Radiographic Progression on Radiographs of Hands and Feet During the First 3 Years of Rheumatoid Arthritis Measured According to Sharp’s Method (van der Heijde Modification)

DÉSIRÉE M.F.M. van der HEIJDE, MIEK A. van LEEUWEN, PIET L.C.M. van RIEL, and LEVINUS B.A. van de PUTTE

ABSTRACT. Objective. To determine the development rate of erosions and joint space narrowing in a cohort of patients during the first 3 years of rheumatoid arthritis (RA).

Methods. All consecutive patients fulfilling the American Rheumatism Association criteria and seen within the 1st year of the disease were followed prospectively with biannual radiographs of hands and feet. One hundred and forty-seven patients were followed for 2 years and 90 patients for 3 years. Erosions and joint space narrowing were scored with a modified version of Sharp’s method (van der Heijde modification). Wilcoxon’s rank sum test was used to test differences between joints and between erosions and joint space narrowing.

Results. On average, at 3 year followup most groups of joints showed about 8% of the maximum possible score. In most groups of joints about 20% of the joints were affected. At the start more foot joints were affected than hand joints. However, the increase in the number of affected joints and in the radiographic damage was similar in hands and feet. Consequently, the predominance of affected foot joints was still present after 3 years. The progression in the 3rd year was statistically significantly less compared to the 1st year. This was more pronounced for the number of affected joints than for radiographic damage.

Conclusion. We found that 70% of the patients showed radiographic damage after 3 years and this group could already be selected after 1 year. Overall, ± 18–20% of the joints were affected after 3 years, with relatively few abnormalities per joint (± 8% of the maximum possible score was reached). The rate of progression in the 1st year was significantly more than in the 2nd and 3rd years, indicating a flattening of the curve of radiographic progression.

Key Indexing Terms:
RADIOGRAPHY
EROSIONS
RHEUMATOID ARTHRITIS
PROSPECTIVE FOLLOWUP

Damage seen on radiographs is a result of the chronic inflammatory process in rheumatoid arthritis (RA). This damage is largely irreversible. The course of radiographic damage is mostly determined on radiographs of hands and feet. Many different scoring methods are used. These methods describe various abnormalities such as erosions, joint space narrowing, cyst formation, malalignment, etc. Some methods are mainly qualitative and others are more quantitative. Most studies on the course of radiographic damage are cross sectional in patients with varying disease duration. Just a few prospective followup studies in patients with recent onset RA are available. Both differences in scoring methods and the different populations of patients with RA that have been studied makes it hard to draw conclusions on the course of abnormalities caused by RA as they are seen on radiographs. We describe the course of damage in hands and feet measured by a quantitative scoring method (van der Heijde modification of Sharp’s method) in patients with RA during the 1st 3 years of their disease.

MATERIALS AND METHODS

Patients. All consecutive patients meeting the American Rheumatism Association criteria (1958) were followed prospectively at the departments of rheumatology of the University Hospital Nijmegen and the University Hospital Groningen in the Netherlands. The disease duration of all patients at entry was shorter than 1 year; no patient was treated with 2nd line drugs before the start of the study.

All consecutive patients (n = 147) whose followup was at least 2 years were included in this analysis: 90 patients had a followup duration of at least 3 years. Characteristics at entry: 64% female, age 51 ± 15 yr, 82% IgM rheumatoid factor (IgM-RF) ≥ 5 IU; 59% HLA-DR4, erythrocyte sedimentation rate (ESR) 43 ± 31 mm after 1 h, number of swollen joints 14 ± 8.

During the study they were treated with various second line agents: only nonsteroidal antiinflammatory drugs (NSAID) n = 14; hydroxychloroquine n = 85; sulfasalazine n = 57; auranofin n = 10; aurothioglucose n = 61; methotrexate n = 15; azathioprine n = 12; penicillamine n = 22; corticosteroids n = 37.

Radiographic analysis. Biannual radiographs of hands and feet in posteroan-
ter view were made with a Lanex Single/Fine Kodak screen and an Ortho-
M Kodak film (high speed). The radiographs were scored in random order
by one observer in each department, in chronological order per patient. The
observers were not informed of the patients' identities. The number of ero-
sions (E) and joint space narrowing (N) were scored according to a modi-
fied version of Sharp's method1. This scoring method has been validated
and described before2.

