ABSTRACT
Psoriatic arthritis (PsA) leads to structural damage that can be an important driver for disability and handicap associated with the disease. Serial radiographs, usually of hands and feet, facilitate follow-up documentation of development of these changes. Semi-quantitative scoring methods are designed to measure the degree of radiographically detectable joint damage, and of changes over time. Several radiographic scoring methods that had been developed originally for rheumatoid arthritis have been adopted for the use in PsA.

Four different scoring methods used in PsA are presented with instructions on how to use them: modified Steinbrocker global scoring method; PsA scoring method based on Sharp method for RA; Sharp van der Heijde modified method; and PsA Ratingen score (PARS). Available data on the reliability, sensitivity to change, and use in clinical trials, of these four methods are presented.

Introduction
Psoriatic arthritis (PsA) develops in about 20% of patients with psoriasis. The arthritis most frequently involves the hands and feet, but all other joints such as ankles, knees, elbows and shoulders may be affected. In contrast to rheumatoid arthritis, involvement of the distal interphalangeal (DIP) joints and an asymmetric pattern is common. Different patterns of distribution are described in hands and feet with symmetrical polyarthritis similar to rheumatoid arthritis or isolated transversal DIP involvement or involvement of single digital rays. An oligoarthritic pattern with only large joint involvement may occur as well (1). Enthesal involvement is frequent and contributes to pain and limited motion.

As in rheumatoid arthritis, structural damage is the consequence of continuing inflammation that can destroy cartilage and bone leading to functional impairment and disability. Joint damage can be measured on conventional radiographs. Cartilage destruction leads to joint space narrowing, bony destruction to erosions, in some patients to gross osteolysis and mutilation. Enthesal inflammation and ligament destruction can result in malalignment and subluxation. Gross osteolysis may be observed in severely destroyed joints, sometimes giving the appearance of a widened rather than a narrowed joint space. The characteristic ‘pencil-in-cup’ phenomenon is the result of a pincalled proximal bone protruding into the patelliform destroyed as well as expanded distal bone of the joint.

In psoriatic arthritis, osteoproliferative changes often accompany bony erosions; sometimes they may develop as the only radiographically detectable feature. Proliferative changes are considered to be pathognomonic for PsA and are therefore also included as the most important radiographic feature in classification criteria for the disease (2). An extreme consequence of proliferation is intraarticular osseous fusion or spontaneous ankyloses of joints affecting predominantly DIP, proximal interphalangeal (PIP) and wrist joints. Bone proliferation also occurs at sites of insertion of tendons and ligaments and at sesamoid bones in hands and feet.

Prevention of structural damage is an extremely important goal, as one can assume that, as in RA (3), long-term functional status of the patient is correlated to the amount of radiographically detected joint damage. Radiographically determined damage therefore is predictive for longterm outcomes and even for mortality in more severely affected patients (4). Measurement of radiographic damage, therefore, is a primary method to assess the efficacy of drug treatments. Only drugs that inhibit radiographic progression are regarded to be truly disease-modifying drugs (DMARDs).

Scoring methods
Several semiquantitative scoring methods to measure radiographic progression have been developed for RA.
These methods have also been used for psoriatic arthritis with some modifications to include the joints that may be involved more frequently in PsA and to consider the specific changes of the disease. A comprehensive overview over established methods was reported 10 years ago, which is updated in this article (5). Four methods are described briefly.

**Modified Steinbrocker Global Scoring Method**

This method was developed at the University of Toronto Psoriatic Arthritis Clinic, based on two types of criteria developed in 1949 by Steinbrocker and colleagues (6), which classified patients with RA for both functional and radiographic damage. The radiographic criteria were classified into 4 stages:

- **0** = normal
- **1** = juxta-articular osteoporosis or soft tissue swelling
- **2** = erosions
- **3** = erosion and joint space narrowing or subluxation
- **4** = total joint destruction, either lysis or ankylosis

The Toronto group used this classification not only for the mostly affected joint, but also for 40 joints in the hands and feet: all DIP, PIP, and metacarpophalangeal (MCP) joints of the hands with the wrist as one joint, and all metatarsophalangeal joints (MTPs) and the IP of the big toe. Thus the total score may range from 0 to 160 (7).

