Radiographic progression in rheumatoid arthritis: Does it reflect outcome? Does it reflect treatment?

D van der Heijde

Abstract
Although there is clear face validity that structural damage is related to outcome, it is difficult to prove this. Recently, more evidence became available that structural damage, as assessed on plain films, is indeed related to disease activity in clinical trials and therefore can be used to assess the effect of treatment. Also, a relationship between structural damage and outcome, mainly defined as physical disability, was established. Several examples of findings in recent publications are presented which lead to the following conclusions. There is a relation between the response to treatment measured as clinical disease activity and measured as radiographic progression in most clinical therapeutic trials. A strong relation between local inflammation and progression of damage in the individual joint is present. This is robust evidence for the hypothesis that inflammation leads to structural damage. There is a good relation between the damage in small and large joints as assessed on plain films. Damage measured in small joints is a good substitute for overall damage. Disease activity is always strongly correlated with functional disability throughout the disease course. There is an increasing relation between disability and structural damage with increasing disease duration.

Even for lay people it is crystal-clear that rheumatoid arthritis (RA) has an impact on functional outcome, just by looking at the hands of many patients with RA or seeing how they move around. Also the view of a radiograph of a hand with severe structural changes caused by RA gives the same impression to untrained eyes. So there is clear face validity that radiographic changes caused by RA have a relation with outcome. The question is how to define outcome. Many suggestions have been made. One well known suggestion is that of Fries et al with the five Ds: death, discomfort, disability, drug (therapeutic) toxicity, and dollar cost. In addition to this, people judge other aspects, also, as part of the outcome spectrum, such as the need for surgery, loss of work, and use of health resources. All these aspects of outcome were mentioned by participants of the OMERACT 5 conference on the relation between radiographic damage and outcome.

For the purpose of this paper outcome will be defined, mainly, as functional disability. The basic hypothesis is that disease activity in RA leads to loss of functional ability, and to structural damage, which can be visualised on radiographs. Structural damage in its turn also has a direct effect on functional ability (fig 1). This relationship between function and damage was the subject of a recent review by Scott et al. Only a few papers describe the relationship between functional disability and structural damage in longitudinal follow up studies. The general conclusion was that in early disease, this relationship is weak, but in advanced disease it is much stronger.

Why is it so difficult to prove a relationship when it seems so obvious that one must exist? An important reason might be in the way we measure the various aspects and how we try to link them together thereafter. For example, disease activity can be assessed in various ways, and as single or as composite variables. Important measures are the joint counts, which are assessed in both small and large joints. Other variables are pain and acute phase reactants which, however, may also be influenced by non-rheumatic causes. Moreover, often only measures of change such as the American College of Rheumatology (ACR) response criteria are used.
are used, especially in the analysis of therapeutic trials. These response criteria are not very useful in establishing a relationship between disease activity and radiographic damage as they give no information on the actual level of disease activity. They were developed to show good discriminative power between treatment groups but are by no means suitable to assess relationship between disease activity and function.

Functional disability is usually assessed with the Health Assessment Questionnaire (HAQ). Although this instrument works well for groups of patients, when means and medians are used, a different pattern emerges when individual patients are assessed. With stable medians over a five year follow up, individual patients show great variation over time. This is especially true in early disease. As in disease activity, disability is influenced by the involvement of both small and large joints. Moreover, both disease activity and structural damage have an impact on disability. But there are also other factors that have a major impact on function—for example, coping strategies, age, social class.

We want to determine the relationship between disease activity and function, on the one hand, with structural damage, on the other. Structural damage is usually assessed in small joints only. Most scoring methods are based on erosions, or on a combination of erosions and joint space narrowing. However, other structural damage may also have a role, such as the rupture of ligaments.

When one considers all these differences in the way we assess disease activity, function, and structural damage, the various influences on these aspects by factors not related to RA, and the simple statistics, such as correlations, often used to establish possible relationships, it is surprising that a relationship can be found at all. We present here some examples in which a clear relationship was found. This is by no means a complete and objective review of all available reported data, which can be found in the review by Scott et al., but it provides examples of recent findings that defend the hypothesis of a relation between disease activity, function, and structural damage, notwithstanding all the difficulties existing in this type of analyses. Many studies in the literature were examined which were unable to find such a relationship, which may partly be due to methodological issues.

**Relation between disease activity and radiographic damage in clinical trials**

Van Gestel et al compared the performance of the ACR 20 and EULAR response criteria in relation to radiographic progression in a randomised, double blind, parallel clinical trial. Patients were divided into three groups corresponding to a good, moderate, or no response according to the EULAR response criteria and into two groups corresponding to a response or no response according to the ACR criteria without taking the treatment group into account. There was a good correlation between the response status and the progression in Sharp score (fig 2).

A similar effect was shown in the COBRA trial, a randomised trial comparing a combination of step-down corticosteroids, methotrexate, and sulfasalazine (COBRA), with sulfasalazine alone. The median progression in Sharp score in the combination group was 2 (range 0–28), with an accompanying decrease in disease activity score (DAS) of 2.1. For the sulfasalazine group these figures were 4 (range 0–44) and 1.3, respectively. So a larger decrease in disease activity assessed by the DAS was reflected in a smaller progression in radiographic damage.

