Do We Really Need to Evaluate Entire Cervical Spines for Squaring Score in Modified Stoke Ankylosing Spondylitis Spinal Score?

TAE-JONG KIM, HEE-SUN KIM, KYUNG-BIN JOO, SUNGJUN KIM, and TAE-HWAN KIM

ABSTRACT. ***Objective.*** To evaluate squaring of cervical spine in ankylosing spondylitis (AS) and to suggest whether assessment for squaring of entire cervical spines in the modified Stoke AS Spinal Score (mSASSS) is really needed.

Methods. The study group comprised 125 healthy subjects and 135 patients with AS. Two experienced radiologists assessed the presence or absence of squaring in the individual cervical vertebrae twice independently. The frequencies of squaring of each vertebral site were obtained, and the differences between the control and AS groups were statistically analyzed.

Results. Kappa intra- and interobserver reliability coefficients were 0.76 and 0.61, respectively. The frequency of spinal squaring for each vertebra in the control group was as follows: the 2nd (lower border 100%), 3rd (upper border 85.2%–91.0%, lower border 91.0%–93.4%), 4th (upper border 44.3%–49.2%, lower border 57.4%–63.1%), 5th (upper border 5.0%–6.6%, lower border 8.4%–15%), 6th (upper border 2.7%–7.3%, lower border 4.7%–8.6%), 7th cervical vertebra (upper border 39.7%–45.6%, lower border 51.5%–54.1%), and the 1st thoracic vertebra (upper border 56.0%–62.5%). Although there were statistically significant differences at 4th, 5th, 6th, and upper border of 7th cervical vertebra, a high proportion of normal 4th and 7th cervical vertebra were scored as squared.

Conclusion. Only the 5th and 6th cervical spines are useful for scoring squaring of cervical spine in the mSASSS. (First Release Jan 15 2008; J Rheumatol 2008;35:477–9)

Key Indexing Terms:
ANKYLOSING SPONDYLITIS SQUARING
MODIFIED STOKE ANKYLOSING SPONDYLITIS SPINAL SCORE

Ankylosing spondylitis (AS) is characterized by inflammatory back pain due to sacroiliitis and spondylitis, and the formation of syndesmophytes leading to ankylosis[1]. These defects cause severe disability and loss of work capacity[2]. Symptom control has been accomplished by physical therapy and nonsteroidal antiinflammatory drugs (NSAID) for decades. But recently, treatment with biologic agents has revolutionized the management of AS[3-5]. Therefore, an appropriate radiographic scoring method is essential to investigate whether these new agents influence the structural process. Although some methods[6,7] have been suggested, the modified Stoke AS Spinal Score (mSASSS) has been identified as the most sensitive scoring method[8,9] for evaluation of spinal changes.

In assessment of mSASSS, the spinal squaring is an important factor of scoring. However, it is not easy to distinguish normal from squared vertebrae. It has been recognized that a normal 3rd cervical vertebra cannot be distinguished from a squared one. But the ability to recognize squaring in other sites where mSASSS is scored has not been investigated. Moreover, it is time-consuming to assess all 12 sites of cervical radiography.

In light of this, we evaluated squaring of each cervical vertebra to determine whether assessment of all 12 cervical sites in mSASSS is really necessary.

MATERIALS AND METHODS

This study was conducted using cervical radiographs of 135 patients with AS diagnosed according to the modified New York criteria[10] and 125 age-matched healthy controls having no joint symptoms. Patients who had cervical symptoms for at least 10 years were recruited because radiological spinal signs related to AS can evolve for up to 10 years[11]. Two hundred sixty lateral radiographs of the cervical spines were then randomly mixed to minimize subjective bias. The lower border of the 2nd cervical vertebra down to and including the upper border of the 1st thoracic vertebra were viewed. In interpreting cervical vertebrae, we excluded any images of vertebra that were not properly recorded or if the image had poor quality.

From the Department of Rheumatology and Department of Radiology, The Hospital for Rheumatic Diseases, Hanyang University College of Medicine; and the Department of Radiology, Hanyang University GURI Hospital, Hanyang University College of Medicine, Seoul, Korea.