Erosions were counted in the 10 metacarpophalangeal (MCP) joints, the
8 proximal interphalangeal (PIP) joints, the 2 interphalangeal joints of
the thumbs, the right and left 1st metacarpal bone, the right and left radius
and ulnar bones, the right and left trapezium and trapezoid (as one unit; multi-
angular), right and left navicular bones, right and left lunate bones, the ten
metatarsophalangeal (MTP) joints, and the 2 interphalangeal joints of the
great toes. Joint space narrowing was assessed in the 10 MCP joints, the 8
PIP joints, right and left 3rd, 4th and 5th carpometacarpal joints, right and
left multangular-navicular joints, right and left capitae-navicular-lunate
joints, right and left radiocarpal joints, the 10 MTP joints and the 2 inter-
phalangeal joints of the big toes. Erosions were scored if there was a dis-
crete interruption of the cortical surface; if there was a greater defect, a
score according to the surface of the joint involved was given. Consequently,
for confluent erosions the score could not decrease. In the hands the maxi-
mum erosion score in a joint was 5; in the feet it was 10. Four grades of
joint space narrowing were recognized: 1 = focal or doubtful; 2 = general,
<50% of the original joint space; 3 = general, >50% of the original joint
space or subluxation; 4 = analgesy. Thus, the maximum number of ero-
sions in the hands was 160 and in the feet 120; and the maximum scores
for joint space narrowing were 120 and 48, respectively. Scores for ero-
sions and joint space narrowing were added up to give the total score.

The number of erosions (E) and joint space narrowing (N) of the various
groups of joints were analyzed separately. The percentage of damage was
determined by dividing the actual score by the maximum reachable score
for that group of joints, ensuring comparability between the various groups
of joints and between the erosions and narrowing. Beside the percentages
of the maximum score, the number of affected joints (separate for E and
N) were counted. This was again expressed as the percentage of the cor-
responding number of joints. In all analyses the percentage of damage and
the percentage of maximum affected joints was used.

The following groups of joints were analyzed separately and in combi-
nation: PIP joints, MCP joints, carpometacarpal, multangular-navicular,
capitate-navicular-lunate and radiocarpal joints (WRIST), MTP and the 2
interphalangeal joints, HANDS (PIP + MCP + WRIST), TOTN (narrowing
of all groups of joints), TOTE (erosions of all groups of joints), TOTE + N
(erosions and narrowing of all groups of joints).

To evaluate the interobserver variation, 96 pairs of radiographs of hands
and feet have been scored by both observers of this study. The Spearman
correlations between the 2 observers were TOTE 0.92, TOTN 0.80,
TOTE + N 0.90. High correlations do not rule out the possibility of sys-
tematic differences. Wilcoxon signed rank tests were performed between
the scores of the 2 observers to check this. One observer scored signifi-
cantly more narrowing than the other observer (mean 15.1 vs 12.1). No
systematic differences could be found in any of the other scores.

The intraobserver variation for both observers was computed in the same
way. Observer 1 scored 25 randomly chosen pairs of radiographs twice with
a minimum time of 6 months between the 2 scores and observer 2 scored
27 pairs. The Spearman correlations were for TOTE 0.94 and 0.99, for
TOTN 0.94 and 0.99, and for TOTE + N 0.96 and 0.99 for the 2 observ-
ers, respectively. No systematic differences were present in any of the scores.

Statistical analysis. The Wilcoxon signed rank test was used to test the differ-
ence between various groups of joints and between erosions and narrowing.
Radiographic damage measured with this method may be static or may
progress, but not decline. To obtain a smooth curve with duration of dis-
tease from the 57 patients over only 2 years and the 90 patients over 3 years
of followup, the following procedure was used to combine the data of the
2 periods. The means of all data of 147 patients were calculated for t = 0
to t = 2 yr. The same was done for the 90 patients for t = 0 to t = 3 yr.

The increment of this group from t = 2 to t = 3 was attached to the level
at t = 2 for all patients. Following the above approach prevented a dip in
the curve, which was not a real dip but only a consequence of a difference
in study population. The curves are presented as a table.

RESULTS

Course of radiographic progression. The radiographic progression in terms of both the percentage of maximum score and the percentage of maximum number of affected joints is summarized in Table 1. This is split for erosions and narrowing and various groups of joints.