The method was compared to the Larsen method that had been developed for RA, using the modification proposed by Rau and Herborn (8). Interobserver intraclass correlation coefficient (ICC) for the original Steinbrocker, the modified Steinbrocker, and Larsen Score were found at 0.86, 0.86, and 0.87, respectively. The intraobserver ICCs for the original Steinbrocker method were 0.90 and 0.86; for the modified Steinbrocker 0.80 and 0.81; and for Larsen method 0.84 and 0.85. There are serious concerns to accept ICC data as quality measures for the reliability of a measurement method, as they are limited by dependence on single extreme cases and can only provide very limited information about the sensitivity to change (9). This problem is highlighted by the equally good ICC results of the original Steinbrocker method that considers just one joint compared to the results of the much more detailed modified Steinbrocker and Larsen Score.

**PsA Scoring Method based on the Sharp Scoring Method for RA**

Radiographic evaluation was performed in the initial studies with biologic agents in PsA using a modification of the Sharp method for RA (10), which include a separate evaluation of erosions and joint space narrowing (JSN). The same joints were scored as in the original method, with the addition of the DIP 2 to 5 joints of both hands. For the erosion score, a mixture of the original instructions for grades 0 to 5 of the Sharp-Score (counting the number of discrete erosions) and of the definitions of the Ratingen Score for RA with every 20% of joint surface destruction leading to an increased grade of the score (11) was used:

- **0** = no erosion
- **1** = one discrete erosion or involvement of less than 21% of the joint area by erosion
- **2** = two discrete erosions or involvement of 21–40% of the joint
- **3** = three discrete erosions or involvement of 41–60% of the joint
- **4** = four discrete erosions or involvement of 61–80% of the joint
- **5** = extensive destruction involving more than 80% of the joint.

Attempts were made to capture extreme bone destruction as the “pencil-in-cup” phenomenon scored separately with a grade of 6 and gross osteolysis with a score of 7, but these higher scores were not included in the sum scores. Therefore, the erosion score of the modified Sharp-Score for PsA has a range of 0–210 in the hands, 0–60 in the feet, and 0–270 in total.

Joint space narrowing was scored, according to grades 0 to 4 using the same definitions in the same joints as in RA, with the addition of the 8 DIP joints of the hands. As gross osteolysis may result in substantial joint space widening, this finding was scored separately with a score of 5 leading to the following definitions:

- **0** = normal joint
- **1** = asymmetrical and or minimal narrowing
- **2** = definite narrowing with loss of up to 50% of the normal space
- **3** = definite narrowing with loss of 51–99% of the normal space
- **4** = absence of a joint space, presumptive evidence of ankylosis
- **5** = widening

The metatarsophalangeal joint of the big toe was not scored because of the frequent osteoarthritic changes in this joint. The sum score did not include grade 5; therefore, the JSN-score ranges from 0 to 160 in the hands, 0 to 40 in the feet and 0 to 200 in total.

The total score for erosion and joint space narrowing of the modified Sharp Score for PsA may vary between 0 and 470 (5).

Other radiographically detectable changes in PsA, such as periostitis and tuft resorption are recorded and scored separately, but not included in the score value. The modified Sharp Score for PsA was used in a clinical trial of etanercept in PsA. Although the statistical evaluation of the reliability of the scoring method in this study using ICCs also was limited, radiographic progression at 12 months as measured with the modified Sharp Score for PsA was inhibited significantly in the etanercept group (annualised mean change of -0.03 unit) compared with a worsening of 1.00 unit in the placebo group (p=0.0001) (12), thereby documenting that the scoring method detected relevant differences between treatment arms.

