Is radiographic progression a realistic outcome measure in clinical trials with early inflammatory arthritis?

R. Landewé, D. van de Heijde

Department Internal Medicine/Rheumatology, University Hospital Maastricht, The Netherlands.

Robert Landewé, MD; Désirée van de Heijde, MD.

Please address correspondence to:
Robert Landewé, MD, University Hospital, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands.
E-mail: RLAN@SINT.AZM.NL

Clin Exp Rheumatol 2003; 21 (Suppl. 31): S37-S41.

Key words: Rheumatoid arthritis, radiography, early inflammatory arthritis, joint-damage, early inflammatory arthritis.

ABSTRACT

Radiographic progression is one of the most important outcome measures in rheumatoid arthritis (RA) clinical trials, because it reflects historic disease activity, is associated with loss of function over time, and can be reliably assessed. Trials involving patients with early inflammatory arthritis (EIA) will differ from those focusing on RA patients in many respects. They include a heterogeneous spectrum of patients, and some of them will have self-limiting arthritis, or arthritis with a low likelihood of ever becoming erosive. Furthermore, because of the early presentation a high proportion of patients with a high likelihood of erosions will still be non-erosive at presentation; and since EIA trials will aim at permanent clinical remission induced by therapy, the signs of progression will be very subtle.

Current radiographic scoring methods may not be sensitive to the small changes that are expected to occur in EIA trials. This makes radiographic progression a rather unlikely single primary outcome in such trials. However, "permanent clinical remission" (with or without therapy) appears to be a most realistic outcome in such trials, and radiographic stability (the demonstration of "no progression") may serve as a key criterion in establishing whether the endpoint of permanent clinical remission is actually met. The moment at which the first erosion develops is also important in making the correct diagnosis and has implications for the prognosis. We propose here a number of recommendations for the use of radiographic progression as one of the obligatory outcome measures in clinical trials with EIA.

Introduction

Radiographic progression is considered one of the most important outcome measures in rheumatoid arthritis (RA), because it assesses the potential of (new) anti-rheumatic drugs to prevent structural damage or to slow its occurrence.

Measuring structural damage is important for several reasons. First, many studies have now convincingly shown that the level of radiographically demonstrable damage of the joints is increasingly associated with a significant and irreversible loss of function once the disease proceeds (1-3). Therefore, in order to preserve function in RA patients, it seems rational to aim at drug treatment to minimise radiographic progression. Secondly, and perhaps even more importantly, radiographic damage is the result of chronic inflammation in the joints (3-6), which implies that radiographic damage reflects the level of "historic" disease activity in a particular patient. The relationship between disease activity and radiographic progression is far from absolute (both patients with high levels of disease activity without radiographic damage, and patients without measurable disease activity and steady progression are occasionally observed), but in general the demonstration of a slowing or arrest of radiographic progression during a time interval will imply that disease activity was low during this time interval. Third, radiographic progression has some properties that make it an attractive measurement instrument in clinical trials with RA patients: radiographic progression generally occurs slowly but gradually, without a high level of variation, or stays at a zero-level. The level of radiographic damage at a certain time point can be fairly precisely predicted from the previous recording, and progression can be reliably (reproducibly) measured, with an acceptable level of inter-reader variability and other types of error (7).

As stated in a number of articles earlier in this supplement, early inflammatory

arthritis (EIA) is not the same as RA. In the present article, we will elaborate on the question as to whether radiographic progression will become a useful outcome measure in clinical trials with EIA, as it is in RA. We will therefore compare different features of RA and EIA, judge the appropriateness of current scoring methods with respect to this goal, and outline how radiography can best be used in EIA.

Do the RA scoring methods allow the assessment of radiographic progression in EIA?