T-J. Kim, MD, Instructor of Medicine; H-S. Kim, MD, Clinical Fellow; T-H. Kim, MD, PhD, Associate Professor, Department of Rheumatology, The Hospital for Rheumatic Diseases, Hanyang University College of Medicine; K-B. Joo, MD, PhD, Professor, Department of Radiology, The Hospital for Rheumatic Diseases, Hanyang University College of Medicine; S. Kim, MD, PhD, Assistant Professor, Department of Radiology, Hanyang University GURI Hospital, Hanyang University College of Medicine.

Address reprint requests to Dr. T-H. Kim, Department of Rheumatology, The Hospital for Rheumatic Diseases, Hanyang University College of Medicine, 17 Haengdang-Dong, Sungdong-Gu, Seoul 133-792, Korea. E-mail: thkim@hanyang.ac.kr

Accepted for publication October 22, 2007.

Cervical spine squaring. Squaring was defined as present when an imaginary line, drawn with a transparent ruler, between the upper and lower border of each vertebral body overlayed by 50% or more the surface of the vertebra[8]; or when the surface of the vertebra was convex, approaching the method of Ralston, *et al*[12].

Reliability. Cervical radiographs were blindly assessed twice by 2 experienced bone and joint radiologists (KBJ, SK) to validate the results. The time between each scoring was 10 weeks in order to ensure the autonomy of the scores. A kappa statistic was used to determine the significance of intra- and interobserver variability. For interpretation of kappa, a value > 0.75 indicates excellent agreement, a value between 0.40 and 0.75 indicates fair to good agreement, and a value < 0.40 indicates poor agreement[13].

Analysis of radiographs. First, the presence or absence of squaring in individual vertebrae of 125 controls (all male) was analyzed descriptively. Second, cervical radiographs of 135 AS patients (128 male, 7 female) were compared to the control group, using chi-square or Fisher's exact tests. Odds ratio (OR) and 95% confidence interval (95% CI) were estimated for each squaring. Cohen's kappa statistic was then used to assess intra- and interobserver agreement. P values < 0.05 were regarded to be statistically significant.

RESULTS

Cohen's kappa intra- and interobserver reliability coefficients were 0.58–0.89 and 0.47–0.75, respectively (Table 1).

The mean age (SD) of healthy controls was 35.9 (6.7) years. The frequency of spinal squaring for each vertebra in the control group was as follows (Table 2): the 2nd (lower border 100%), the 3rd (upper border 85.2%–91.0%, lower border 91.0%–93.4%), the 4th (upper border 44.3%–49.2%, lower border 57.4%–63.1%), the 5th (upper border 5.0%–6.6%, lower border 8.4%–15%), the 6th (upper border 2.7%–7.3%, lower border 4.7%–8.6%), the 7th cervical vertebra (upper border 39.7%–45.6%, lower border 51.5%–54.1%), and the 1st thoracic vertebra (upper border 56.0%–62.5%).

The mean age (SD) in the AS group was 35.0 (8.7) years; mean disease duration was 15 (3.4) years. The frequency of spinal squaring in this group was (Table 2): the 2nd (lower border 100%), the 3rd (upper border 82.0%–82.7%, lower border 90.2%–91.7%), the 4th (upper border 63.2%–68.0%, lower border 77.4%–81.1%), the 5th (upper border 20.5%–25.6%, lower border 31.1%–36.8%), the 6th (upper border 21.7%–27.3%, lower border 28.3%–35.9%), the 7th cervical vertebra (upper border 62.7%–66.4%, lower border 68.3%–69.8%), and the 1st thoracic vertebra (upper border 82.9%–89.1%).

Consistently, the differences in squaring between the AS and control groups were statistically significant for the 4th, 5th, 6th, and upper border of 7th cervical vertebra. A high proportion of normal 4th and 7th cervical vertebra was scored as squared.

Table 1. Kappa values for intra- and interobserver reliability. Poor kappa < 0.40; fair to good 0.40 ≤ kappa ≤ 0.75; excellent kappa > 0.75.

	Intraobserver		Interobserver	
	Observer 1	Observer 2	I	II
Kappa (range)	0.71* (0.58–0.82)	0.82* (0.76–0.89)	0.64* (0.56–0.75)	0.57* (0.47–0.68)

* p < 0.001.

