Damage and Progression on Radiographs in Individual Joints: Data from Pivotal Randomized Controlled Trials

ROBERT B.M. LANDEWÉ, VIBEKE STRAND, PHILIP G. CONAGHAN, and DÉSIRÉE van der HEIJDE

ABSTRACT. *Objective.* Radiographic progression is usually assessed by Sharp-based methods (van der Heijde-modified Sharp score and the Genant-modified Sharp score). The aim of this study was to evaluate, in a range of randomized controlled trials (RCT), the presence of erosions and joint space narrowing (JSN) in all individual joints, as well as progression in these joints, and to determine if any redundancy exists due to infrequently involved joints.
Methods. Four databases of rheumatoid arthritis RCT that were all scored according to van der Heijde's modification of the Sharp score were included in a descriptive analysis.
Results. Irrespective of different readers, different patient populations, and different disease durations per trial, similar patterns emerged. Both erosions and JSN occurred in all sites. Erosions occurred most frequently in the feet, preferentially in 5th metatarsophalangeal joint (MTP-5). JSN occurred most frequently in the wrist. Change from baseline in erosions and JSN followed the pattern of involvement at baseline, so that MTP-5, and to a lesser extent MTP-3 and MTP-4, preferentially showed progression in erosive damage. Joints in the wrist showed highest tendency to worsen over time with respect to JSN.
Conclusion. These data indicate that both erosions and JSN must be assessed for damage, and that a more abbreviated joint count cannot be used for radiographic scoring. (J Rheumatol 2011;38:2018–22; doi:10.3899/jrheum.110417)

Key Indexing Terms:
RHEUMATOID ARTHRITIS   RADIOGRAPHY   DISEASE PROGRESSION
EROSION   JOINT SPACE NARROWING

Radiographic damage is an important outcome in rheumatoid arthritis (RA)[1], because it is the consequence of joint inflammation[2] and is independently associated with irreversible impairment of physical function[3]. In order to grant a claim of prevention of structural damage, regulatory authorities require data proving that a particular medicine can retard the progression of radiographic damage during a 6 to 12-month period. The main reason that radiographic progression is measured in hands and feet is 2-fold: (1) RA classically involves the small joints of hands, wrists, and feet; and (2) radiographs of hands and feet cover the highest number of involved joints using the lowest number of radiographs. At the group level, radiographic damage in hand and feet joints associates acceptably with damage in large weight-bearing joints with a dominant influence on physical function[4].

In order to measure radiographic progression in clinical trials, several instruments have been developed. These instruments can largely be divided into Larsen-based methods and Sharp-based methods. Larsen proposed his method in 1977[5], and modifications were subsequently proposed by Scott[6] and by Rau[7]. Sharp proposed his Sharp score in 1971, with modifications proposed by himself[8], van der Heijde (van der Heijde-modified Sharp score)[9], and Genant (Genant-modified Sharp score)[10].

The main difference between Larsen-based and Sharp-based methods is that Larsen applies a global grading of joint damage, while Sharp quantifies number and size of erosions per joint and amount of joint space narrowing (JSN). The measurement of radiographic progression in the context of a short-term randomized controlled trial (RCT) requires that a method have a high level of reliability and sensitivity to change, and Sharp-based methods have been shown to outperform Larsen-based methods, in that they have a better "signal-to-noise ratio"[11]. As a consequence, radiographic progression in modern clinical trials is usually assessed by Sharp-based methods, mostly the van der Heijde-modified Sharp score and the Genant-modified Sharp score. While both methods have been successful in demonstrating differences in radiographic progression in separate trial arms, some concerns remain. Scoring is time-consuming, especially in the case of significant damage or a high level of progression.

From the Department of Internal Medicine/Rheumatology, Maastricht University Medical Center and Caphri Research School, Maastricht, The Netherlands; Division of Immunology/Rheumatology, Stanford University School of Medicine, Palo Alto, CA, USA; Section of Musculoskeletal Disease, University of Leeds and NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds, UK; and Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands.