However, taking into account change only, as is done with ACR response criteria, may easily lead to misleading conclusions. This is illustrated in the ATTRACT trial, comparing placebo with infliximab infusions on background treatment of methotrexate. Comparing the radiographic progression in the control group with the infliximab treated patients showed a treatment effect in the group of ACR responders but also in that of ACR non-responders. This observation opened a lot of discussion about the dissociated effect on disease activity and radiographic progression. This is possible, but it might also be explained as due to a misinterpretation of the data.

This can be illustrated with further analyses of the COBRA trial. When the COBRA data were looked at in the same way, a similar effect was also seen—namely, a treatment effect present in the ACR responders but also in the ACR non-responders. However, when the actual disease activity during the entire period of the trial was examined in detail the following was shown (fig 3). In the group of ACR responders, the time integrated DAS for the COBRA treated patients was lower than that for the sulfasalazine treated patients. The same was true for the ACR non-responders. In conclusion, there was a difference in disease activity during the entire trial period in favour of the COBRA treated patients, which was in line with the lower progression in radiographic damage.
damage. This phenomenon was hidden by looking at the ACR response, which is a measure of change only.

In summary, the data show that there is a relation between the response to treatment measured as clinical disease activity and measured as radiographic progression in most clinical therapeutic trials.

**Relation between disease activity, structural damage, and function**

In patients with early disease there is a good correlation between several disease activity measures and function measured by the HAQ. Van Leeuwen et al reported a better correlation between painful joints and the HAQ (0.54) than between swollen joints and the HAQ (0.32) in an inception cohort of patients after three years’ follow up.11 The relation between HAQ and the increase in Sharp score was rather weak (0.31). The relation of swollen joints with the Sharp progression score was 0.47. Both correlations with the swollen joints were improved by using a weighted swollen joint count, in which larger joints get more weight. The correlation with the HAQ improved to 0.41 (instead of 0.32) and with the Sharp progression score to 0.57 (rather than 0.47).

Recently, 12 year follow up data from another inception cohort of female patients, aged between 20 and 50 at onset, was published.12 The relation between disease activity (assessed as DAS), HAQ, and Sharp score was assessed cross sectionally at baseline, after three, six, and 12 years. Throughout this follow up there was a strong relation between the DAS and HAQ, ranging from 0.51 to 0.68. On the other hand, there was an increasing relationship between the HAQ and the Sharp score from 0.22 at baseline to 0.57 at the 12 year follow up time. These data support the hypothesis presented in fig1.

In conclusion: disease activity is always strongly correlated with functional disability throughout the disease course. There is an increasing relation between disability and structural damage with increasing disease duration.

**Relation between damage in small and large joints**

Radiographic scoring methods which are applied in clinical trials use films of small joints only. The hands and wrists are always included, but more and more (and preferably) the feet are also part of the examination. However, to be a valid representation of the total damage caused by RA, there needs to be a strong relation between the damage in small and large joints. This was assessed by Scott et al in a cross sectional group of patients.13 They showed a good relation between the damage in hands and wrists and that in the other joints. More recently, similar data were published from an inception cohort of women followed up for 12 years, mentioned earlier.14 At the 12 year follow up time, 80% of the patients showed erosive changes in hands or feet. In 70% of the patients at least one large joint was abnormal, with 54% of the large joints showing erosive changes. If the results were limited to those patients with erosions in hands and feet, 69% of the patients had at least one erosive large joint. The overall correlation between damage in small and large joints was 0.76. An important finding was that none of the patients without erosions in hands or feet, had an erosive large joint.

In an earlier prospective follow up study a good relation was also found between small and large joints.15 There was also a relation between the damage in the small joints and number of joint replacements, which is another way of defining outcome.

These data lead to the following conclusions: there is a good relation between the damage in small and large joints. Therefore, damage assessed in small joints is a good substitute for overall damage.
Relation between inflammation and damage in individual joints

So far, all studies have compared the overall damage in all (small) joints with disease activity and function. However, an intriguing question to answer and important to support the hypothesis is whether there is a relation between inflammation and damage in individual joints. This analysis was again performed in the COBRA trial and presented at the ACR meeting by Boers et al.16 The conclusion of that study was that the presence at baseline of damage, swelling, and pain in a particular joint predicted, independently and strongly, progression of damage in that joint (p<0.001). Each swelling point (assessed on a scale of 0–2) tripled the risk, and each damage point (range 0–8) and pain point (range 0–3) doubled the risk. When the cumulative scores of pain and swelling over the entire trial period were used they predicted damage even more strongly. The prediction was present for both erosions and narrowing scores. Thus there is a strong relation between local inflammation and progression of damage in the individual joint. This is robust evidence for that part of the hypothesis which suggests that disease activity leads to structural damage.

In conclusion, more and more evidence is becoming available from clinical trials and cohort studies that there is a clear relation between inflammation (disease activity) and damage. Also the relation between structural damage and outcome, mainly assessed as physical function, is emerging from the data. This implies that structural damage, assessed on plain films, can indeed be used as a surrogate for outcome.