MRI of the wrist in early rheumatoid arthritis can be used to predict functional outcome at 6 years

N Benton, N Stewart, J Crabbe, E Robinson, S Yeoman, F M McQueen

Objectives: To determine whether magnetic resonance (MR) scans of the dominant wrist of patients with early rheumatoid arthritis (RA) can be used to predict functional outcome at 6 years' follow up.

Methods: Dominant wrist MR scans were obtained in 42 patients with criteria for RA at first presentation. Patients were followed up prospectively for 6 years, and further scans obtained at 1 year (42 patients) and 6 years (31 patients). Two radiologists scored scans for synovitis, tendonitis, bone oedema, and erosions. The Stanford Health Assessment Questionnaire (HAQ) score, indicating functional outcome, and standard measures of disease activity were assessed at 0, 1, 2, and 6 years. The physical function component of the SF-36 score (PF-SF36) was also used as a functional outcome measure at 6 years.

Results: Baseline MR parameters, including bone oedema score and the total baseline MR score, were predictive of the PF-SF36 at 6 years ($R^2 = 0.22, p = 0.005$ and $R^2 = 0.16, p = 0.02$, respectively). The PF-SF36 score correlated strongly with the HAQ score at 6 years ($r_s = 0.725, p < 0.0001$); none of the baseline MR parameters predicted the 6 year HAQ score. The total MR score obtained at 1 year was predictive of the 6 year HAQ ($R^2 = 0.04, p = 0.01$). Standard clinical and radiographic measures at baseline were not predictive of the 6 year PF-SF36, but when combined in a model with baseline MR oedema score, prediction increased from 0.09 to 0.23, or 23% of the 6 year variance.

Conclusion: MR imaging of the wrist in patients with early RA can help to predict function at 6 years and could be used to plan aggressive management at an earlier stage.

The ability to predict accurately prognosis at presentation in patients with rheumatoid arthritis (RA) is becoming an increasingly important part of clinical management. It is recognised that as many as 80–90% of those patients who develop erosive disease will begin to do so within the first 2–3 years after symptom onset. Significant functional disability is also common in relatively early disease, occurring by 5 years in 47% of a cohort of patients from the Norfolk Arthritis Register. With the advent of biological and conventional disease modifying antirheumatic drugs (DMARDs) possessing powerful anti-erosive effects, it may be possible to greatly reduce or even completely prevent joint damage during this crucial period and reduce the risk of an adverse functional outcome.

Traditional factors predicting long term clinical disability include the level of disability as measured by the Stanford Health Assessment Questionnaire (HAQ score), seropositivity for rheumatoid factor, the presence of rheumatoid nodules increased age at onset and, in some studies, female sex. Whether disease activity measured at baseline can predict long term disability remains controversial, but the area under the curve for C reactive protein (CRP) measurements over the first 3 years of disease has been shown to predict radiological progression. A close relationship exists between erosive joint damage and clinical disability in late disease (>10–12 years), but over the first 3–6 years functional capacity is more strongly associated with joint inflammation.

We studied an inception cohort of 42 patients with early RA, from presentation (median of 4 months from symptom onset) to 6 years, using clinical assessments of disease activity and function as well as plain radiography and magnetic resonance (MR) scans of the dominant carpus, scored according to our own validated scoring system. We have previously shown that the MR erosion score at baseline is predictive of x ray erosion scores at 1, 2, and 6 years. In this paper we examine the possibility that a baseline MR examination might be used to predict functional disability in the medium term (6 years from onset) as measured by the physical function component of the SF-36 score (PF-SF36) and the HAQ score. We also compared the prognostic power of baseline MR parameters with traditional clinical measures such as the CRP and conventional radiography.