R.B.M. Landewé, MD, PhD, Department of Internal Medicine/ Rheumatology, Maastricht University Medical Center and Caphri Research School (currently, Department of Clinical Immunology and Rheumatology, Academic Medical Center, University of Amsterdam, and Atrium Medical Center, Heerlen, The Netherlands; V. Strand, MD, Division of Immunology/Rheumatology, Stanford University School of Medicine, Palo Alto, CA, USA; P.G. Conaghan, MBBS, PhD, FRACP, FRCP, Section of Musculoskeletal Disease, University of Leeds and NIHR Leeds Musculoskeletal Biomedical Research Unit; D. van der Heijde MD, PhD, Department of Rheumatology, Leiden University Medical Center.

Address correspondence to Dr. Landewé; E-mail: landewe@rlandewe.nl

Measurement error, reflected by disagreement among independent readers, jeopardizes precision and may decrease the signal-to-noise ratio, which in turn leads to impaired discrimination. Theoretically, time to score could be decreased and precision could be increased by deleting those joints that neither contribute to damage nor to progression in RCT, by deleting joints that inappropriately add to measurement variation, or by leaving out either the domain of erosions or the domain of joint space narrowing.

The aim of this study was to evaluate and compare the frequency of both involvement and progression of radiographic erosions, and JSN in individual joints. By including various datasets with populations with different levels of severity, generalizability of results could be tested.

MATERIALS AND METHODS

Radiographic data from baseline and 12-month timepoints from 4 RA RCT that employed van der Heijde-modified Sharp scoring were used for this analysis after permission was kindly granted by the sponsors. The following trial databases were included: (1) The Combination of Methotrexate (MTX) and Etanercept in Active Early Rheumatoid Arthritis (COMET) trial[12], which included patients with early RA naive to MTX with a disease duration < 2 years (mean duration 0.8 yrs). (2) The Trial of Etanercept and Methotrexate with Radiographic Patient Outcomes (TEMPO)[13], which included patients with established RA, the majority being MTX-naive and whose mean disease duration was 6.7 years at inclusion. (3) The RA Prevention of Structural Damage-1 (RAPID-1) trial[14]. This trial, comparing certolizumab plus MTX and MTX monotherapy, included patients with active, established RA despite a stable dose of MTX who had mean disease duration at baseline of 6.2 years. (4) The RA Prevention of Structural Damage-2 (RAPID-2) trial[15], comparing certolizumab plus MTX and MTX monotherapy, and including patients with active RA despite MTX. The mean disease duration at baseline was 6.2 years. All trials were scored by panels of different readers.

Analysis. The percentage of patients with involvement by erosions or JSN per joint was extracted from each database. Data were extracted per reader. Some databases (COMET, TEMPO) allowed a distinction between left and right sides while for others (RAPID-1 and RAPID-2) left and right joints were pooled. Per-joint data were extracted for baseline only (containing data of all treatment arms) and for change between baseline and endpoint. For the latter analysis, use was made of the treatment arm that did not contain tumor necrosis factor (TNF)-blocking treatment (usually the MTX monotherapy arm) because under TNF-blocking therapy, the normal relationship between disease activity and radiographic progression is lost[16] and radiographic progression is almost completely blocked.

In order to visualize erosion and JSN involvement on a per-joint basis, histograms were plotted for all 4 trials. These histograms express all individual joints on the x-axis and the percentage of patients with involvement on the y-axis. Joint patterns of involvement were clustered according to 4 subgroups: proximal interphalangeal (PIP) joints, metacarpophalangeal (MCP) joints, wrist joints, and metatarsophalangeal joints.

RESULTS

In general, all different readers showed similar patterns across the trials, both with respect to joint involvement at baseline, and with respect to change from baseline in the MTX monotherapy arms.