Several scoring methods are available, with which it is possible to assess the level of radiographic progression. In clinical trials, the Larsen scoring method (8) and even more frequently the (modified) Sharp scoring method (9) are used. These scoring methods assess the level of damage (erosions and joint space narrowing) per joint on X-rays of the hands and feet according to prespecified criteria (e.g. the Sharp method assigns an erosion score of 0 to 5, and a joint space narrowing score of 0 to 4 per small hand joint), but only in those joints that have proven to be susceptible to change, and that can be reliably assessed. An important and perhaps even decisive difference between RA and EIA is that patients with EIA often present with arthritis of only one or a few joints, that certainly will not always include the small joints of the hands or feet (10). The implication is that the possibility of identifying radiographic damage in EIA by routine X-rays of hands and feet will be lower compared to RA, which has many more joints involved. This could interfere with the statistical power needed to detect differences in radiographic progression in clinical treatment trials, which will be discussed further below.

In RA, radiographic progression in the hands and feet has been shown to appropriately reflect the progression of damage in the larger joints (11, 12), which are considered to be more important in determining physical function than the small joints of the hands and feet. This is why scoring X-rays of the hands and feet, a properly validated assessment technique, suffices in clinical trials with RA patients. It is,however, anything but obvious that these observations can be extended to EIA. Apart from an RA-like type of arthritis, with a relatively high tendency to show radiographic damage in the hands and feet, the spectrum of EIA may include all kinds of transient arthritides, undifferentiated chronic monoarthritides, as well as spondyloarthropathy-like types of oligoarthritis (often with involvement of the large joints of the lower extremities), that may have a far lower tendency to radiographic damage in the hands and feet. As a consequence, the conventional validated scoring methods for assessing radiographic damage in RA will pick up damage in an significantly lower proportion of patients in an EIA cohort.

One may consider the option of using x-rays of the affected (large) joints in order to score progression (the so-called "index joint"). There are however numerous arguments against this option. First, it takes more time to see changes in large joints compared to small joints, which makes it a less sensitive measure in early disease; second, scaling properties limit their usefulness; third, little information on the usefulness of an index joint to assess structural damage is available; and fourth, the low likelihood of radiographic damage in single joints other than those located in the hands and feet (except perhaps for the RA-like type of EIA) interferes with the sensitivity of picking up a signal at all, although it should be noted that the sensitivity of the accepted scoring methods primarily depends on the sum score of all the scored joints.

Patterns of radiographic progression in EIA

EIA cohorts have a number of characteristics that differentiate them from RA cohorts. First, as already mentioned, the spectrum of EIA may include a fairly high number of different presentations of arthritis, which makes an EIA cohort much more heterogeneous than an RA cohort. Second, the duration of complaints/signs will be much shorter in an EIA cohort. Third, the level of disease activity, measured by joint counts, levels of acute phase reactants and other parameters, will be lower than in "classical" RA cohorts (especially RA treatment cohorts) of clinical trials. These disparities will have important theoretical consequences for the assessment of radiographic progression as an outcome measure in clinical trials with EIA patients.

The heterogeneity of an EIA cohort implies that only a proportion of patients will present with an RA-like type of (poly)arthritis with an accompanying high level of inflammatory activity, *i.e.*, a patient cohort prone to radiographic progression. In the Leiden early arthritis cohort (described elsewhere in this issue), which included patients with arthritis in at least one joint, 60% of patients proved to have self-limiting disease (13). Only 15% of patients had erosions in hand and/or foot joints at baseline (10), and only a minority of the patients who did not have erosions at baseline developed them over a period of 2 years. The classical RA clinical trial with radiographic progression as the primary outcome includes patients with a high level of inflammatory activity, and often with an unfavourable prognosis, such as the presence of erosions at baseline and/or rheumatoid factor. The criterion of erosions – still the most important determinant of further damage – is included in the set of classification criteria for RA (14), so that every cohort of RA patients (observational or trial cohort), unlike a cohort with EIA patients, bears a natural likelihood for radiographic progression.

Many authors have emphasised the average linear course of radiographic progression in RA (15-17), with which they refer to a common causal background. It is unlikely that this observation can be extended to EIA. First, the average linear course in RA does not reflect the individual patient, for there may be a heterogeneity of patterns of progression forming a spectrum with "no progression" and "accelerating progression" at either extreme, and all patterns of acceleration and deceleration in between (18).

