the modified Sharp score at each measurement time was skewed. A followup was assessed by Spearman correlation coefficient. The distribution of the modified Sharp score at each measurement time was skewed. A boxcox transformation prevented departure from normality. The transformed outcome variable was then analyzed using a linear regression model. Candidate variables for multivariate analysis were age, sex, country, duration of followup, and baseline variables reaching the 0.10 alpha level of significance in univariate analysis. The final model was checked for stability using stepwise forward and backward procedures. All analyses were performed using SAS software with alpha level = 0.05 for significance.

RESULTS
Patient characteristics at baseline and at followup are presented in Table 1. A total of 318 patients (70.4% were women) aged 52.4 years, with a mean (SD) of 2 (1.2) years’ disease duration at baseline, were followed over an average 30 months. Baseline and followup median (quartiles) modified Sharp scores were 3 (0–11) and 9 (1–27), respectively, with 35.9% and then 22.3% of patients, respectively, with no radiological damage.

Table 2 describes the prognosis univariate analysis. There was a significant univariate relationship of a higher modified Sharp followup score with a longer duration of followup (r = 0.18), a higher modified Sharp score at baseline (0.82), a higher ESR (0.30), a higher HAQ score (0.21), and a poor physician global assessment by Karnofsky index (0.24). RF seropositivity was not related with modified Sharp score at followup in univariate nonparametric analysis.

DISCUSSION
This investigation, following 318 patients, is one of the largest cohort studies of prognostic factors for radiographic damage in RA. Our results reveal good performance of a combination of characteristics in the first years of disease, i.e., baseline radiographic score, time from disease diagnosis, RF positivity, higher ESR, worse patient overall estimation of health, and longer duration of followup. According to the final model R², these variables explained 76.8% of the variance of the outcome variable.
The representativeness of this European sample has been reported. Briefly, sampling representativeness was ensured on a geographical basis, conforming to local public and/or private practices and referral systems in each county and country in order to comprehensively cover each health care system. Among the incident and prevalent cases identified in each catchment area in each country, the characteristics of patients included in the cohort did not differ from those not included. In this cohort, we studied a subsample of 318 patients with RA with 2 years’ mean disease duration at entry among 516 patients with baseline radiographs available. Those without followup hand radiographs did not differ by age, sex, clinical characteristics, or baseline radiograph scores, thus ruling out a selection bias based on clinical severity of disease during followup.

For this population, the multivariate linear regression model revealed that time from disease diagnosis, RF positivity, ESR, patient’s overall evaluation of health at baseline, radiographic articular damage at baseline, and duration of followup were prognostic factors, i.e., were associated with the level of joint destruction after an average 2.5 years of disease progression. The negative adjusted regression coefficient of the time from disease diagnosis in the multivariate model can be interpreted as follows: the longer the time from onset at a given radiographic damage at baseline, the less the radiographic damage after a given duration of followup. This illustrates a decrease in the progression of radiographic damage over time, which differs from recent findings. The baseline radiographic score is strongly associated with score at followup, which is responsible for a high part of the variance explained in the multivariate model. Thus, additional variables add lower contribution to the model. At onset, disease activity expressed by ESR has already proved to be a poor prognostic factor, as is CRP. Many studies show that RF positivity is commonly a poor prognostic factor. Other prognostic factors were inconsistent in some studies, such as older age, tender joint count, grip strength, cumulative joint inflammation, HLA-Dw4, early radiological damage, nodes, disease duration, and low hemoglobin. The heterogeneity of samples and the different methodologies explained discrepant results across previous studies. Several studies involved early RA, with disease development for less than 2 years. Since our study covers RA of less than 5 years at inclusion, it allows extension of the predictive value of these factors, i.e., disease activity, RF positivity, and baseline joint damage, to a longer time frame after disease diagnosis. The treatment regimen in this observational cohort has been documented. They did not enter the final multivariate models for predicting final joint damage, possibly because their effect might be too heterogeneous or too weak compared to other determinants. This might change in the future with new therapies.

Among identified determinants, disease activity assessed by ESR adds moderate contribution to the model, while baseline joint damage contributes for the major part to the outcome. The new therapies might be active modifiers over these 2 prognostic factors, thus changing the disease profile in the future.

Other studies have proposed consideration of genetic markers as prognostic factors. However there is a debate about the added value of genetic predictors over and above that for RF, since RF likely has a higher or similar level of prediction, with a lower cost.

Our study confirms the prognostic role of a set of features (RF positivity, ESR, patient’s overall evaluation of health at baseline, radiographic articular damage at baseline, shorter time from disease diagnosis, and duration of followup) identified in smaller cohorts. It indicates which disease characteristics should be focused on in the early years of disease to determine which patients are at higher risk of developing a severe disease. These results are also of practical interest for identifying patients with RA who may be candidates for new aggressive therapies, thus needing early referral and early treatment. The multinational structure of the study and the control of the method indicate that the results can be generalized to other settings and patient cohorts.

**ACKNOWLEDGMENT**

The authors thank the rheumatologists S. van der Burg (Martini Hospital,
REFERENCES