On average, at 3 year followup most groups of joints show about 8% of the maximum score. The narrowing of the PIP is somewhat less (4%) and the narrowing of the wrist much higher (17%). The percentage of maximum number of joints is about 18-20% for most groups. Again fewer PIP show narrowing (10%) and more joints in the wrist show narrowing and more MTP show erosions (30% and 28%, respectively) compared with the other groups of joints. In general, the percentage of number of affected joints is 2-3 times the percentage of maximum score, meaning that many joints show little damage. In Table 2 the medians with 90% confidence intervals are given for the percentage of maximum score for some overall scores. The medians give information supplementary to the means. The medians are substantially lower than the means (Table 1, i.e., negatively skewed). This indicates that many patients have few or no radiographic abnormalities, while patients who show radiographic damage have many abnormalities.

Erosions vs narrowing. The average percentage of total nar-
rrowing (TOTN) over all time points was significantly larger than the average percentage of total erosions (TOTE, signed rank test p = 0.007), meaning that the TOTN was higher than the TOTE during the whole followup period (Table 1). This was mostly a consequence of the evident narrowing in the wrist. The increase in radiographic damage is given in Table 3, split up for the 1st, 2nd, and 3rd years. The in-
crease over the whole period was again more in the percen-
tage of total narrowing compared with the percentage of total erosions (signed rank test p <0.01).

Hands vs feet. The mean total score of the hands
(HANDE + N) during the whole period was not signifi-
cantly different from the mean total score of the feet (MTPE
+ N). On the contrary, more foot joints than hand joints
were affected (Table 1, signed rank test p = 0.01). This was
the same if only MCP were compared with MTP. The in-
crease in score and in number of affected joints was similar
for both hands and feet. This means that the predominance of
affected foot joints present at the start did not increase or
decrease.

Progression of radiographic damage in Year 1, 2, 3. We tested
for differences in progression of radiographic damage in the
first, 2nd, or 3rd year of RA (Table 3). The increase in per-
centage of maximum of the total score of the hands and of

van der Heijde, et al: Radiographs in early RA

1793
Table 1. Radiographic progression of groups of joints in hands and feet (0, 1.0, and 2 yr, n = 147; 3 yr, n = 90, mean values)

<table>
<thead>
<tr>
<th></th>
<th>% of Maximum Score</th>
<th>% of Maximum Number Joints</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-1.0</td>
<td>2.0-3.0</td>
</tr>
<tr>
<td>TOTE</td>
<td>0.4</td>
<td>3.4</td>
</tr>
<tr>
<td>TOTN</td>
<td>0.2</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>4.2</td>
</tr>
<tr>
<td>WRISTE</td>
<td>0.3</td>
<td>2.3</td>
</tr>
<tr>
<td>WRISTN</td>
<td>1.1</td>
<td>7.6</td>
</tr>
<tr>
<td>HANDE</td>
<td>0.4</td>
<td>2.9</td>
</tr>
<tr>
<td>HANDN</td>
<td>0.9</td>
<td>5.0</td>
</tr>
<tr>
<td>HANDE+N</td>
<td>0.6</td>
<td>3.8</td>
</tr>
<tr>
<td>MTPN</td>
<td>2.1</td>
<td>5.6</td>
</tr>
<tr>
<td>MTP + N</td>
<td>1.6</td>
<td>4.8</td>
</tr>
<tr>
<td>TOTE</td>
<td>0.9</td>
<td>3.6</td>
</tr>
<tr>
<td>TOTN</td>
<td>1.2</td>
<td>5.2</td>
</tr>
<tr>
<td>TOTE+N</td>
<td>1.0</td>
<td>4.2</td>
</tr>
</tbody>
</table>

PIPE = erosions of PIP joints, PIPN = narrowing of PIP joints, MCPE = erosions of MCP joints, MCPN = narrowing of MCP joints, WRISTE = erosions of wrist joints, WRISTN = narrowing of wrist joints, HANDE = erosions of PIP, MCP and wrist joints, HANDN = narrowing of PIP, MCP and wrist joints, HANDE+N = summation of erosions and narrowing in PIP, MCP and wrist joints, MTPN = erosions of MTP and 1st IP joints, MTPE = erosions of MTP and 1st IP joints, MTPN = narrowing of MTP and 1st IP joints, MTPE + N = summation of erosions and narrowing of MTP and 1st IP joint, TOTE = erosions of all hand and foot joints, TOTN = narrowing of all hand and foot joints.