The modified Sharp Score for PsA also was used in a slightly modified version for the ADEPT-trial of adalimumab in PsA (13). For this trial the os pisiforme and triquetrum were added but scored as one unit in the erosion score, and the space between the scaphoid and lunate bones for JSN were included while the interphalangeal space in the big toes was no longer evaluated. Also, the erosion score was expanded to include grades 6 for “pencil in cup” deformity and 7 for osteolysis. These values were incorporated into the total scores, resulting in a maximum modi-
Radiographic scoring methods in PsA / S. Wassenberg

Table I. An overview of the four Scoring methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Joints considered</th>
<th>Score range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Modified Steinbrocker (Toronto)</strong></td>
<td>Hands: 28 joints (wrist as one joint, IP 1 neglected) graded 0-4, Feet: 12 joints (including IP 1, graded 0-4, Hands and feet: 40 joints</td>
<td>0-160</td>
</tr>
<tr>
<td></td>
<td>Erosions: Hands: 40 joints including 7 bones in the wrist graded 0-5*, Feet: 12 joints graded 0-5*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Joint Space Narrowing: Hands 0-210, Feet 0-60, Hands + feet 0-270</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Erosions: Hands 0-40, Feet 0-40, Hands + feet 0-200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Joint Space Narrowing: Hands 0-160, Feet 0-100, Hands + feet 0-200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total Score: Hands 0-370, Feet 0-470</td>
<td></td>
</tr>
</tbody>
</table>

| Sharp Method for PsA (Etanercept Trial)** | Hands: 44 joints (+triquetrum/pisiforme) graded 0-7*, Feet: 10 joints graded 0-7* |             |
|                                          | Erosions: Hands 40 joints (+naviculare/lunatum) graded 0-4**, Feet: 8 joints (-IP 1) graded 0-4** |             |
|                                          | Joint Space Narrowing: Hands 0-308, Feet 0-70, Hands + feet 0-378                  |             |
|                                          | Erosions: Hands 0-160, Feet 0-32, Hands + feet 0-192                              |             |
|                                          | Joint Space Narrowing: Hands 0-468, Feet 0-102, Hands + feet: 0-570                |             |
|                                          | Total Score: Hands 0-468, Feet 0-102, Hands + feet: 0-570                          |             |

| Sharp Method for PsA (Adalimumab Trial)† | Hands: 40 joints graded 0-5, Feet: 12 joints graded 0-10 |                 |
|                                         | Erosions: Hands 40 joints graded 0-4, Feet 12 joints graded 0-4 |             |
|                                         | Joint Space Narrowing: Hands 0-200, Feet 0-120, Hands + feet 0-320 |             |
|                                         | Erosions: Hands 0-160, Feet 0-48, Hands + Feet 0-208 |             |
|                                         | Joint Space Narrowing: Hands 0-468, Feet 0-102, Hands + Feet: 0-570 |             |
|                                         | Total Score: Hands 0-468, Feet 0-102, Hands + Feet: 0-570 |             |

| Van der Heijde Method for PsA*          | Hands: 30 joints graded 0-5, Feet: 10 joints graded 0-5 |                 |
|                                       | Erosions: Hands 30 joints graded 0-4, Feet: 10 joints graded 0-4 |             |
|                                       | Joint Space Narrowing: Hands 0-150, Feet 0-50, Hands + feet 0-200 |             |
|                                       | Erosions: Hands 0-120, Feet 0-40, Hands + feet 0-160 |             |
|                                       | Joint Space Narrowing: Hands 0-270, Feet 0-90, Hands + feet: 0-360 |             |
|                                       | Total Score: Hands 0-270, Feet 0-90, Hands + feet: 0-360 |             |

| Psoriatic Arthritis Rating (PARS)       |Hands: 30 joints graded 0-5, Feet: 10 joints graded 0-5 |                 |
|                                       | Erosions: Hands 30 joints graded 0-4, Feet: 10 joints graded 0-4 |             |
|                                       | Joint Space Narrowing: Hands 0-150, Feet 0-50, Hands + feet 0-200 |             |
|                                       | Erosions: Hands 0-120, Feet 0-40, Hands + feet 0-160 |             |
|                                       | Joint Space Narrowing: Hands 0-270, Feet 0-90, Hands + feet: 0-360 |             |
|                                       | Total Score: Hands 0-270, Feet 0-90, Hands + feet: 0-360 |             |

1Other features typical for PsA like periostitis and tuft resorption are recorded but not included in the score.
2Grades 6 for pencil in cup lesions and 7 for gross osteolysis are recorded in both trials but counted only in the Adalimumab-Trial.
3Grade 5 for widening is recorded separately but not counted for the sum scores.
4Pencil in cup and gross osteolysis are recorded separately. If present they are counted with the maximum score for both erosion and joint space narrowing.