MATERIALS AND METHODS

Patient group and clinical assessments
An inception cohort of 42 patients with early RA has been studied since symptom onset. Details of recruitment, baseline demographics, and clinical assessments have been described. To summarise briefly, all patients fulfilled the 1987 American College of Rheumatology criteria for RA and had had symptoms for 6 months or less (median 4 months) at entry to the study. All patients were assessed clinically for disease activity using a tender joint count (temperomandibular, acromioclavicular, sternoclavicular, shoulders, elbows, wrists, metacarpophalangeal (MCP) joints, proximal interphalangeal (PIP) joints, hips, knees, ankles, midtarsal joints, metatarsophalangeal joints and PIP joints of the foot; maximum possible = 60), swollen joint count (maximum possible = 58, hips excluded), Ritchie score, pain score, HAQ score, erythrocyte sedimentation rate, CRP, and disease activity score (DAS) at baseline, 1, 2 and 6 years. The DAS was calculated using three variables according to the method described by van der Heijde et al. Functional outcome was measured using the HAQ score and the PF-SF36. The HAQ
was measured at baseline, and at 1, 2, and 6 years. The PF-SF36 was assessed only at 6 years.

Of the original 42 patients enrolled, all were fully evaluated at baseline and at 1 year with x-ray examination, magnetic resonance imaging (MRI), and clinical data. At 6 years, four patients were untraceable or had moved overseas and four withdrew from further participation. Thus a total of 34 patients had clinical and functional assessments at that point. Three of these patients did not have an MR evaluation for technical reasons (two patients had internal fixation of the wrist, one patient had a permanent pacemaker). Table 1 summarises the demographic, clinical, and current drug treatment details for the 6 year cohort.

**MR scans**

An MR scan of the dominant wrist was obtained using a 1.5 Tesla MR scanner (GE Signa Horizon) with a dedicated wrist coil (Medical Devices). The same scanner was used at baseline, 1 year and 6 years. The method of MR scanning has been previously described. In brief, the hand was placed in the wrist coil where it fitted snugly by the patient’s side with the palm facing the body, the thumb pointing anteriorly. Each sequence of the 6 year follow up study was planned using localising sequences to match the first study sequence as closely as possible. Parameters were identical for all studies. The field of view was 8 cm and included the distal radioulnar, radiocarpal, and midcarpal joints as well as the metacarpal bases. The small field of view was chosen to optimise resolution. The MCP joints were not included in the field of view. Coronal and axial T1 sequences were performed, followed by axial fat suppressed fast spin echo T2, then coronal fat suppressed T1. Erosions were only scored if visible in two planes, with a cortical break seen in at least one plane. MR scans were also scored for bone oedema, synovitis, and tendinitis as has been previously described. A total MRI score for the carpus was derived from the sum of scores for erosions, bone oedema, synovitis, and tendinitis. MR scans were scored independently by the same two musculoskeletal radiologists who scored baseline and 1 year scans. They were unaware of the clinical and radiological data. Scans at 6 years were scored without reference to baseline or 1 year scans. At each of the three times, validation of the scoring system has been performed for all MRI parameters.

**Radiographic assessments**

The radiographic assessments have been fully described in a recent publication. Modified Sharp scores were obtained from baseline anteroposterior films of the hands and feet in the 34 patients assessed for functional outcome at 6 years. Scoring was performed by two observers (a rheumatologist and a radiology registrar) after preliminary training. Interobserver reliability has been reported. Scores of the two observers were averaged before use in data analysis.

**Statistical analyses**

Spearman’s correlations were calculated to investigate relationships between functional, clinical, and MR measures at baseline and 6 years. Linear and logistic regressions were used to investigate relationships between baseline and 6 year measures. The R² quoted in the logistic regression is a generalised coefficient of determination. The adjusted R² was used in exploratory model building to produce subsets of independent baseline variables that best predicted outcomes at 6 years in linear regressions. Square root and log transformations were used respectively for 6 year MR erosion scores and CRP scores to satisfy normality assumptions. The 6 year HAQ scores were grouped as <=0.25 and >0.25.

**RESULTS**

How do clinical, functional, and MR features change over 6 years?

