CONCISE REPORT

Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system

M C W Creemers, M J A M Franssen, M A van 't Hof, F W J Gribnau,
L B A van de Putte, P L C M van Riel

....

Objective: To develop and validate an extensive radiographic scoring system for ankylosing spondylitis (AS).

Methods: The Stoke Ankylosing Spondylitis Spinal Score (SASSS) was modified by adding a score for the cervical spine and defining squaring. This modified SASSS (mSASSS) is the sum of the lumbar and cervical spine score (range 0–72). 370 lateral views of the lumbar and cervical spine were used for development of the mSASSS, standardisation of observers, and for studying reliability. In a 48 week NSAID study of 57 patients, change over time and construct validity were studied.

Results: Interobserver correlations of the lumbar and cervical spine scores were good ($r > 0.95$). The interobserver duplicate error was 0.55 in a range from 0 to 36. The mean change in the cervical and lumbar spine scores between weeks 0 and 48 of all patients was 1.45 (range 0–6.0) and 1.06 (0–5.0), respectively (paired t testing, $p < 0.001$). Change in radiological score was seen in 36/57 (63%) patients (lumbar and cervical spine 11, cervical spine 12, lumbar spine 13 patients).

Conclusion: The mSASSS is useful for assessing extensive radiographic damage in AS. It is reliable, detects changes over 48 weeks, and shows a satisfactory face and construct validity.

Treatment of patients with ankylosing spondylitis (AS) with anti-tumour necrosis factor agents has shown short term efficacy on symptoms and signs of the disease.[1][2] However, these treatments are costly and safety in the long term is not known. Therefore it is important to know whether these agents, in addition to their beneficial effects on signs and symptoms, do also influence the structural process. Appropriate radiographic scoring methods are thus necessary to assess the disease, which mostly affects the whole axial spine.[3] In the early 1990s we developed and validated a more extensive radiological scoring system for AS, in which next to the lumbar spine the cervical spine was encompassed (face validity). This radiographic scoring system is a modification of the Stoke Ankylosing Spondylitis Spinal Score (SASSS),[4] which was the only published radiographic scoring system in AS when we developed this method. This extended radiological scoring system has been used in a 48 week non-steroidal anti-inflammatory drug (NSAID) study in patients with AS to describe radiographic features, to evaluate change after a follow up of 48 weeks, and construct validity. In the past, commonly used treatments like NSAIDs and disease modifying antirheumatic drugs such as sulfasalazine[5] did not influence the course of the disease. Probably because of this there was less interest in these kinds of outcome measure in AS, and our study[6] has not been published before.

We here present the original study of this extended radiographic scoring method—called the modified Stoke Ankylosing Spondylitis Spinal Score (mSASSS)—briefly, because recent studies with the mSASSS have shown promising results[7][8] compared with the other two published radiographic scoring systems.[4][9]

METHODS

The study was conducted with radiographs of patients with AS, diagnosed according to the modified New York criteria.[10] Three hundred and seventy lateral views of the lumbar and cervical spine were used for development of the scoring system and standardisation of the two observers (MF, MC). Radiographs were scored blinded for the patient and the previous score.

Reliability

Radiographs, which had not been used previously, were scored twice by two observers, after standardisation of the observers who were unaware of the patient and the previous score. The time between the duplicate scores was 8 weeks in order to guarantee independency of the scores. Interobserver correlations and interobserver duplicate error ($\sqrt{\Sigma d_i^2 / 2n}$) of the different scores were calculated. Intraobserver correlation and intraobserver error were only studied if the interobserver correlation was < 0.95.

To study whether this method could detect *changes over time* radiographs of 57 patients with AS (43 male, 13 female; mean disease duration 10 years; 54 HLA-B27 positive) participating in a 48 week NSAID[6][11] trial were used. Radiographs were obtained at weeks 0 and 48 of the trial and scored sequentially blinded for the patient and treatment.

Construct validity was studied in this NSAID study by calculating Pearson correlations at weeks 0, 12, and 48 between the mSASSS and assessed clinical variables, if necessary transformed to obtain normality. The following variables were assessed at weeks 0, 2, 4, 8, 12, 18, 24, 30, 36, 42, and 48: physicians' assessed variables (performed by a single observer): occiput to wall distance, chest expansion, 10 cm Schober test, lumbar flexion index,[12] lumbar lateral flexion,[13] 53 swollen joint count, enthesis index,[14] and mobility of the cervical spine in all three planes; patients' assessed variables: duration of morning stiffness, 100 mm visual analogue scale (VAS) score for spinal pain during the day, VAS spinal pain during the night, and VAS general wellbeing; and laboratory variables: erythrocyte sedimentation rate, C reactive protein, haemoglobin, leucocytes, and immunoglobulin A. In addition, every 12 weeks the Dutch Functional Index[15] was completed.

....

