

# Selection of a Method for Scoring Radiographs for Ankylosing Spondylitis Clinical Trials, by the Assessment in Ankylosing Spondylitis Working Group and OMERACT

DÉSIRÉE VAN DER HEIJDE, ROBERT LANDEWÉ, and the ASAS Working Group

**ABSTRACT.** Radiographs are important for assessing structural damage in patients with ankylosing spondylitis (AS); this technology was selected by the international ASessment in Ankylosing Spondylitis (ASAS) Working Group as an important domain for assessing outcome in clinical trials. The selection of a preferred scoring method based on the OMERACT filter still needs to be undertaken. This article describes the scientific basis for and process of making the selection. Finally, both the ASAS Working Group and OMERACT participants voted for the modified Stoke AS Spinal Score (mSASSS) as the preferred scoring method for use in clinical trials in AS. (*J Rheumatol* 2005; 32:2048–9)

*Key Indexing Terms:*

IMAGING ANKYLOSING SPONDYLITIS SCORING METHODS RADIOGRAPHS

Radiographs of the spine were selected by the international ASessment in Ankylosing Spondylitis (ASAS) Working Group as an important outcome domain in clinical trials<sup>1</sup>. But at the time of selection and during the fourth OMERACT conference on radiographic outcome in AS, there was too little information on scoring methods, in relation to various aspects of the OMERACT filter, to select a preferred method<sup>2</sup>. Further validation of the methods was put on the research agenda as one of the most important topics.

The 3 existing scoring methods — the Bath Ankylosing Spondylitis Radiology Index (BASRI), Stoke Ankylosing Spondylitis Spinal Score (SASSS), and the modified SASSS (mSASSS)<sup>3–5</sup> — have been compared in only one study to date. This research<sup>6</sup> served as the basis for selection of a preferred scoring method by the ASAS Working Group and OMERACT. A specific module at OMERACT 7 was devoted to this task.

During the plenary introduction the 3 scoring methods were presented, with a focus on the 3 aspects of the OMERACT filter: truth, discrimination, and feasibility<sup>7</sup>. Thereafter 2 discussion groups were asked to rank the 3 scoring methods for each of the 3 aspects of the OMERACT filter separately, and then the rankings were combined. The results

from the discussion groups were presented to the plenary audience, immediately followed by a vote. Conclusions of the published comparison of the 3 scoring methods are summarized here<sup>6</sup>.

## Truth

To assess the aspect “truth” of the OMERACT filter (“does the method assess what should be assessed”), we assessed which anatomical areas were most frequently involved and contributed most to assessing progression. Moreover, the relation with spinal mobility was appraised. The main conclusions on the aspect “truth” were:

1. SI joints on radiographs are not useful for measuring change in a clinical trial: a large proportion of patients already have the maximum score relatively early in the course of the disease; most important, there is only a little progression in a small proportion of patients. During a 4-year followup only 9% of the patients showed any change at all.
2. Radiographic progression of hip joints is not a useful outcome in a clinical trial, as only a small number of patients have involvement of the hips, and again only very few patients show measurable change over time (8% of the patients over a 4-year period).
3. The BASRI method was able to detect progression in 18% of the patients in the cervical spine and in 23% of the patients in the lumbar spine. For the mSASSS these progressions at these sites were 43% and 41%. The SASSS assesses only lumbar spine, and it detected progression in 46% of patients. These data show that the SASSS methods are able to pick up progression in a higher percentage of patients than the BASRI. Moreover, progression is present

---

*From the Department of Internal Medicine, Division of Rheumatology, University Hospital Maastricht, and CAPHRI Research Institute, University Maastricht, Maastricht, The Netherlands.*

*D.M.F.M. van der Heijde, MD, PhD, Professor of Rheumatology;  
R.B.M. Landewé, MD, PhD, Associate Professor of Rheumatology.*

*Address reprint requests to Prof. D. van der Heijde, University Hospital Maastricht, Department of Internal Medicine, Division of Rheumatology, PO Box 5800, 6202 AZ Maastricht, The Netherlands.  
E-mail: dhe@sint.azm.nl*

in the cervical spine in a considerable number of patients, indicating that the cervical spine is essential to be included in a scoring method.