Most progression may occur within the first 2 years (19, 20), but occasionally RA patients develop erosions only

years after the first clinical presentation (21). Notwithstanding this heterogeneity, the concept of linearity of radiographic progression in RA can still be defended, by referring to a common pathophysiological background. The greater heterogeneity of arthritis presentations in an EIA cohort, however, makes it likely that radiographic progression will show a bimodal, rather than a more continuous distribution pattern (as in RA):Probably, the largest proportion of patients (>70%?) will not show progression at all, and will not have a tendency to become erosive over time. Therefore, radiographic progression will be a feature of a minority only. It will probably make no scientific sense (no face validity) to try and describe the average course of radiographic progression in the entire EIA cohort in one model (either linear or not).

There is a (theoretical) second reason to argue against the linearity of radiographic progression, both in RA and EIA cohorts. The observations of linearity of progression in RA were made in cohorts of patients treated with conventional disease modifying drugs (DMARDs). These drugs may slow radiographic progression, as demonstrated at the group level, but cannot stop it, and the average course of radiographic progression may be a gradually increasing one. Putatively, the introduction of TNF-blocking drugs has changed this picture: TNF-blocking drugs have now consistently shown that they are able to arrest further progression, both in early and in advanced RA (22, 23). The consequence will be that radiographic progression in anti-TNF treated patients will be different as compared to non-anti-TNF-treated patients, and the rationale to describe radiographic progression in averaged terms will be lost.

With these promising results in mind, future clinical trials with EIA patients will undoubtedly have the provocative aim to prevent the occurrence of erosions (or this may be the consequence of the treatment), even in patients with a high likelihood of erosive disease. This argument may further add to the validity of a binomial distribution pattern, rather than a continuous one, for radiographic progression in EIA (trial) cohorts.

One of the reasons that we outline this binomial distribution pattern here is that it may have consequences with respect to statistical power. Binomial distributions force a statistical test of comparing proportions, and such tests are inherently less powerful than tests for continuous variables in underscoring small between-group differences (e.g., a test for the comparison of mean changes in DAS in two groups is more powerful than a test to compare the proportion of responders based on the DAS in the two groups).

How sensitive are current scoring methods?

One may question – in view of the aforementioned arguments – whether current scoring methods are sensitive enough to pick up early changes in EIA. As already stated, damage will occur sporadically, rather than being dispersed as in RA, and a large proportion of patients will not show any damage at all. Assessing radiographic damage and progression is subject to various sources of measurement error, including inter-reader variability and error due to small differences in positioning and lighting. There may be a rather unfavourable signal-to-noise ratio in EIA, particularly because the signal is expected to be low, and the measurement noise cannot be reduced accordingly. Making use of the concept of the smallest detectable difference (SDD), we showed that a patient's score should change by 4 to 5 Sharp units before he is unequivocally (i.e., beyond measurement error) adjudicated as having progressive disease (24, 25). The latter may be necessary in order to allow a binomial analysis (progressors versus non-progressors), as argued previously. In trials using EIA patients and the usual follow-up period (1-2 yrs), it is not to be expected that many will show this level of progression unless they develop RA. On the other hand, and in contrast with expectations, we were able to demonstrate that experienced readers may pick up changes in radiographic progression in a cohort of early RA patients after only 3 months of follow-up, which adds to the high sensitivity of the current scoring methods (personal observation). Moreover, in EIA it might be of major importance to determine the moment that a patient becomes erosive, as this event may establish the diagnosis and prognosis. The determination of erosiveness should be performed with a number of erosions that is well below the SDD for progression of disease.

How important is radiographic progression in EIA?

We have argued thus far that the make-up of an EIA cohort and the features of the radiographic scoring methods may limit the application of radiographic progression as a single outcome measure. There is however no good reason to adopt the thesis that radiographic progression is less relevant in EIA patients with a propensity to develop joint damage than in RA patients. On the hypothesis that we are able to define (predict) which EIA patients will become erosive over time, the demonstration of "no progression" will be the ultimate proof of efficacy in this group of patients. The problem is that it is very difficult at presentation to predict future erosiveness.