Abbreviations: AS, ankylosing spondylitis; mSASSS, modified Stoke Ankylosing Spondylitis Spinal Score; NSAID, non-steroidal anti-inflammatory drug; VAS, visual analogue scale

Table 1 Means (SD), inter- and intraobserver correlations, duplicate errors, and p values of t testing

| | Observer | Score (systematic differences) | p Value | Interobserver | |
				Correlation	Duplicate error
Cervical spine (n = 26)					
	1	11.78 (13.32)	0.03	0.99	0.55
	2	11.88 (13.48)			
Lumbar spine (n = 46)					
	1	6.12 (11.35)	0.70	0.99	0.99
	2	6.36 (11.32)			

The mSASSS system contained a score for the lumbar spine and a score for the cervical spine. The total score is the sum of both scores (range 0–72):

- *Lumbar spine*: the scoring system, developed by Taylor *et al*,[4] was applied to the lower border of the 12th thoracic vertebra, all five lumbar vertebrae, and the upper border of the sacrum, which were viewed on the lateral radiograph. The corresponding nominal scoring system was used: 0 = no abnormality; 1 = erosion, sclerosis or squaring; 2 = syndesmophyte; and 3 = total bony bridging at each site. Radiological abnormalities, not related to AS, such as osteophytes, and sites not clearly visible on the radiograph were not considered for scoring. Radiographs were only taken into account if no more than three scoring sites were missing. Squaring was defined as present if a line, fictively drawn with a transparent ruler, from the upper and lower border of each vertebral body overlayed 50% or more with the surface of the vertebra, starting at either the upper or lower border; or if the surface of the vertebra was convex, approaching the method of Ralston *et al*.[16] To deal efficiently with missing observations the total score (range 0–36) was calculated as 12 times the mean score of all scoring sites.

- *Cervical spine*: The scoring system was identical to that of the lumbar spine. The lower border of the 2nd cervical vertebra up to and including the upper border of the 1st thoracic vertebra were viewed on the lateral radiograph. Owing to the original straight shape of the lateral surface of the 3rd cervical vertebra it was decided that this vertebra should not be scored for squaring, although erosions and sclerosis were scored.

RESULTS AND DISCUSSION
Reliability
In total 46 radiographs of the lumbar and 26 of the cervical spine were scored (table 1). The interobserver correlations of the lumbar and cervical spine scores were good ($r>0.95$). Although the cervical spine score showed a statistically significant difference between the scores of the two observers, the interobserver duplicate error of 0.55 in a range from 0 to 36 can be considered as relatively small. Intraobserver correlation and intraobserver duplicate error were not calculated because the interobserver correlation was 0.99.

Radiographs of patients in the NSAID study showed total bony bridging from C7 to Th1, C2 to C3, C3 to C4, and from Th12 to L1. The 12th thoracic, the 1st lumbar, and the 4th cervical vertebrae showed most frequently radiological involvement. In total 51/57 (89%) patients showed involvement of the cervical spine and 37/57 (65%) patients of the lumbar spine at baseline. Differences between men and women were not statistically significant if adjusted for disease duration, which was 11 (SD 8) years and 6 (4) years for men and women, respectively.

Change over time
The mean change in scores between weeks 0 and 48 of all patients was 1.45 (range 0–6.0) in the cervical spine score and 1.06 (0–5.0) in the lumbar spine score; all changes were significant (paired *t* testing, $p<0.001$). Change in score was seen in 36 (63%) patients—in 11 patients in both the lumbar and cervical spine, in the cervical spine only in 12 patients, and in the lumbar spine only in 13 patients (fig 1).

The radiographic scores changed statistically significantly. However, it should be noted that the radiographs were scored in sequential order in this longitudinal study, allowing for a positive difference only, for reduction of measurement error.

Construct validity
Chest expansion, occiput to wall distance, cervical anteflexion, cervical retroflexion, cervical rotation, cervical lateral flexion, lumbar flexion index, and lumbar lateral flexion showed significant Pearson correlations with the mSASSS at weeks 0, 12, and 48 (p value varied from <0.05 to <0.0005).

In conclusion, this mSASSS is a useful method for assessing radiographic damage in AS: it is reliable, can detect change over 1 year, and has an acceptable construct validity. It shows that radiological involvement is present in the cervical spine when lumbar spine involvement is not found.
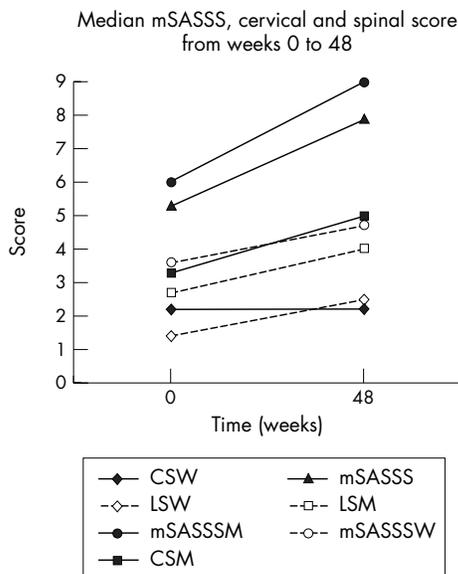
Figure 1 Radiographic changes from weeks 0 to 48 for all patients, men and women separately. CSW, cervical spine women; CSM, cervical spine men; LSW, lumbar spine women; LSM, lumbar spine men; mSASSS, modified Stoke Ankylosing Spondylitis Score in all patients; mSASSSW, mSASSS women; mSASSSM, mSASSS men.