4. The correlations between the scoring methods and various spinal mobility measures ranged from 0.58 to 0.76, and were similar for all 3 scoring methods.

Based on all these aspects of "truth" together, a method including the cervical spine is preferred, and it is not useful to include scoring of SI joints and hips in a clinical trial of patients with established AS.

### Discrimination

For the aspect of discrimination, intra/interobserver reliability of change scores and sensitivity to change were analyzed. Interobserver reliability for measuring change in a 2-year interval was better for the mSASSS (intraclass correlation coefficient 0.82) than for the SASSS (ICC 0.44) or the BASRI (ICC 0.51). The lower reliability of the SASSS was mainly due to disagreement between readers in scoring the posterior site of the lumbar spine. As mentioned, the mSASSS was able to identify progression in the highest percentage of patients, and also the greatest amount of progression per patient. The BASRI method suffered from a substantial ceiling effect (the highest possible score was found in 15% of the patients for the cervical spine and in 23% of the patients for the lumbar spine, as compared to 4% and 5%, respectively, if assessed by the mSASSS).

Overall, the mSASSS performed best for the various features of discrimination.

### Feasibility

The BASRI method takes somewhat less time to perform, but the radiation exposure is highest for the BASRI because of the additional anteroposterior radiograph at lumbar spine.

### Voting and conclusions

The recommendations from discussion groups for the research agenda were to collect more data on sensitivity to change of the mSASSS over longer followup periods. A

second recommendation was to develop a cutoff of a maximum level of radiographic damage for inclusion in a clinical trial with mSASSS as an outcome.

All participants were invited to vote on the following question: "Based on all the features of the OMERACT filter, which method for scoring radiographs in AS for use in clinical trials do you prefer." Of all voting participants, 81% were in favor of the modified SASSS, 4% voted for the BASRI, 1% for the SASSS, 10% did not know, and 4% had no preference. These results endorsed the preference for the modified SASSS as already concluded by the ASAS group.

Consequently, the combined ASAS/OMERACT recommended scoring method for assessing structural damage in the spine in clinical trials is the modified SASSS. Use of the mSASSS in future trials will give us more insight into both performance of the method and the potency of drugs to inhibit the progression of structural damage in AS.

### REFERENCES

1. van der Heijde D, Bellamy N, Calin A, Dougados M, Khan MA, van der Linden S. Preliminary core sets for endpoints in ankylosing spondylitis. *J Rheumatol* 1997;24:2225-9.
2. van der Heijde D, van der Linden S, Dougados M, Bellamy N, Russell AS, Edmonds J. Ankylosing spondylitis: plenary discussion and results of voting on selection of domains and some specific instruments. *J Rheumatol* 1999;26:1003-5.
3. Calin A, Mackay K, Santos H, Brophy S. A new dimension to outcome: application of the Bath Ankylosing Spondylitis Radiology Index. *J Rheumatol* 1999;26:988-92.
4. Dawes PT. Stoke Ankylosing Spondylitis Spine Score. *J Rheumatol* 1999;26:993-6.
5. Creemers M, Franssen M, van 't Hof M, Gribnau F, van de Putte L, van Riel P. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. *Ann Rheum Dis* 2005;64:127-9. Epub 2004 Mar 29; doi.10.1136/ard.2004.020503
6. Wanders AJB, Landewé RBM, Spoorenberg A, et al. What is the most appropriate radiological scoring method for Ankylosing Spondylitis? A comparison of the available methods based on the OMERACT filter. *Arthritis Rheum* 2004;50:2622-32.
7. Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for Outcome Measures in Rheumatology [editorial]. *J Rheumatol* 1998;25:198-9.