Radiographic Scoring in Rheumatoid Arthritis
A Short Introduction to the Methods

Tuulikki Sokka, M.D., Ph.D.

Abstract
Several methods have been introduced for the scoring of plain radiographs in patients with rheumatoid arthritis. The first methods characterized global damage in hands and wrists in a qualitative rather than quantitative approach. The Sharp and the Larsen scoring methods were introduced in the 1970s. Modifications of these methods over time have resulted in contemporary scoring systems that are described herein.

Radiographs provide a permanent measure of damage in rheumatoid arthritis (RA). Plain radiographs of hands and feet have been important in the evaluation of the RA disease course and its possible modification over the last 60 years. The long-term severity of RA was recognized in longitudinal studies of clinical cohorts that indicated continuous radiographic progression over follow-ups in excess of 20 years. The true efficacy of disease-modifying antirheumatic drugs (DMARDs) traditionally has been viewed as their capacity to slow down radiographic damage. That being said, disease activity measures that may indicate short-term levels of change and differences between active and control treatment are the primary outcome measures in most clinical trials. Furthermore, documentation of milder radiographic progression of RA at this time, compared to previous decades, indicates improved outcomes of RA. This is at least, in part, due to improved treatment strategies and available DMARDs and biological agents.

Modern technologies, including magnetic resonance imaging (MRI) and ultrasound, are valuable and more sensitive than radiographs in detecting early structural changes in joints and surrounding structures. However, availability and costs may limit the use of these technologies in daily clinical practice. Clinical trials rely on plain radiographs instead of other imaging technology and present radiographic differences between the active drug and comparator that are statistically highly significant but indicate minor absolute differences in both groups. Therefore, it remains important for a clinical rheumatologist to understand the scoring of plain radiographs and the history of the scoring methods, as described in this article.

Selected Methods to Score Radiographs
Initially, radiographs were scored using the Steinbrocker method, with a global damage score to hands and wrists on a four-point scale from I (minimal damage) to IV (severe damage). The grade was determined by the worst change in any joint and, therefore, the score was biased toward the most severely affected joint. Furthermore, the narrow scale of grades only from I to IV was not sufficiently detailed to assess change in status. The Kellgren method was similar to the Steinbrocker method: a global grade was given as the summation of abnormalities for all the joints in both hands and wrists.

The two most widely used measures of radiographs are based on the work of Sharp and Larsen that provide a continuous quantitative scale of more than 100 units, rather than a limited qualitative measure of radiographic damage. The Sharp method involves separate scores for erosions and joint space narrowing, while the Larsen method is based on a global score of each joint.

The Sharp Method
The Sharp method originally included radiographs of hands and wrists and counted several features, such as periosteal reaction, cortical thinning, osteoporosis, sclerosis, osteophyte...
radiographs. Initially, articular osteoporosis and soft tissue swelling were included in the scoring but were omitted later due to technical reasons. Furthermore, the score of wrists was suggested to be multiplied by five due to technical reasons. Furthermore, the score of wrists was suggested to be multiplied by five or divided in four sections, which would each be scored, but these modifications were dropped later (Larsen, personal communication).

In addition to reference films, Larsen has introduced the radiographic score ranges from 0 to 448. The minimal clinically important difference (MCID) is roughly 1% of the maximum for both of the methods. Overall, the Larsen method is more easily scored and less time-consuming compared to the Sharp method.

### Limitations of Radiographs

Although radiographs provide optimal documentation of joint destruction, several important limitations are seen in the application of radiographs as a measure of clinical status in patients with RA. Radiographs change slowly in most people with RA. At least 6 months to a year might be required to assess changes in an individual patient, although changes may be detectable over months in large groups of patients. Modern treatment of RA requires that many patients be treated prior to radiographic damage. Radiographs correlate significantly with radiographic score ranges from 0 to 448. The minimal clinically important difference (MCID) is roughly 1% of the maximum for both of the methods. Overall, the Larsen method is more easily scored and less time-consuming compared to the Sharp method.