The modification based on the Sharp method for RA scores the same joints and definitions as seen in RA, with the addition of the DIP joints 2 to 5 of both hands (14). The erosion score can vary between 0 and 5 in the hands and 0 and 10 in the feet:

0 = no erosions
1 = discrete erosion
2 = large erosion not passing the middle-line
3 = large erosion passing the middle-line
4 = sum of grades adding to 4
5 = sum of grades adding to 5

In the feet 0–5 scores are calculated for each side (proximal and distal) of the joint therefore the maximum score can reach 10. The joint space narrowing score can vary between 0 and 4 in hands and feet:

0 = normal
1 = focal or general narrowing up to a maximum of 25% of the original joint space
2 = definite narrowing loss of less than 50% of the original joint space
3 = definite narrowing of more than 50% of the original joint space or subluxation
4 = no detectable joint space or ankylosis or complete luxation.

Gross osteolysis and pencil-in-cup is scored separately in the same joints. If present these changes are scored with the maximum score for both erosions and joint space narrowing. Thus the erosion score can vary between 0 and 200 in the hands and 0 and 120 in the feet; the JSN score between 0 and 160 in the hands and 0 and 48 in the feet. The total erosion score may sum up to a maximum of 320, the total JSN score to a maximum of 208, and the combined total score to a maximum of 528.

The method has been analysed for reliability and used to assess treatment response in several trials with different biologic agents in PsA (15–19). As examples the most relevant results of two of these studies are discussed.

In the IMPACT 2 trial (15) of infliximab compared to control, the inter- and intra-reader consistency of the x-ray readings as determined by ICC was between 0.97 and 1. Patients in the active treatment group had significantly less structural damage progression with 24 weeks mean ± SD changes in PsA-modified SHS from baseline being signifi-
Significantly smaller in the infliximab group (-0.70±2.53) compared to the placebo group (0.82±2.62) \( p<0.001 \). Also a significantly lower proportion of patients in the infliximab group had a change in the total Sharp/van der Heijde score greater than the SDC (2.7) when compared with the placebo group (3% vs. 12%; \( p<0.017 \)). No between-group differences were observed for the number of involved joints, and the number of patients who had the PsA-specific radiographic features of pencil-in-cup and gross osteolysis deformities.

In the GO-REVEAL trial (16) of Goli-mumab 50 and 100 mg versus control, the consistency of the x-ray readings demonstrated agreement with ICCs of 0.94, 0.93, and 0.93 at baseline, week 24, and week 52 respectively. At 24 weeks, mean ± SD changes from baseline in PsA-modified SHS were significantly less in patients receiving golimumab 50 mg (-0.16±1.31) compared to the control group (0.27±1.26) \( p=0.011 \). The difference between the golimumab 100 mg group (-0.02±1.32) and the control group at week 24 approached statistical significance (\( p=0.086 \)). Significantly fewer patients in the golimumab 50 mg group had radiographic progression defined by change greater than the SDC (1.56) compared to the placebo group (3.8% vs. 12.8%; \( p=0.030 \)).

### Psoriatic Arthritis Ratingen Score (PARS)

This method was developed based on the Rau and Herborn modification of the Larsen Score (8). DIP joints 2-5 of the hands were added. Altogether it includes 40 joints of the hands and feet (DIP 2-5 of the hands, 2 IPs of the thumbs, 8 PIPs of the hands, 10 MCPs, both wrists, 2 IPs of the great toes and MTPs 2 to 5). All joints are scored separately for destruction and proliferation (20). The destruction score (DS) is based on the amount of joint surface destruction on a 0–5 scale:

\[
\begin{align*}
0 &= \text{normal} \\
1 &= \text{one or more definite erosions with an interruption of the cortical plate of >1 mm but destruction of <10% of the total joint surface} \\
2 &= \text{destruction of 11–25% of joint surface} \\
3 &= \text{destruction of 26–50% of joint surface} \\
4 &= \text{destruction of 51–75% of joint surface} \\
5 &= \text{destruction of more than 75% of joint surface}
\end{align*}
\]