Thus, it should be emphasised that inclusion of the cervical spine score in the mSASSS[7 8] adds important information, which should be used in the evaluation of treatments such as anti-tumour necrosis factor agents. The mSASSS should be validated further in long term studies, paying particular attention to other aspects of validity such as predictive validity and validation in respect of clinical variables.

. . . .

Authors' affiliations

M C W Creemers, L B A van de Putte, P L C M van Riel, Department of Rheumatology, University Medical Centre St Radboud, Nijmegen, The Netherlands

M J A M Franssen, Department of Rheumatology, St Maartenshospital, Nijmegen, The Netherlands

M A van 't Hof, Department of Medical Statistics, Catholic University, Nijmegen, The Netherlands

F W J Gribnau, Department of Clinical Pharmacology, University Medical Centre St Radboud, Nijmegen, The Netherlands

Acting editor: Frank A Wollheim

Correspondence to: Dr M C W Creemers, Department of Rheumatology, University Medical Centre St Radboud, PO Box 9101, 6500 HB Nijmegen, The Netherlands; m.creemers@reuma.umcn.nl

REFERENCES

1 **Braun J**, Brandt J, Listing J, Zink A, Alten R, Burmester, *et al.* Long-term efficacy and safety of infliximab in the treatment of ankylosing spondylitis. An open observational extension study of a three-month, randomized, placebo-controlled trial. *Arthritis Rheum* 2003;**48**:2224–33.
2 **Brandt J**, Khariouzov A, Listing J, Haibel H, Sörensen, Grassnickel L, *et al.* Six-month results of a double-blind, placebo-controlled trial of etanercept treatment in patients with ankylosing spondylitis. *Arthritis Rheum* 2003;**48**:1667–75.
3 **Carette S**. The natural disease course of ankylosing spondylitis. *Arthritis Rheum* 1983;**26**:186–90.
4 **Taylor HG**, Wardle T, Beswick EJ, Dawes PT. The relationship of clinical and laboratory measurements to radiological change in ankylosing spondylitis. *Br J Rheumatol* 1991;**30**:330–5.
5 **Creemers MCW**, van Riel PLCM, Franssen MJAM, van de Putte LBA, Gribnau FWJ. Second-line treatment in seronegative spondyloarthropathies. *Semin Arthritis Rheum* 1994;**24**:71–81.
6 **Creemers MCW**. Clinical and methodological studies in ankylosing spondylitis. The Netherlands: Catholic University Nijmegen, 1994 (Thesis.).
7 **Wanders AJB**, Landewé RBM, Van der Linden S, Dougados M, Mielants H, Van der Tempel, et al. What is the most appropriate scoring method for ankylosing spondylitis in clinical trials? A comparison of the available methods based on the OMERACT filter [abstract]. *Arthritis Rheum* 2003;**48**(suppl):S440.
8 **van der Heijde DMF**, Landewé RBM, Spoorenberg A, Wanders A, van der Linden S, De Vlam K, *et al.* Can a historical cohort of patients with ankylosing spondylitis containing data of 2-year radiographic progression serve as a control group to assess inhibition of structural damage by TNF-blockers [abstract]. *Arthritis Rheum* 2003;**48**(suppl):S441.
9 **MacKay K**, Mack C, Brophy S, Calin A. The Bath ankylosing spondylitis radiology index: a new, validated approach to disease assessment. *Arthritis Rheum* 1998;**41**:2263–70.
10 **van der Linden S**, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis; a proposal for modification of the New York criteria. *Arthritis Rheum* 1984;**27**:361–8.
11 **Creemers MCW**, van 't Hof MA, Franssen MJAM, van de Putte LBA, Gribnau FWJ, van Riel PLCM. Disease activity in ankylosing spondylitis: selection of a core set of variables and a first step in the development of a disease activity score. *Br J Rheumatol* 1996;**35**:867–73.
12 **Adrichem AJM**, van der Korst JK. Assessment of the flexibility of the lumbar spine. *Scand J Rheumatol* 1973;**2**:87–91.
13 **Domján L**, Nemes T, Bálint GP, Tóth Z, Gömör B. A simple method for measuring lateral flexion of the dorsolumbar spine. *J Rheumatol* 1990;**17**:663–5.
14 **Mander MM**, Simpson JM, McLellan A, Walker D, Goodacre JA, Dick WC. Studies with an enthesis index as a method of clinical assessment in ankylosing spondylitis. *Ann Rheum Dis* 1987;**46**:197–202.
15 **Creemers MCW**, van 't Hof MA, Franssen MJAM, van de Putte LBA, Gribnau FWJ, van Riel PLCM. A Dutch version of the functional index for ankylosing spondylitis. Development and validation in a long-term study. *Br J Rheumatol* 1994;**33**:842–6.
16 **Ralston SH**, Urquhart GDK, Brzeski M, Sturrock RD. A new method for the radiological assessment of vertebral squaring in ankylosing spondylitis. *Ann Rheum Dis* 1992;**51**:330–3.