### Limitations of Radiographs

Although radiographs provide optimal documentation of joint destruction, several important limitations are seen in the application of radiographs as a measure of clinical status in patients with RA. Radiographs change slowly in most people with RA. At least 6 months to a year might be required to assess changes in an individual patient, although changes may be detectable over months in large groups of patients. Modern treatment of RA requires that many patients be treated prior to radiographic damage. Radiographs correlate significantly with radiographic score ranges from 0 to 448. The minimal clinically important difference (MCID) is roughly 1% of the maximum for both of the methods. Overall, the Larsen method is more easily scored and less time-consuming compared to the Sharp method.

### The Larsen Method

The clinical observation of a male running to a bus who had RA and a maximum Steinbrocker damage score of 4 was described by Larsen as the initial basis for a more detailed scoring method. The Larsen method includes both erosions and joint space narrowing in each joint as a single score, on a scale of 0 to 5 according to reference radiographs. Initially, articular osteoporosis and soft tissue swelling were included in the scoring but were omitted later due to technical reasons. Furthermore, the score of wrists was suggested to be multiplied by five or divided in four sections, which would each be scored, but these modifications were dropped later (Larsen, personal communication). In addition to reference films, Larsen has introduced the following guidelines for scoring:

- **0** = Intact bony outlines and normal joint space;
- **1** = Erosion less than 1 mm in diameter or joint space narrowing;
- **2** = One or several small erosions, diameter more than 1 mm;
- **3** = Marked erosions;
- **4** = Severe erosions, where there is usually no joint space left, and the original bony outlines are partly preserved; and
- **5** = Mutilating changes, where the original bony outlines have been destroyed.

All synovial joints can be included in the Larsen score, and joints that are scored should, therefore, be listed, as well as the maximum score applied. Kaarela and Kautiainen suggested including 10 metacarpophalangeal joints and wrists, and the second to the fifth MTP joints in the scoring, with a range of 0 to 100. Rau and Herborn introduced a modified Larsen score that includes a quantitative estimate of the percentage of loss of the joint surface and is known as a Ratingen score.

The Sharp and Larsen scores are correlated significantly. The smallest detectable difference (SDD) is the smallest change that can be reliably discriminated from the measurement error of the scoring method, which is 5.0 for the van der Heijde modified Sharp score (0 to 448) and 5.8 for the Larsen score (0 to 200). The minimal clinically important difference (MCID) is roughly 1% of the maximum for both of the methods. Overall, the Larsen method is more easily scored and less time-consuming compared to the Sharp method.

### Limitations of Radiographs

Although radiographs provide optimal documentation of joint destruction, several important limitations are seen in the application of radiographs as a measure of clinical status in patients with RA. Radiographs change slowly in most people with RA. At least 6 months to a year might be required to assess changes in an individual patient, although changes may be detectable over months in large groups of patients. Modern treatment of RA requires that many patients be treated prior to radiographic damage.

Radiographs correlate significantly with radiographic score ranges from 0 to 448. The minimal clinically important difference (MCID) is roughly 1% of the maximum for both of the methods. Overall, the Larsen method is more easily scored and less time-consuming compared to the Sharp method.

### Limitations of Radiographs

Although radiographs provide optimal documentation of joint destruction, several important limitations are seen in the application of radiographs as a measure of clinical status in patients with RA. Radiographs change slowly in most people with RA. At least 6 months to a year might be required to assess changes in an individual patient, although changes may be detectable over months in large groups of patients. Modern treatment of RA requires that many patients be treated prior to radiographic damage.
and clinical observational studies requires basic knowledge of the scoring methods of radiographs.

**Disclosure Statement**
The author has no financial or proprietary interest in the subject matter or materials discussed, including, but not limited to, employment, consultancies, stock ownership, honoraria, and paid expert testimony.

**References**