DISCUSSION

The most characteristic bony changes in AS are growth of new bone and formation of syndesmophytes[1], possibly leading to ankylosis and spinal fusion. Measuring structural damage is crucial to assess the potential of drugs used in AS to modify disease progression. Several radiographic scoring methods have been used: the Bath AS Radiologic Index (BASRI)[6], the Stoke AS Spinal Score (SASSS)[7], and the mSASSS. After testing the radiographic scoring methods of all 3 aspects of the Outcome Measurement in Rheumatology Clinical Trial (OMERACT) filter, mSASSS seems to be the most appropriate method for scoring radiologic progression[14]. The mSASSS system generates a score for both the lumbar and cervical spine, and the total score is the sum of both. Scores are 0 for

Table 2. Comparison of squaring in control and AS groups. The table shows frequencies of each spinal squaring examined by 2 observers. Consistently, differences in squaring between the AS and normal groups were statistically significant for the 4th, 5th, 6th, and upper border of 7th cervical vertebra. However, a high proportion of 4th and 7th cervical vertebra on controls were scored as squared.

	Observer 1				Observer 2			
	AS (%)	Control (%)	p	OR (95% CI)	AS (%)	Control (%)	p	OR (95% CI)
C2 lower	122 (100)	121 (100)	NS	NS	133 (100)	122 (100)	NS	NS
C3 upper	100 (82.0)	104 (85.2)	NS	NS	110 (82.7)	111 (91.0)	NS	NS
lower	110 (90.2)	111 (91.0)	NS	NS	122 (91.7)	114 (93.4)	NS	NS
C4 upper	83 (68.0)	60 (49.2)	< 0.01	2.19 (1.30–3.70)	84 (63.2)	54 (44.3)	< 0.01	2.16 (1.31–3.57)
lower	99 (81.1)	77 (63.1)	< 0.01	2.51 (1.40–4.51)	103 (77.4)	70 (57.4)	< 0.01	2.55 (1.48–4.39)
C5 upper	25 (20.5)	8 (6.6)	< 0.01	3.64 (1.57–8.44)	34 (25.6)	6 (5.0)	< 0.001	6.58 (2.65–16.33)
lower	38 (31.1)	10 (8.4)	< 0.001	4.93 (2.32–10.47)	49 (36.8)	18 (15)	< 0.001	3.31 (1.79–6.10)
C6 upper	26 (21.7)	3. (2.7)	< 0.001	9.87 (2.90–33.64)	36 (27.3)	8 (7.3)	< 0.001	4.78 (2.12–10.80)
lower	34 (28.3)	5 (4.7)	< 0.001	8.07 (3.02–21.52)	47 (35.9)	9 (8.6)	< 0.001	5.97 (2.76–12.90)
C7 upper	69 (62.7)	31 (39.7)	< 0.01	2.55 (1.41–4.63)	77 (66.4)	36 (45.6)	< 0.01	2.35 (1.31–4.24)
lower	69 (68.3)	40 (54.1)	NS	NS	67 (69.8)	34 (51.5)	< 0.05	2.17 (1.13–4.16)
T1 upper	34 (82.9)	20 (62.5)	NS	NS	41 (89.1)	14 (56.0)	< 0.01	6.44 (1.90–21.79)

NS: not significant. C: cervical vertebra, T: thoracic vertebrae.

normal; 1 for an erosion, sclerosis, or squaring; 2 for a syndesmophyte; and 3 for total bony bridging, as developed by Taylor, *et al*[15]. However, although squaring is an important factor for scoring it is often difficult to assess whether squaring is present. A potential solution would be measuring each cervical vertebra to determine the presence or absence of squaring compared to a normal population.

To verify the reliability of our measurements, inter- and intraobserver agreements were assessed based on kappa coefficient. Kappa statistics are appropriate for testing whether agreement exceeds chance levels for binary and nominal ratings, and the kappa values indicate the proportion of agreement beyond that expected by chance. In this study, good intra- and fair interobserver agreement were found in the assessment of squaring.