Table 2. Medians and 90% confidence intervals of the radiographic damage in groups of joints of hands and feet (0, 1.0, and 2.0 yr: n = 147; 3.0 yr: n = 90; expressed as % of maximum score)

<table>
<thead>
<tr>
<th></th>
<th>0-1.0</th>
<th>1.0</th>
<th>2.0</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>HANDE+N</td>
<td>0.0 (0-0.2)</td>
<td>1.4 (0.4-2.5)</td>
<td>3.2 (1.5-4.0)</td>
<td>3.2 (0.8-6.5)</td>
</tr>
<tr>
<td>MTPE+N</td>
<td>0.0 (0-0.2)</td>
<td>1.8 (0.3-2.4)</td>
<td>2.4 (1.2-3.6)</td>
<td>3.6 (1.8-5.4)</td>
</tr>
<tr>
<td>TOTE</td>
<td>0.0 (0-0.2)</td>
<td>1.4 (0.6-2.5)</td>
<td>3.2 (1.5-4.3)</td>
<td>5.0 (1.8-6.8)</td>
</tr>
<tr>
<td>TOTN</td>
<td>0.0 (0-0.3)</td>
<td>1.8 (0.6-3.6)</td>
<td>3.6 (1.8-7.2)</td>
<td>3.6 (0.6-6.6)</td>
</tr>
<tr>
<td>TOTE+N</td>
<td>0.2 (0-0.5)</td>
<td>2.0 (1.4-3.2)</td>
<td>4.2 (2.4-5.0)</td>
<td>4.4 (1.6-7.9)</td>
</tr>
</tbody>
</table>

Table 3. Progression in the 1st, 2nd (n = 147) and 3rd years (n = 90) and comparison of the progression in the 1st with the progression in the 2nd and 3rd years (Wilcoxon signed rank test)

<table>
<thead>
<tr>
<th>Progression</th>
<th>Mean % of Maximum Score</th>
<th>Mean % of Maximum Number Joints</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-1 yr</td>
<td>1-2 yr</td>
</tr>
<tr>
<td>HANDE</td>
<td>2.5</td>
<td>1.8†</td>
</tr>
<tr>
<td>HANDN</td>
<td>4.1</td>
<td>2.9†</td>
</tr>
<tr>
<td>HANDE+N</td>
<td>3.2</td>
<td>2.3†</td>
</tr>
<tr>
<td>MTPE</td>
<td>3.0</td>
<td>1.8†</td>
</tr>
<tr>
<td>MTPN</td>
<td>3.5</td>
<td>1.6†</td>
</tr>
<tr>
<td>MTPE+N</td>
<td>3.2</td>
<td>1.8*</td>
</tr>
<tr>
<td>TOTN</td>
<td>4.0</td>
<td>2.5†</td>
</tr>
<tr>
<td>TOTE</td>
<td>2.7</td>
<td>1.8†</td>
</tr>
<tr>
<td>TOTE+N</td>
<td>3.2</td>
<td>2.1†</td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < 0.01; † NS (compared with the progression in the first year).
the feet was significantly less in the 3rd year compared with the first year. The increase in the feet was also significantly less in the 2nd year. The progression of the number of affected joints was significantly less in the 2nd year and 3rd year in comparison with the first year in all groups of joints. The lesser progression in the 2nd and the 3rd year was related to the number of affected joints rather than the degree of damage.

DISCUSSION

Radiographic damage is the irreversible result of chronic joint inflammation and is often considered as an important factor in judging the course of the disease and efficacy of slow acting antirheumatic drugs (SAARD). Curiously, the natural course of the radiographic damage is not known. Ideally, a prospective followup of early RA is needed to study the "natural" course. However, this is not the natural course sensu strictu, because medication is always included. Many cross sectional studies on this are based on RA of widely varying disease duration. If these investigations are carried out in a department of rheumatic diseases, selection bias may play an important role: e.g., patients with a good prognosis, and thus, little radiographic damage, may no longer attend the rheumatic clinic, while patients with very severe disease may have died and therefore not been included in analysis.

We agree with Sharp that it is a misconception that, to be useful, the same radiological scores must be found by different individuals or when used repeatedly by the same observer. Papers on the reproducibility of radiographic scoring reveal that differences between observers may be quite large but that these differences are generally consistent. Therefore interobserver correlations are usually high. In our study, the correlations between the 2 observers for the narrowing score, erosion score, and total score were 0.80, 0.92, and 0.90, respectively. These correlations are quite high and fully in accordance with the literature. In a study by Sharp, et al on the reproducibility of multiple observer scoring of radiologic abnormalities, correlations ranged from 0.53 to 0.96. The high correlations between the 2 observers and the very high consistency of each observer (correlations ranging from 0.94 to 0.99) indicate that the results of the radiographic progression may be considered reliable. The absolute figures of the radiographic abnormalities are more indicative, particularly for the narrowing score as there was a significant difference in scoring between the 2 observers. Taking into account the pitfalls of radiographic scoring, this study gives a good indication of radiographic progression, whereas the absolute figures of the scores might be different if read by other observers.