Visser *et al.* proposed a prediction algorithm for patients presenting with EIA, which performed very well in their cohort of patients with EIA with respect to the prediction of erosions (13). Unfortunately however, cross-validation in other EIA cohorts showed less impressive results. Nonetheless, for group comparisons in clinical trials, such an algorithm will probably create subgroups with a certain likelihood to develop erosiveness.

Another matter that should be discussed here, because it emphasizes the importance of radiographic progression as an outcome of interest, is that EIA cohorts were initiated because of the realization that RA is often diagnosed too late (26). The most important argument was that radiographic damage was often already present at diagnosis, and proved to be the most important predictor of future damage despite treatment. The rationale behind EIA

cohorts (and trials) is the hypothesis that the occurrence of damage can be prevented by starting (aggressive) therapies at a time point at which no damage is apparent. A more provocative hypothesis is that the disease may even be cured before it really occurs (referred to as "the prevention of RA", or "from care to cure"). Rational outcome measures in such kinds of studies will be "permanent clinical remission" defined as the absence of disease activity in the absence of therapy. In such scenarios, it is of the greatest importance to use radiographs as the ultimate proof of clinical remission for several reasons. First, we know that a subgroup of RA patients may show progression despite a state of clinical remission (27). Second, the presence of clinical remission cannot be assessed continuously during a clinical trial,and clinical remission at some time point is not equivalent to the absence of disease activity over a time interval. We and others have shown in RA patients that the level of disease activity, but especially changes ("peaks") in disease activity, determines the rate of radiographic progression.

So, assuming that "absence of disease" becomes a relevant outcome in future trials with EIA, an important element in the ultimate proof will be the absence of structural damage ('non-erosive') or absence of radiographic progression (e.g. a radiographic progression less than the SDD).

Magnetic resonance imaging (MRI) and ultrasound (US)

Many retain MRI and US to be more sensitive imaging modalities than conventional radiography for the detection of early (structural) joint damage. This is probably correct, but sensitivity is at the cost of specificity, and it remains unclear whether all abnormalities that are detected earlier by MRI as compared to radiography are really relevant with respect to structural changes later on (construct validity). Other major problems today relate to its low feasibility and insufficient (inter-observer) reliability (28, 29).

MRI and US may become important in the diagnosis and follow-up of patients with EIA, since these modalities recognise structures other than bone alone. It is possible that the presence of synovitis will be followed once by MRI/US, but this is clearly beyond the scope of this article. At this time, radiographs remain sufficient to recognize the absence of complete remission, as noted above.

Summary of the pro's and con's

We have described how radiographic progression in EIA cohorts might differ from that in RA cohorts, especially RA trial cohorts. The proportion of patients with radiographic damage at baseline will be lower compared to RA cohorts, as will be the proportion of patients with progression, due to their earlier presentation and to a higher level of heterogeneity, with a mixture of patients without an intrinsic propensity to radiographic damage. As a consequence, radiographic progression may show a bimodal distribution pattern, rather than a continuous one as in RA. It is uncertain whether currently available scoring methods are sensitive enough to pick up the subtle changes in radiographic damage, and because of the expected bimodal distribution of progression scores, it does not seem rational to use radiographic progression as a primary endpoint in clinical trials of EIA.

On the other hand, EIA trials will also include a number of patients with a high likelihood of radiographic progression, at least in RA, but probably also in EIA. Patients may show progression of damage despite the absence of inflammatory activity, and results from early arthritis clinics have shown that the presence of erosions can be predicted to some degree. Since the aim of future trials in patients with EIA will undoubtedly be the absence of any signs of disease ("cure" in stead of "care"), radiography will be indispensable to provide the ultimate evidence that the disease is in permanent clinical remission.

Recommendations for future trials in EIA

Based on the above arguments, we now propose a number of recommendations with respect to the use of radiographic endpoints in clinical trials with EIA patients.

First, it seems unwise to use radiographic progression, as assessed on X-rays of the hands and feet, as the single primary outcome measure in EIA trials.