The RAPID-2 trial was scored by 3 instead of 2 readers as in the other trials, with each read overlapping one-third of the radiographic reading; this made it impossible to compare the

performances of these readers within the trial, but the general pattern that emerged was entirely similar. Since RAPID-1 and RAPID-2 were very similar trials in design, we chose to include only RAPID-1 data here. For reasons of brevity, we chose to show the data of only the COMET trial in this article as a representative example while mentioning essential data of the other trials in the text.

Figure 1 shows the percentage of involvement of joints with erosions at baseline (entire trial population). In general, involvement of the left side (dark bars) is similar to involvement of the right side (light bars), although individual joints show some variation. Foot joints were most frequently involved (14%–16% of all patients had MTP-5 involvement), followed by MCP joints, wrist joints, and PIP joints. Within the clusters and across all trials, MTP-5 was most frequently involved at baseline, followed by other MTP joints, although at different rates. For example, MTP-5 involvement was far higher in RAPID-1 (37%) and TEMPO (45%) compared to COMET (18%), reflecting the far longer disease duration in the former 2 trials.

If one looks on a per-cluster basis at the 3 trials, of the MCP cluster, MCP-2 was most frequently involved [11% in COMET (Figure 1), 30% in RAPID-1, and 28% in TEMPO]. Of the wrist cluster and the PIP cluster a dominant pattern of joint involvement could not be identified. Of note, none of the joints that were scored at baseline had involvement in zero percent of the patients, neither in the COMET trial nor in the RAPID-1 trial and the TEMPO trial.

Figure 2 (COMET) shows the percentage of joints with change in erosions from baseline. Change was predominantly seen in the foot cluster, and to a lesser extent in the MCP cluster. In line with patterns of involvement at baseline, change in the COMET trial was most frequently seen in MTP-3 to MTP-5 followed by MCP-2 to MCP-3 (Figure 2). But the picture in the RAPID-1 and TEMPO trial was slightly different: MTP-5 was scored as changed in approximately 10% of the patients in the COMET trial (Figure 2), but in only 7% of the TEMPO trial, and a little bit more than 1% of the RAPID-1 trial. Although change was rare in some individual joints, we could not find joints in any trial in which change in erosions was never detected.