Figure 1 summarises the pattern of change of MR, clinical, and functional parameters over the 6 years of observation. MR features have already been described but are presented here for comparison with changes in function. No significant change in MR synovitis score was seen, although a trend downwards was observed from a median level of 8 at baseline to 6.5 at 6 years. Similarly, the DAS fell from a median of 4.1 at baseline to 2.6 at 1 year, but then remained stable at 2.4 at 2 years and 2.7 at 6 years. MR bone oedema also remained constant over 6 years for the group as a whole, although scores fluctuated in individual patients, but MR bone erosions progressed significantly (p=0.0001). There was a “J” shaped curve to progression for the HAQ score, which fell from a baseline median of 0.6 to 0.16 at 1 year, paralleling the change in DAS, followed by a slow rise from 2 years to 0.33 at 6 years, paralleling the change in MR erosion score.

What baseline parameters predict 6 year MRI erosions?

Data were examined to determine which baseline measures were most predictive of the 6 year MR erosion score. Table 2 shows the results obtained. The baseline MR bone oedema

---

**Table 1** Demographic, clinical, and radiological characteristics of 34 patients assessed at 6 years of follow up

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female:male</td>
<td>21:13</td>
</tr>
<tr>
<td>Median age at diagnosis (years)</td>
<td>48.5</td>
</tr>
<tr>
<td>IgM RF ever positive during follow up</td>
<td>30 (88%)</td>
</tr>
<tr>
<td>Disease activity, median (range)</td>
<td>0.33 (0–1.6)</td>
</tr>
<tr>
<td>HAQ score</td>
<td>5–15</td>
</tr>
<tr>
<td>Swollen joint count</td>
<td>3.5 (5–21)</td>
</tr>
<tr>
<td>DAS</td>
<td>2.7 (0–4.6)</td>
</tr>
<tr>
<td>ESR (mm/1st h)</td>
<td>25 (12–59)</td>
</tr>
<tr>
<td>CRP (mg/l)</td>
<td>7 (3–69)</td>
</tr>
<tr>
<td>Radiographic characteristics</td>
<td></td>
</tr>
<tr>
<td>Joint erosions on x ray (hands and feet)</td>
<td>28 (82%)</td>
</tr>
<tr>
<td>Median total modified Sharp score</td>
<td>27</td>
</tr>
<tr>
<td>Drugs</td>
<td></td>
</tr>
<tr>
<td>NSAID</td>
<td>20 (59%)</td>
</tr>
<tr>
<td>DMARD</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>21 (62%)</td>
</tr>
<tr>
<td>Methotrexate (7.5–17.5 mg/week)</td>
<td>14</td>
</tr>
<tr>
<td>Sulfasalazine (1–3 g/day)</td>
<td>9</td>
</tr>
<tr>
<td>Hydroxychloroquine (400 mg/day)</td>
<td>2</td>
</tr>
<tr>
<td>Leflunomide (20 mg/day)</td>
<td>2</td>
</tr>
<tr>
<td>Azathioprine (100 mg/day)</td>
<td>1</td>
</tr>
<tr>
<td>Combination</td>
<td>6</td>
</tr>
<tr>
<td>Ever DMARD use</td>
<td>31 (91%)</td>
</tr>
</tbody>
</table>

RF, rheumatoid factor; HAQ, Health Assessment Questionnaire; DAS, disease activity score; ESR, erythrocyte sedimentation rate; CRP, C reactive protein; NSAID, non-steroidal anti-inflammatory drug; DMARD, disease modifying antirheumatic drug.
score was strongly predictive of the 6 year MR erosion score ($\beta$ (SE) = 0.40 (0.13), $p = 0.005$), and the baseline score for the total MR score also reached significance ($0.15 (0.05)$, $p = 0.003$). Baseline clinical measures were not predictive of the MR erosion score at 6 years, apart from the CRP ($\beta$ (SE) = 0.038 (0.018), $p = 0.04$). Figure 2 shows an example of erosive progression seen on MR scans in a patient who was not compliant with DMARD treatment. This example shows how early bone marrow oedema on MR evolves into erosive change.