Second, patients included in any EIA trial should be stratified according to their likelihood of developing erosive changes over time. We urgently need validated algorithms for prediction, such as the one developed by Visser. Such a stratification would allow between group comparisons in patients with a high possibility of erosive damage, under the condition that the statistical power of such a stratum is sufficiently high (sufficient patient numbers) to allow an appropriate comparison.

Third, serial radiographs of the hands and feet should be taken in every patient, not to serve as a single outcome, but as a check that disease (activity) is really absent. These X-rays should be scored by one of the (preferably most sensitive) available scoring methods, and cut-offs for progression (yes versus no) should be based on methods for gauging measurement error (e.g. SDD). We do not recommend single joint scores (index joints), because of the lack of proof of validity (scaling, sensitivity to change, reproducibility).

Fourth, radiographs can be used to assess the occurrence of the first erosion. This moment is important in making the correct diagnosis and has implications for the prognosis.

References

1. DROSSAERS-BAKKER KW, DE BUCK M, VAN ZEBEN D, ZWINDERMAN AH, BREEDVELD FC, HAZES JM: Long-term course and outcome of functional capacity in rheumatoid arthritis: The effect of disease activity and radiologic damage over time. *Arthritis Rheum* 1999; 42: 1854-60.
2. SCOTT DL, SYMMONS DP, COULTON BL, POPERT AJ: Long-term outcome of treating rheumatoid arthritis: Results after 20 years. *Lancet* 1987;1:1108-11.
3. WELSING PM, VAN GESTEL AM, SWINKELS HL, *et al.*: The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. Modified sharp method: factors influencing reproducibility and variability. *Arthritis Rheum* 2001; 44: 2009-17.
4. PLANT MJ, WILLIAMS AL, O'SULLIVAN MM, LEWIS PA, COLES EC, JESSOP JD: Rela-

tionship between time-integrated C-reactive protein levels and radiologic progression in patients with rheumatoid arthritis. *Arthritis Rheum* 2000; 43: 1473-7.