In our prospective followup of 147 patients with early RA, we found a considerable percentage of radiographic damage. After 3 years 20% of the total numbers of hand and foot joints were affected. The percentage of maximum damage was 8.4%, meaning that many joints were affected with relatively little damage per joint. Möttönen found 2.3% of the small joints of the hands and feet eroded at the start and 13.1% after 2 years of followup of 58 patients with early RA. These figures are very similar to our results: 2.7% at the start and 16.4% after 2 years (Table 1).

The medians, as presented in Table 2, give some (indirect) information on the number of patients with abnormalities. The medians are considerably lower than the means, indicating that many patients have little radiographic damage. However, if patients have radiographic abnormalities, they show considerable damage.

In our group 30% of the patients showed <1% joint damage in hands and feet after 3 years' followup. This figure was the same after 2 years. Only 1 patient (out of 90) who had <1% damage at year 1 showed substantial progression in the next 2 years. Therefore, patients with little radiographic damage can already be selected after 1 year of followup. The percentages of patients without abnormalities in the literature are similar to our study: Brook and Corbett 23%, Möttönen 24%, and Eberhardt 28%. In 2 cross sectional studies on longstanding RA by Thould and Fuchs, the percentages of patients without abnormalities were only 9 and 4.5%, respectively. However, this may be due to selection of patients: patients with good prognoses no longer attend the rheumatological department.

Some studies included only the radiographic progression of hands, although several studies indicate that foot joints may be affected at an early stage of disease and erosiveness was even found twice as much in feet as in fingers. We also found that more foot than hand joints were affected during the whole followup period. At the start this difference was already present. The progression in the feet and hands was similar, so that the difference did not increase during the 3 year followup. This means that the damage in the foot joints developed in an earlier stage than in the hand joints. The distinct narrowing in the wrist is striking. As a consequence of this pronounced narrowing in the wrist, the percentage of narrowing is larger than the percentage of erosions over the whole period.

It is striking that most progression was seen in the 1st year. The increase in damage was less in the 2nd and 3rd years. This declining progression in the 2nd and 3rd years was more pronounced in the number of affected joints than in the radiographic score. Thus in the 2nd and 3rd year relatively few new joints became damaged but the increase in joint damage continued on a relatively higher level (but less than in the 1st year). This observation means that a damaged joint is more prone to become more seriously damaged than an unaffected joint to become damaged. This progression of the severity of an affected joint may be due to several factors. First, it may be much harder to diminish the pannus in a joint than to prevent a new inflammation in a "virgin" joint. Second, biomechanical factors may play a role in the progres-

von der Heijde, et al: Radiographs in early RA

1795
sion, independent of inflammation. An additional factor might be that fewer joints are "at risk" to become damaged. This seems less important because ± 80% of the joints can become damaged.

The decline in progression after the 1st year causes a flattening curve of the radiographic course. In our study, also, a relatively large amount of radiographic damage in the 1st 3 years is present. Other studies also found a considerable quantity of radiographic damage in the first few years and a flattening of the progression thereafter 6-13. Möttönen, et al found a linear progression in the hands and the greatest progression in the feet between 6 and 18 months during a followup of 2 years 5.

We can only speculate on the cause of the decline of the progression rate. One of the possibilities is the effect of SAARD 14. In our group, most patients were treated with SAARD in the very early phase of the disease. This approach may explain the diminishing progression of radiographic damage. However, this can only be answered definitely in controlled, randomized clinical trials.

Not knowing the cause of this difference in progression rate, we need to be cautious with the use of radiographs as an outcome variable in trials of SAARD. It is of great importance in the design of such a study that the patients have a comparable disease duration. Especially in the early phase of the disease, a relatively small difference in disease duration may be confounding. However, during the 1st 3 years of disease duration a ceiling effect of the scoring method seems no problem as only 8-10% of the maximum possible score is reached.

In conclusion, 70% of the patients showed radiographic damage after 3 years of followup and this group could already be selected after 1 year. Overall, ± 18-20% of the joints were affected after 3 years, with relatively few abnormalities per joint (± 8% of the maximum score was reached). The rate of progression in the 1st year was significantly greater than in the 2nd and 3rd years, indicating a flattening of the curve of radiographic progression.

REFERENCES