Figure 3 (COMET) shows the percentage of joints with JSN at baseline per joint and per cluster. Here, the cluster of the wrist was primarily affected, followed by foot joints, MTP joints, and PIP joints. This pattern was more or less recognizable in the other trials, although the most frequently involved joint in the wrist cluster (and in other clusters) differed across trials. In the COMET trial (Figure 3), the joint between os multangulum and os naviculare was most frequently involved (approximately 10% of the patients). The radiocarpal joint was most frequently affected in the RAPID-1 trial (23%) and in the TEMPO trial (40%). Importantly, JSN was occasionally found in all possible sites, although differences in the rate of involvement were found. The picture of change per cluster
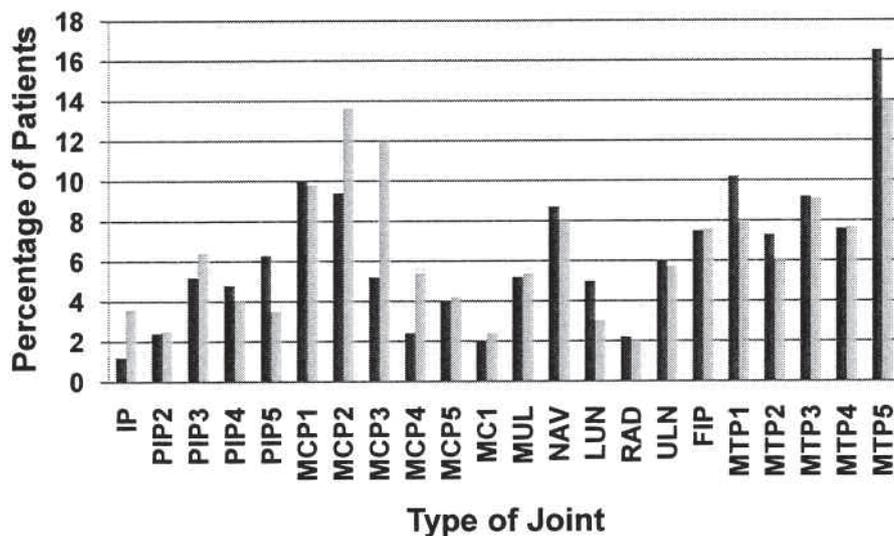
Figure 1. Percentage of patients with erosions on a per-joint level at baseline of the COMET trial (n = 476). All patients irrespective of treatment allocation are included in the analysis. Dark bars: joints of the left side of the body; light bars: joints of the right side. IP: interphalangeal joint; PIP (2, 5): proximal interphalangeal joint of 2nd, 5th ray; MCP (2, 5): metacarpophalangeal joint of 2nd, 5th ray; MC1: metacarpophalangeal joint of 1st ray; MUL: os multangulum; NAV: os naviculare; LUN: os lunatum; RAD: radius; ULN: ulna, FIP: first interphalangeal joint of the foot; MTP (1, 5): metatarsophalangeal joint of 1st, 5th ray.
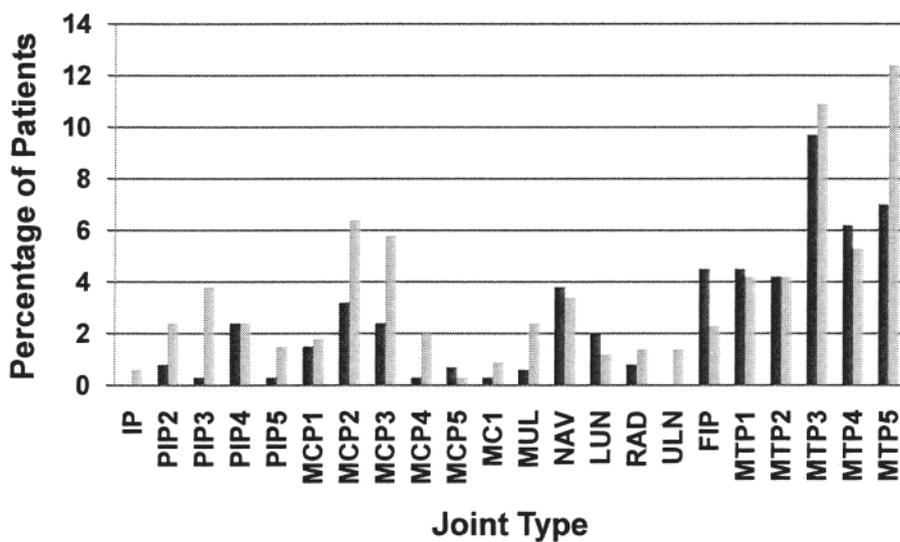

Figure 2. Percentage of patients with change from baseline in erosions on a per-joint level in the methotrexate-only arm of the COMET trial (n = 135). Dark bars: joints of the left side of the body; light bars: joints of the right side. IP: interphalangeal joint; PIP (2, 5): proximal interphalangeal joint of 2nd, 5th ray; MCP (2, 5): metacarpophalangeal joint of 2nd, 5th ray; MC1: metacarpophalangeal joint of 1st ray; MUL: os multangulum; NAV: os naviculare; LUN: os lunatum; RAD: radius; ULN: ulna; FIP: first interphalangeal joint of the foot; MTP (1, 5): metatarsophalangeal joint of 1st, 5th ray.

(Figure 4) followed the picture of involvement at baseline, in that the wrist sites most frequently showed change in the COMET trial (up to 5% in the joint between the os capitatum and the os naviculare) and in the RAPID-1 trial (up to 1.5% in the joint between the os capitatum and the os naviculare). In the TEMPO trial, change in JSN was seen slightly more frequently in MCP-2 (approximately 4% of the patients).