5. VAN LEEUWEN MA, VAN RIJSWIJK MH, VAN DER HEIJDE DM, *et al*.: The acute-phase response in relation to radiographic progression in early rheumatoid arthritis: a prospective study during the first three years of the disease. *Br J Rheumatol* 1993; 32: 9-13.
6. VAN LEEUWEN MA, VAN RIJSWIJK MH, SLUITER WJ, *et al*.: Individual relationship between progression of radiological damage and the acute phase response in early rheumatoid arthritis. Towards development of a decision support system. *J Rheumatol* 1997; 24: 20-27.
7. LASSERE MN, VAN DER HEIJDE D, JOHNSON KR, BOERS M, EDMONDS J: Reliability of measures of disease activity and disease damage in rheumatoid arthritis: implications for smallest detectable difference, minimal clinically important difference, and analysis of treatment effects in randomized controlled trials. *J Rheumatol* 2001; 28: 892-903.
8. LARSEN A, DALE K, EEK M: Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. *Acta Radiol Diagn Stockh* 1977; 18: 481-91.
9. SHARP JT, YOUNG DY, BLUHM GB, *et al*.: How many joints in the hands and wrists should be included in a score of radiologic abnormalities used to assess rheumatoid arthritis ? *Arthritis Rheum* 1985; 28:1326-35.
10. VAN DER HORST-BRUINSMA IE, SPEYER I, VISSER H, BREEDVELD FC, HAZES JM: Diagnosis and course of early-onset arthritis: results of a special early arthritis clinic compared to routine patient care. *Br J Rheumatol* 1998; 37:1084-8.
11. SCOTT DL, COULTON BL, POPERT AJ: Long term progression of joint damage in rheumatoid arthritis. *Ann Rheum Dis* 1986; 45: 373-8.
12. DROSSAERS-BAKKER K, KROON H, ZWINDERMAN A, BREEDVELD F, HAZES J: Radiographic damage of large joints in long-term rheumatoid arthritis and its relation to function. *Rheumatology* 2000; 39: 998-1003.
13. VISSER H, LE CESSIE S, VOS K, BREEDVELD FC,HAZES JM: How to diagnose rheumatoid arthritis early: a prediction model for persistent (erosive) arthritis. *Arthritis Rheum* 2002; 46: 357-65.
14. ARNETT FC, EDWORTHY SM, BLOCH DA, *et al*.: The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988; 31: 315-24.
15. FEX E, JONSSON K, JOHNSON U, EBERHARDT K: Development of radiographic damage during the first 5-6 yr of rheumatoid arthritis. A prospective follow-up study of a Swedish cohort. *Br J Rheumatol* 1996; 35: 1106-15.
16. HULSMANS HM, JACOBS JW, VAN DER HEIJDE DM, VAN ALBADA KUIPERS GA, SCHENK Y, BIJLSMA JW: The course of radiologic damage during the first six years of rheumatoid arthritis. *Arthritis Rheum* 2000; 43: 1927-40.
17. DROSSAERS-BAKKER KW, AMESZ E, ZWINDERMAN AH, BREEDVELD FC, HAZES JM: A comparison of three radiologic scoring systems for the long-term assessment of rheumatoid arthritis: Findings of an ongoing prospective inception cohort study of 132 women followed up for a median of twelve years. *Arthritis Rheum* 2000; 43: 1465-72.
18. PLANT MJ, JONES PW, SAKLATVALA J, OLLIER WER, DAWES PT: Patterns of radiological progression in early rheumatoid arthritis: Results of an 8 year prospective study. *J Rheumatol* 1998; 25: 417-426.
19. PINCUS T, FUCHS HA, CALLAHAN LF, NANCE EP JR, KAYE JJ: Early radiographic joint space narrowing and erosion and later malalignment in rheumatoid arthritis: A longitudinal analysis. *J Rheumatol* 1998; 25: 636-40.
20. VAN DER HEIJDE DM, VAN LEEUWEN MA, VAN RIEL PL *et al*.: Biannual radiographic assessments of hands and feet in a three-year prospective followup of patients with early rheumatoid arthritis. *Arthritis Rheum* 1992; 35: 26-34.
21. BUKHARI M,HARRISON B, LUNT M,SCOTT DG, SYMMONS DP, SILMAN AJ: Time to first occurrence of erosions in inflammatory polyarthritis: results from a prospective community-based study. *Arthritis Rheum* 2001; 44: 1248-53.
22. BATHON JM, MARTIN RW, FLEISCHMANN RM *et al*.: A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. *N Engl J Med* 2000; 343:1586-93.
23. LIPSKY PE, VAN DER HEIJDE DM, ST CLAIR EW, *et al*.: Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. *N Engl J Med* 2000; 343: 1594-602.
24. BRUYNESTEYN K, VAN DER HEIJDE D, BOERS M *et al*.:Minimal clinically important difference in radiological progression of joint damage over 1 year in rheumatoid arthritis: preliminary results of a validation study with clinical experts. *J Rheumatol* 2001; 28: 904-10.
25. LASSERE MN, VAN DER HEIJDE D, JOHNSON K *et al*.: Robustness and generalizability of smallest detectable difference in radiological progression. *J Rheumatol* 2001; 28: 911-3.
26. HUIZINGA TW, MACHOLD KP, BREEDVELD FC, LIPSKY PE,SMOLENJS: Criteria for early rheumatoid arthritis: from Bayes' law revisited to new thoughts on pathogenesis. *Arthritis Rheum* 2002; 46: 1155-9.
27. MOLENAAR ET, VOSKUYL AE, DIJKMANS BA: Functional disability in relation to radiological damage and disease activity in patients with rheumatoid arthritis in remission. *J Rheumatol* 2002; 29: 267-70.
28. GOLDBACH-MANSKY R, WOODBURN J, YAO L, LIPSKY PE: Magnetic resonance imaging in the evaluation of bone damage in rheumatoid arthritis:A more precise image or just a more expensive one ? *Arthritis Rheum* 2003; 48: 585-9.
29. WAKEFIELD RJ, GIBBON WW, EMERY P: The current status of ultrasonography in rheumatology. *Rheumatology (Oxford)* 1999; 38: 195-8.