**Do MR scores correlate with function at baseline or 6 years?**

Correlations between MR scores and functional outcome were examined at baseline and at 6 years (table 3). The HAQ score was measured at both times, whereas the PF-SF36 was only measured at 6 years when it was highly correlated with the HAQ ($r_s = -0.725$, $p<0.0001$). At baseline, the MR bone oedema score correlated most strongly with the HAQ ($r_s = 0.44$, $p = 0.004$). The total MR score also correlated with the HAQ ($r_s = 0.31$, $p = 0.04$) and the MR erosion score, nearly so ($r_s = 0.44$, $p = 0.004$ and $r_s = 0.29$, $p = 0.06$, respectively). However, there was no correlation between MR synovitis and the HAQ score at baseline.

At 6 years again there was good correlation between the total MR score and measures of function (for HAQ, $r_s = 0.48$, $p = 0.006$; for PF-SF36, $r_s = -0.46$, $p = 0.009$). Of the individual MR parameters scored, bone erosion was the most strongly associated with function: (for HAQ, $r_s = 0.48$, $p = 0.007$; for PF-SF36, $r_s = -0.48$, $p = 0.006$). The degree of correlation between bone marrow oedema and function at 6 years had fallen and failed to reach significance.

**Can early MR parameters be used to predict 6 year function?**

Table 4 shows the baseline and 1 year MR predictors of 6 year HAQ and PF-SF36 scores. The baseline MR bone oedema score was predictive of the 6 year PF-SF36 ($R^2 = 0.22$, $p = 0.005$), as was the baseline total MR score ($R^2 = 0.16$, $p = 0.02$), but not individual scores for erosions, tendonitis, or synovitis. This same predictive pattern was also seen on analysis of the 1 year MR scores. None of the baseline MR scores and only the total MR score at 1 year predicted the 6 year HAQ ($R^2 = 0.04$, $p = 0.01$).

**Are clinical and radiological baseline and 1 year parameters useful predictors of 6 year function?**

Of the baseline clinical and radiographic scores, only the Ritchie index and the HAQ score were predictive of the 6 year HAQ ($R^2 = 0.2$, $p = 0.05$ for both) and none predicted the 6 year PF-SF36 (table 5). Data describing the predictive value of 1 year clinical scores are also shown. The 1 year Ritchie index, DAS, and the HAQ predicted the 6 year HAQ but only DAS and HAQ scores predicted the 6 year PF-SF36.

**A model incorporating baseline clinical and radiographic and MR measures is the best predictor of 6 year function**

The strongest baseline MR predictor of functional outcome at 6 years (as measured by the PF-SF36) was MR bone oedema. When this measure was added to a regression model which included baseline clinical (CRP, DAS, and HAQ) and radiographic measures (modified Sharp score) for the prediction of 6 year PF-SF36, it significantly increased the overall variance ($R^2 = 0.09$ for clinical and x ray measures $v R^2 = 0.23$ for clinical and x ray measures plus MR bone oedema).

**DISCUSSION**

Functional outcome in RA has traditionally been measured by the HAQ score. The HAQ measures functional ability in eight domains and focuses on patient centred outcomes rather than biological parameters and outcomes perceived by the physician. The HAQ is widely used in clinical practice and clinical trials but may have some limitations in that it emphasises physical activities using the arm and it tends to be dominated by effects on larger joints. The SF-36 is a

**Table 2 Prediction of 6 year MR erosion score by baseline MR and clinical scores**

<table>
<thead>
<tr>
<th>Baseline scores</th>
<th>$\beta$ (SE)</th>
<th>$p$ Value</th>
<th>$r_s^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR synovitis</td>
<td>0.24 (0.11)</td>
<td>0.03</td>
<td>0.15</td>
</tr>
<tr>
<td>MR synovitis+tendonitis</td>
<td>0.18</td>
<td>0.03</td>
<td>0.15</td>
</tr>
<tr>
<td>MR bone oedema</td>
<td>0.40 (0.13)</td>
<td>0.005</td>
<td>0.23</td>
</tr>
<tr>
<td>MR erosion score</td>
<td>0.34 (0.17)</td>
<td>0.06</td>
<td>0.12</td>
</tr>
<tr>
<td>Total MR score</td>
<td>0.15 (0.05)</td>
<td>0.003</td>
<td>0.27</td>
</tr>
<tr>
<td>CRP</td>
<td>0.04 (0.02)</td>
<td>0.04</td>
<td>0.14</td>
</tr>
<tr>
<td>DAS</td>
<td>0.73 (0.38)</td>
<td>0.07</td>
<td>0.11</td>
</tr>
</tbody>
</table>