DISCUSSION

The results of this analysis showed a remarkably consistent picture that can be summarized as follows: irrespective of different readers, different patient populations, and different disease durations in each trial, similar patterns emerged. Both erosions and JSN occur in all sites that are scored according to van der Heijde's modification of the Sharp score. Erosions
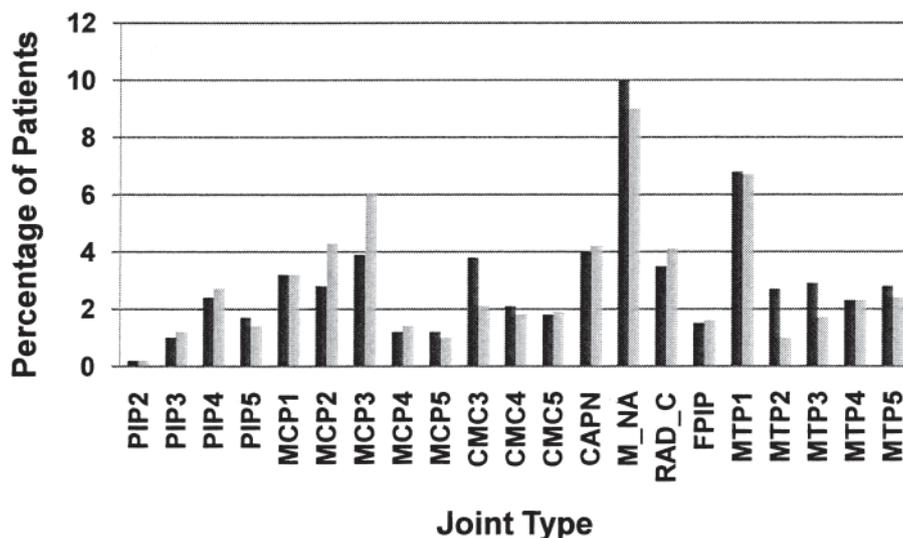
Figure 3. Percentage of patients with erosions on a per-joint level at baseline of the COMET trial (n = 476). Dark bars: joints of the left side of the body; light bars: joints of the right side. PIP (2, 5): proximal interphalangeal joint of the 2nd, 5th ray; MCP (2, 5): metacarpophalangeal joint of the 2nd, 5th ray; CMC (3, 5): carpometacarpophalangeal joint of the 3rd, 5th ray; CAPN: joint between os capitatum and os naviculare; M_NA: joint between os multangulum and os naviculare; RAD_C: joint between radius and os capitatum; FPIP: first interphalangeal joint of the foot; MTP (1, 5): metatarsophalangeal joint of the 1st, 5th ray.
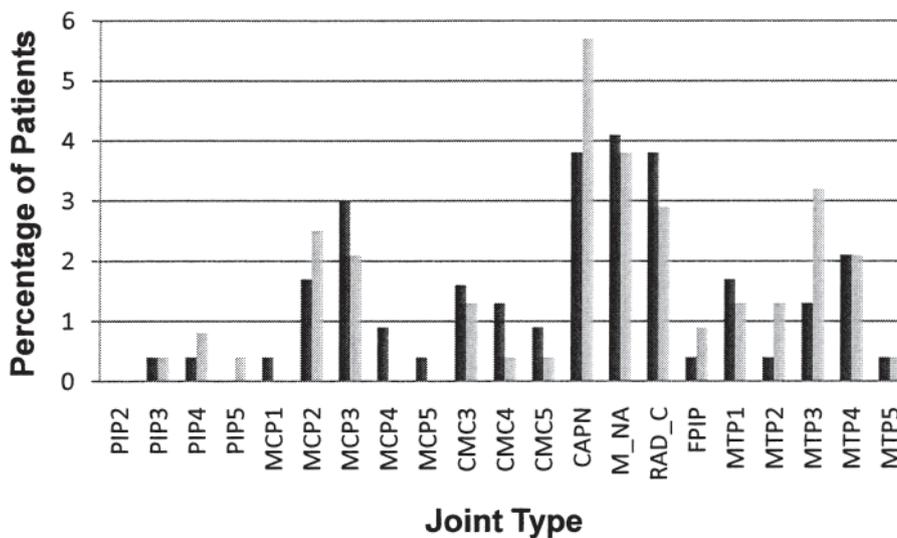