The proliferation score (PS) considers any kind of bony proliferation typical for PsA on a 0–4 scale:

\[
\begin{align*}
0 &= \text{normal} \\
1 &= \text{bony proliferation measured from the original bone surface of 1-2 mm or, if the margins of the proliferation cannot be distinguished from the original bone surface, clearly identifiable bone growth not exceeding 25% of the original diameter of the bone} \\
2 &= \text{bony proliferation of 2–3 mm or bone growth between 25–50%} \\
3 &= \text{bony proliferation >3 mm or bone growth >50%} \\
4 &= \text{bony ankylosis}
\end{align*}
\]

The DS (0-200) and the PS (0–160) are added to give the total score TS (0–360) for each patient. The method was validated using complete sets of x-rays of 20 patients with active psoriatic arthritis taken at a mean time interval of 3 years (20). Reliability was measured using a hierarchical analysis of variance model (ANOVA), which compared progression with intra- and inter-reader variation. The ratio of change-STD (progression) to intra- or inter-reader-STD (measurement error) with values >1 determines the likelihood of detecting progression (9).

Ratios of were found with 3.3 (reader 1), 2.0 (reader 2) and 3.8 (both readers) for the DS, 2.2, 4.2 and 2.7 for the PS and 3.6, 2.8 and 3.9 for the TS. Comparing the change over time of the DS with the change of the PS revealed that there was only a weak correlation between both features, suggesting that
proliferation develops at least in part independently from destruction. Measuring both features separately therefore adds significant information compared to a global simple score.

This method has not yet been used in a clinical trial, but was used in a single-center observational study in early PsA patients that was performed in Sweden with 197 PsA patients, of whom 72 had radiographs scored at baseline and 5 years later (21). The mean change in the proliferation score of 1.8 was highly significant and contributed even more to the change of the total score of 3.1 than the change of the erosion score of 1.25. Baseline and 5-year scores were highly correlated (for total scores: Spearman rho 0.752, p=0.000). The baseline total score was correlated with ESR (rho: 0.364, p=0.004) and 5-year score with swollen joint count (rho 0.310, p=0.016). Male gender (OR 4.42, 95% CI: 0.35–8.49, p=0.034) and higher total baseline radiographic scores (OR: 2.23, 95% CI 1.80–2.65, p=0.000) were the only predictors of radiographic abnormalities after 5 years.

Discussion

Four radiographic scoring methods have been used in different studies of PsA. All have been proven to capture radiographic change with reasonable precision. The two methods based on the Sharp Score and the van der Heijde modified Sharp Score documented significant differences in radiographic progression in controlled clinical trials. No direct comparison of the performance of the four methods in capturing radiographic change has been reported. As a component of an OMERACT initiative, an exercise was performed with reading of the radiographs of the IMPACT 1 trial by the four methods, but the overall radiographic change in this trial was so small that differences in the reliability and the sensitivity to change between the methods could not be determined (unpublished data). Therefore one cannot decide based on data which method should be preferred.

It remains an open question whether inclusion of proliferative changes or of other specific features of PsA would add important information that would lead to a better capacity to assess joint progression in PsA. In the ADEPT trial, IMPACT2 and the GO-REVEAL trials, some of these phenomena such as penicil in cup and gross osteolysis were unusual, and could therefore not contribute meaningful incremental information about change of radiographic scores. Proliferative changes are very common and are regarded to be specific for PsA. Therefore they are recognised for the classification of the disease (2). Until now the PARS is the only scoring method that focuses on these changes. The finding of the Swedish observational cohort (21) using the PARS that the proliferation score contributed more to the observed change than the erosion score over a longer period of 5 years suggests that it is desirable to focus further research on the relevance of proliferative changes for radiographic progression in PsA. It is also necessary to define the relevance of different radiographic features measured by radiographic scoring methods for the most important outcome parameter for the patient that is physical function.

References