MR, magnetic resonance; CRP, C reactive protein; DAS, disease activity score.
general health status questionnaire that has been well validated in the normal population and patients with a number of chronic conditions including RA. The SF-36 also has eight domains and is designed to measure overall health related quality of life rather than just physical function. However, the physical functioning component of the SF-36 is a useful adjunct to the HAQ, and the two scores have been shown to be closely correlated in patients with RA. This was confirmed in our cohort of patients at 6 years ($r_s = -0.725$). Therefore both scores were included as primary outcome measures in assessing function.

As far as we know, this study is the first to show that data from MR scans taken at first presentation of RA, can be used to predict function as measured by the PF-SF36 at 6 years. Although the total baseline MR score, which combined MR erosion, oedema, synovitis, and tendonitis scores, reached significance as a predictor, this was not as strong as the baseline MR bone oedema score alone, which predicted 22% of the variance of the 6 year PF-SF36 score ($R^2 = 0.22$, $p = 0.005$). The same MR measures did not predict the 6 year HAQ score, despite a high correlation between the HAQ and PF-SF36 at the 6 year mark. As 11/34 or one third of the patients had a HAQ score of zero at 6 years, assumptions of normality were not met and the data could not be transformed to meet these assumptions. Therefore, rather than perform the analysis with the HAQ score at 6 years as a continuous variable (as was appropriate for the SF-36, which was not skewed), data were dichotomised into two groups with 0.25 as the cut off score. Unfortunately, power was lost as a result of this.

![Sequential MR scans of the dominant carpus over 6 years in a 44 year old patient who was non-compliant with disease modifying treatment. This patient has congenital fusion of the triquetrum and lunate. (A) The baseline $T_1$ weighted coronal scan shows focal areas of bone oedema as low signal within the capitate and the triquetrum/lunate (short arrows). There is extensive synovitis adjacent to the scaphoid (long arrow). (B) Axial $T_1$ weighted image (baseline) showing bone oedema as low signal within the distal radius. (C) Equivalent baseline axial $T_2$ weighted image showing bone oedema within the radius as bright signal. (D) Coronal $T_1$ weighted scan at 1 year showing discrete low signal regions breaching the cortex of the capitate and the radial styloid indicating erosions (arrows). There is also extensive bone oedema within the hamate (circle). (E) Axial $T_1$ weighted image at 1 year showing erosion at the radius. (F) Coronal $T_1$ weighted scan at 6 years showing extensive erosive disease at multiple sites (arrows). (G) Axial $T_1$ weighted image at 6 years showing more extensive erosion at the radius.](http://www.annrheumdis.com)
manoeuvre, which may have led to the negative result for the HAQ score.

Others have examined early clinical and radiographic findings as predictors of functional outcome in RA. Welsing et al studied a cohort of 378 patients with RA over the first 9 years of disease and noted that the HAQ score initially declined during the first 12 months followed by a slow increase over the next 8 years in the same pattern as observed for our cohort. This initial fall in HAQ corresponded with a fall in disease activity as measured by the DAS score, which remained stable from thereon, as was also seen in our cohort. They showed that long term functional outcome was influenced by disease activity at baseline but not by early radiographic damage as measured by the Sharp score. This accords with the findings of Scott et al, who summarised data from a number of studies and showed that in the earliest phases of RA, x-ray damage and HAQ were not related, but significant correlations were found at 5–8 years from disease onset and increased in significance in late disease. Functional outcome is therefore likely to reflect components of inflammatory activity and structural damage to varying degrees depending on disease duration.

Why then should baseline MR be able to predict function in the medium term when plain radiography does not? We have already shown that MR scans can show rheumatoid erosions at the wrist much earlier than plain radiography, and this has been confirmed by others. In this cohort, 45% of the patients had erosions on MR at presentation compared with only 15% on plain radiography. MR erosion scores progressed relentlessly in this cohort, paralleling changes in radiographic erosive damage as measured by the Sharp score. Although we specifically examined the carpus, other groups have confirmed that MR can detect early rheumatoid erosions at MCP joints and also at the feet. Thus joint damage represented by erosive change can be seen earlier using MR scanning and is therefore likely to predict function earlier than radiographic features. Our finding that baseline total MR score predicts 6 year function is in keeping with this argument.