Figure 4. Percentage of patients with change from baseline in joint space narrowing on a per-joint level in the methotrexate-only arm of the COMET trial (n = 135). Dark bars: joints of the left side of the body; light bars: joints of the right side. PIP (2, 5): proximal interphalangeal joint of the 2nd, 5th ray; MCP (2, 5): metacarpophalangeal joint of the 2nd, 5th ray; CMC (3, 5): carpometacarpophalangeal joint of the 3rd, 5th ray; CAPN: joint between os capitatum and os naviculare; M_NA: joint between os multangulum and os naviculare; RAD_C: joint between radius and os capitatum. FPIP: first interphalangeal joint of the foot; MTP (1, 5): metatarsophalangeal joint of the 1st, 5th ray.

occur most frequently in the feet, and MTP-5 is preferentially involved. JSN occurs most frequently in the wrist. In general, changes in erosions and JSN follow the pattern of involvement at baseline, so that MTP-5 and to a lesser extent MTP-3 and MTP-4 preferentially show progression in erosive damage. Similarly, the joints in the wrist that are preferentially affected at baseline show the highest tendency to worsen over time. Such a pattern of individual joint involvement has been described with

the same joint preferences in an observational cohort setting[17].

The main aim of this study was to determine if particular joints could be excluded from scoring due to lack of disease involvement and/or limited progression. In this analysis we could not find joints that were consistently negative in terms of involvement in damage and/or change over time. We had considered the hypothesis that erosions rarely, if ever, are seen in the PIP joints, and that change in erosions was also rare in

these joints. Although involvement and change in erosions was definitely less prevalent in the PIP joints in comparison to other sites, all trials showed change in these joints in a proportion of patients. It was also hypothesized that the carpometacarpal joints of the 3rd, the 4th, and the 5th ray, which are considered joints belonging to the wrist area, rarely if ever show change. However, we found that these joints were also contributing to change in JSN in all trials studied.

Although some joints contributed to change in less than 2% of a particular trial population, it is important to realize that the total Sharp score is a sum score of all separate joint erosion scores and JSN scores, and that the change in total Sharp score in modern RCT with effective treatments rarely, if ever, exceeds 2 units per year. That means that every joint that shows change contributes importantly to the individual total Sharp score change. It is to be expected that future clinical trials, in which a new standard of care in the control group may imply an even lower progression rate, will show less contrast between treatment arms, which poses a greater challenge to the scoring methods in terms of discriminative power. A limitation of this study obviously is that we have not investigated to what extent the removal of particular joints would influence performance of the score in terms of sensitivity to change and discrimination.

An interesting observation in this descriptive study was that change in erosions and JSN was highest in the COMET trial with early DMARD-naive patients, whereas it was much lower in the RAPID trials and in the TEMPO trial, which included patients with more longstanding disease, despite a far higher rate of involvement at baseline. Although the general statement that the pattern of joint change follows the pattern of joint involvement is true, our results seem to confirm that the overall rate of progression per joint in early DMARD-naive RA is higher than in more advanced RA, with the proviso that an appropriate comparison between trials read by different readers is limited. Translated to the level of the individual patient, this means a confirmation of the parabolic configuration of the radiographic damage-over-time curve, with steepest progression in the early phase and leveling off thereafter. This consistent finding has always been explained by a truly higher progression score in the early disease[18].