As a result of this research, we reached several conclusions. First, it is already known that the 3rd cervical vertebra is excluded from squaring score due to the straight shape of its lateral surface[8]. Therefore, it makes sense to exclude 3rd cervical vertebra. Second, there were inconsistent statistical results in the 7th lower cervical and 1st upper thoracic vertebra. In addition, it is difficult to obtain good quality radiographic images of the cervico-thoracic junction area in AS due to bone tissue contrast and anatomical distortion, which cause a lot of missing data. Third, in spite of statistically significant differences in squaring of the 4th, 5th, 6th, and 7th cervical vertebrae between cases with AS and controls, squaring was common on normal 4th and 7th cervical vertebrae, limiting its utility in distinguishing radiographic damage in AS from normal vertebrae. Then, it may be acceptable to exclude 4th and 7th cervical spine and 1st upper thoracic vertebrae while scoring "squaring" by mSASSS. Fourth, given odds ratios at each location, we note that the 5th and 6th cervical vertebrae have great influence on the discrimination between patients and controls. Moreover, there were consistently significant differences between AS and normal in squaring of the 5th and 6th cervical vertebrae.

We conclude it is reasonable to assess only the 5th and 6th cervical vertebrae for scoring squaring, rather than assessing entire cervical spines.

ACKNOWLEDGMENT

The authors thank Nam-Wook Hur, PhD, Epidemiologist, Biomedical Research Group, Hanyang University, for statistical support.

REFERENCES

1. Sieper J, Braun J, Rudwaleit M, Boonen A, Zink A. Ankylosing spondylitis: an overview. Ann Rheum Dis 2002;61 Suppl 3:iii8-18.
2. Ward MM. Functional disability predicts total costs in patients with ankylosing spondylitis. Arthritis Rheum 2002;46:223-31.
3. Brandt J, Khariouzov A, Listing J, et al. Six-month results of a double-blind, placebo-controlled trial of etanercept treatment in patients with active ankylosing spondylitis. Arthritis Rheum 2003;48:1667-75.
4. Zochling J, van der Heijde D, Dougados M, Braun J. Current evidence for the management of ankylosing spondylitis: a systematic literature review for the ASAS/EULAR management recommendations in ankylosing spondylitis. Ann Rheum Dis 2006;65:423-32.
5. Braun J, Brandt J, Listing J, 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.
6. MacKay K, Mack C, Brophy S, Calin A. The Bath Ankylosing Spondylitis Radiology Index (BASRI): a new, validated approach to disease assessment. Arthritis Rheum 1998;41:2263-70.
7. Averns HL, Oxtoby J, Taylor HG, Jones PW, Dziedzic K, Dawes PT. Radiological outcome in ankylosing spondylitis: use of the Stoke Ankylosing Spondylitis Spine Score (SASSS). Br J Rheumatol 1996;35:373-6.
8. Creemers MC, Franssen MJ, van 't Hof MA, Gribnau FW, van de Putte LB, van Riel PL. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. Ann Rheum Dis 2005;64:127-9.
9. Salaffi F, Carotti M, Garofalo G, Giuseppetti GM, Grassi W. Radiological scoring methods for ankylosing spondylitis: a comparison between the Bath Ankylosing Spondylitis Radiology Index and the modified Stoke Ankylosing Spondylitis Spine Score. Clin Exp Rheumatol 2007;25:67-74.
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. Mau W, Zeidler H, Mau R, Majewski A, Freyschmidt J, Deicher H. Outcome of possible ankylosing spondylitis in a 10 years' follow-up study. Clin Rheumatol 1987;6 Suppl 2:60-6.
12. Ralston SH, Urquhart GD, Brzeski M, Sturrock RD. A new method for the radiological assessment of vertebral squaring in ankylosing spondylitis. Ann Rheum Dis 1992;51:330-3.
13. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 1977;33:363-74.
14. Wanders AJ, Landewe RB, Spoorenberg A, et al. What is the most appropriate radiologic scoring method for ankylosing spondylitis? A comparison of the available methods based on the Outcome Measures in Rheumatology Clinical Trials filter. Arthritis Rheum 2004;50:2622-32.
15. 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.