Baseline MR scans disclose more than early structural damage in the form of erosions. Synovitis can be seen and scored and was a predictor of the MR erosion score at 6 years ($r^2 = 0.15$, $p = 0.03$) but not of the total modified Sharp score. This accords with the findings of Østergaard et al, who showed that MR synovitis, measured by estimation of synovial volume, was a predictor of MR erosions after 1 year. Although MR synovitis alone was not predictive of the 6 year functional scores, it was incorporated into the total baseline MR score and might have contributed to its ability to predict the 6 year PF-SF36.

Finally, the most important property of MR imaging in obtaining a prognosis is its ability to show bone marrow oedema. We have shown that MR bone oedema correlated significantly with clinical markers of inflammation at baseline and at 6 years, including both CRP and DAS. Furthermore, using a site-specific analysis, we have shown that bone oedema affecting a specific carpal bone at baseline was highly likely to be associated with MR erosion at that site after 1 year and 6 years (OR = 6.5, 95% CI 2.78 to 18.1).
We have also shown that the baseline MR bone oedema score was predictive of the 6 year total Sharp score in this cohort. Bone oedema alone was also the strongest predictor of functional outcome at 6 years as measured by the PF-SF36, and when MR bone oedema was added to clinical and radiographic scores in a regression model analysis for the prediction of functional outcome, the predictive value was increased significantly.

MR bone oedema is important, therefore, as it contains information about two facets of rheumatoid joint disease likely to have a bearing on functional outcome: initial inflammatory activity, and the potential for erosion and development of structural joint damage.

DMARD treatment over the period of observation may have influenced the degree of joint damage which developed by 6 years, and this in turn may have impacted on function. There was no attempt to standardise management as this was an observational study. However, patients were treated with DMARDs which were commonly used in New Zealand in the late 1990s, including sulfasalazine (a first choice for many clinicians), methotrexate and, in some instances, combination treatment. Tumour necrosis factor a inhibitors were not available. The cohort included a number of patients who were non-compliant with DMARDs. If one assumes that there was a reduction in joint damage in some patients due to their DMARD treatment over 6 years, this would tend to weaken the association between baseline MR bone oedema score reflecting disease activity at the very start of treatment, and the outcome reflected by function at 6 years. Thus the true association may be stronger than we have observed.

In summary, this is the first study to show that early features seen on MRI can predict 6 year functional outcome in RA. MR bone oedema is the most significant predictor of function, probably because of its combined ability to reflect joint inflammation and to predict erosive damage. As early treatment of RA with conventional DMARDs and especially with biological agents may reduce radiographic damage and disability, the study of prognostic indicators applicable to very early disease has assumed considerable importance. We suggest that MRI may be used as a useful adjunct to standard clinical and laboratory determinants of prognosis.

ACKNOWLEDGEMENTS

We acknowledge the assistance of the following clinicians who have referred patients for this study: Dr Mike Butler, Dr David Caughey, Dr Nora Lynch, Dr Alan Doubt, Dr Hamish Hart, Dr Peter Gow, Dr Raoul Stuart, Dr Terry Macedo, Dr Max Robertson, Dr Roger Reynolds, Dr Bob Grigor. We are most grateful to Ms Rika Nel, technician at the Auckland Radiology Group, who supervised the MRI scans. Supported by grants from the Health Research Council of New Zealand, the Arthritis Foundation of New Zealand, the Auckland Medical Research Foundation, Lotteries Health, New Zealand, the Auckland Radiology Group, and Sanofi-Winthrop.

**REFERENCES**

15. McQueen FM, Benton N, Perry D, Crabbe J, Robinson E, Yeoman S, et al. Bone oedema scored on magnetic resonance imaging scans of the dominant corpus at presentation predicts radiographic joint damage at the hands and