In summary, we have shown here in different RCT databases, using individual joint scores for erosion and JSN, that all scored joints contribute to some extent in quantifying change over time. In light of the need to detect subtle changes in future clinical trials in RA, we do not recommend reducing the current joints scored by modified Sharp methodologies.

REFERENCES

1. Strand V, Sharp JT. Radiographic data from recent randomized controlled trials in rheumatoid arthritis: What have we learned? Arthritis Rheum 2003;48:21-34.
2. Welsing PM, Landewe RB, van Riel PL, Boers M, van Gestel AM, van der Linden S, et al. The relationship between disease activity and radiologic progression in patients with rheumatoid arthritis: a longitudinal analysis. Arthritis Rheum 2004;50:2082-93.
3. Welsing PM, van Gestel AM, Swinkels HL, Kiemeney LA, van Riel PL, Laan RF, et al. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. Arthritis Rheum 2001;44:2009-17.
4. 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.
5. Larsen A, Dale K. Standardized radiological evaluation of rheumatoid arthritis in therapeutic trials. In: Dumonde DC, Jasani JK, editors. Recognition of anti-rheumatic drugs. Lancaster: MTP Press; 1977:285-92.
6. Scott D, Houssien D, Laasonen L. Proposed modification to Larsen's scoring method for hand and wrist radiographs. Br J Rheumatol 1995;34:56.
7. Rau R, Herborn G. A modified version of Larsen's scoring method to assess radiologic changes in rheumatoid arthritis. J Rheumatol 1995;22:1976-82.
8. Sharp JT, Young DY, Bluhm GB, Brook A, Brower AC, Corbett M, 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.
9. van der Heijde DM. Plain X-rays in rheumatoid arthritis: overview of scoring methods, their reliability and applicability. Baillieres Clin Rheumatol 1996;10:435-53.
10. Genant HK. Methods of assessing radiographic change in rheumatoid arthritis. Am J Med 1983;75:35-47.
11. Bruynesteyn K, van der Heijde D, Boers M, van der Linden S, Lassere M, van der Vleuten C. The Sharp/van der Heijde method out-performed the Larsen/Scott method on the individual patient level in assessing radiographs in early rheumatoid arthritis. J Clin Epidemiol 2004;57:502-12.
12. Emery P, Breedveld FC, Hall S, Durez P, Chang DJ, Robertson D, et al. Comparison of methotrexate monotherapy with a combination of methotrexate and etanercept in active, early, moderate to severe rheumatoid arthritis (COMET): a randomised, double-blind, parallel treatment trial. Lancet 2008;372:375-82.
13. Klareskog L, van der Heijde D, de Jager JP, Gough A, Kalden J, Malaise M, et al. Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. Lancet 2004;363:675-81.
14. Keystone E, van der Heijde D, Mason D Jr, Landewe R, Vollenhoven RV, Combe B, et al. Certolizumab pegol plus methotrexate is significantly more effective than placebo plus methotrexate in active rheumatoid arthritis: findings of a fifty-two-week, phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group study. Arthritis Rheum 2008;58:3319-29.
15. Smolen J, Landewé RB, Mease P, Brzezicki J, Mason D, Luijtens K, et al. Efficacy and safety of certolizumab pegol plus methotrexate in active rheumatoid arthritis: the RAPID 2 study. A randomised controlled trial. Ann Rheum Dis 2009;68:797-804.
16. Landewé RB, van der Heijde D, Klareskog L, van Vollenhoven R, Fatenejad S. Disconnect between inflammation and joint destruction after treatment with etanercept plus methotrexate: results from the trial of etanercept and methotrexate with radiographic and patient outcomes. Arthritis Rheum 2006;54:3119-25.
17. 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.
18. van der Heijde DM, van Leeuwen MA, van Riel PL, Koster AM, van 't Hof MA, van Rijswijk MH, 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.