Editorial

Spinal Mobility Measures in Spondyloarthritis: Application of the OMERACT Filter

Disease activity in ankylosing spondylitis (AS) is measured by multiple measures including metrology, radiology, laboratory, functioning, and disability[1]. Spinal mobility assessment is advantageous, as loss of mobility can be an early feature and is utilized for classification criteria[2]. Loss of spinal mobility has also been reported to be a prognostic factor[3]. Significant therapeutic advances have changed the way that patients with AS are treated and have shown the potential for disease modification[4]. The Assessment in Ankylosing Spondylitis (ASAS) group has recommended that spinal mobility measures be used as part of the disease controlling antirheumatic therapy criteria[5].

Over the past 50 years, multiple spinal mobility measures and composite scoring systems have been developed and are available for clinical application. Frequently these measurements have not been standardized or assessed for reliability, validity, or sensitivity to change[6]. Moreover, performance of some measures is time-consuming and may not be feasible in clinical practice. Therefore the goal would be to identify a minimum number of individual clinically appropriate measures to accurately assess the axial manifestation of disease. The purpose of this review is to assess the measures that have been used to assess spinal disease in terms of their validity and conformity with the OMERACT filter.

OMERACT FILTER

The OMERACT (Outcome Measures in Rheumatoid Arthritis Clinical Trials) process involves consensus on outcome measures and is based on the "OMERACT filter," composed of 3 key components: truth, discrimination, and feasibility[7]. *Truth* addresses the issues of face, content, construct, and criterion validity — does the instrument measure what it is intended to? Is the result unbiased and relevant? *Discrimination* addresses issues of reliability (inter and intraobserver) and sensitivity to change. *Feasibility* addresses whether the measure can be easily applied or used with regard to time, money, training, scoring, interpretability, and acceptance by physician and patient. Table 1 lists the various areas that have been assessed. This review deals only with those measures that have been tested and reported in the literature (Table 2).

CERVICAL SPINE

Cervical spine movement includes lateral flexion, rotation, and forward flexion.

Lateral flexion of the neck using an inclinometer was found to correlate moderately with radiological change (r = 0.57), with excellent inter- and intraobserver reliability [intraclass correlation coefficient (ICC) 0.90 and 0.96, respectively][8]. A decrease in distance between the tragus and the coronoideus process of the clavicle in maximal lateral bending of the head to the left and right measured by a tape was found to correlate better with radiological change (r = 0.60) and had excellent intra- and interobserver reliability (ICC 0.97 and 0.98, respectively)[8].

Cervical rotation has been tested in several ways. Using a Myrin inclinometer, full rotation from left to right was measured[8], with moderate correlation with radiological change (r = 0.41), and with excellent inter- and intraobserver reliability (ICC 0.98 and 0.96, respectively). With a tape measure, the change in distance between the chin and the coronoideus process in maximal cervical rotation from left to right has been found to correlate better with radiological change (r = 0.57) and have excellent inter- and intraobserver reliability (ICC 0.97 and 0.98, respectively)[8]. Another method using a tape measure assesses the difference in the distance between a mark in the suprasternal notch and the tragus of the right ear when the neck is rotated from left to right[9]. This method provided excellent intraobserver (ICC 0.80 and 0.89) and interobserver reliability (ICC 0.82). The Bath Ankylosing Spondylitis Metrology Index (BASMI)[10] measures cervical

Table 1. Types of metrology measures described/studied.

Area Assessed	Measure
Cervical spine	Rotation
	Lateral bending
	Flexion
	Occiput to wall
	Tragus to wall
	Chin to chest
Thoracolumbar spine	Lateral flexion
	Forward flexion
	Rotation
	Chest expansion
	Vital capacity
	Finger to floor
	Schober
Shoulder mobility	Flexion
	Abduction
Hip mobility	Intermalleolar distance

rotation using a gravity-action goniometer. The mean of the right and left results is calculated and the results are scored as 0 if > 70°, 1 if 20°–70°, and 2 if < 20°. Interobserver reliability was 0.98 and intraobserver reliability 0.99. Thus, cervical rotation reflects changes in axial disease in AS; however, measurements that require specific instruments may be difficult to perform.

Chin to chest: This measure of forward flexion of the neck was found to correlate poorly with radiological disease, thus its relevance to the disease process (truth) is questionable. It was, however, found to be reliable, with inter- and intraobserver agreement measured by ICC of 0.92 and 0.95, respectively[8]. It has not been included as an outcome measure in clinical trials, so sensitivity to change is not available.

Occiput to wall: The distance measured between the occiput and the wall when the patient stands with heels and shoulder against the wall with the back straight has been used to assess patients for many years and is thought to correlate with kyphosis. It was correlated with radiographic change in the study by Viitanen, *et al*[8] (r = 0.49). Interobserver reliability was high (r = 0.92). It also proved to be sensitive to change, with an effect size of 0.25.

Tragus to wall: The distance between the tragus of the ear and the wall is measured when the patient stands with heels and buttocks touching the wall, knees straight, shoulders back, and places the head as far back as possible, keeping the chin in. Correlation with radiographs was high (r = 0.92), and interobserver and intraobserver reliability both were very high (r = 0.99)[10].

Occiput to wall and tragus to wall were studied in a longitudinal observational cohort, and both were comparably reliable[11]. The authors recommended using the occiput to wall because a value of zero clearly distinguishes between patients with normal thoracic spine extension and kyphotic patients.

Lumbar mobility. Forward flexion.

Schober's test: In the original report in 1937[12] the subject stood upright and lumbosacral junction was identified by a skin mark. Another mark was made 10 cm above this and the distance between the 2 was recorded when the patient bent maximally forward. It was thought to reflect the movement in the lumbar spine, but criticism has been that areas of spine used are more susceptible to superficial stretching of the skin and not actual underlying structures, particularly since there was no anatomical definition of the first mark. No studies of reliability or validity were carried out initially. A study by Macrae and Wright[13] tested the original Schober test against their modification (which added an additional mark 5 cm below the first mark), and showed that both the original Schober and their modification reflected lumbar forward flexion (confirmed by radiographs). Their modified Schober correlated with radiographic changes better than the original test (r = 0.97 vs 0.90, respectively)[13]. Testing the original Schober method, a misplacement of 2 cm at the level of the lumbosacral junction against radiographs resulted in an error of up to 15°, whereas the modified Schober method reduced the error to less than 5°.

Modified Schober test: Moll and Wright further refined the Macrae and Wright modification of the Schober test by providing an anchor for the lumbosacral junction using a line across the upper limit of the dimples of Venus[14]. The subject was then asked to bend maximally forward, and the new distance between the upper and lower marks was measured. The distraction of this mark has been found to correlate very closely with anterior flexion measured radiologically (r = 0.97)[13]. Moll and Wright demonstrated a moderate correlation between spinal mobility and duration of symptoms in patients with AS and no correlation with age[15].

The modified Schober is included as one of the measures of the BASMI. However, in the BASMI the first mark is made as a line across the iliac crests, which is at the level of L4 rather than at the lumbosacral junction, which is marked by a line across the dimples of Venus. Unfortunately, the method was not revalidated against radiographs as the original and its initial modification have been. Within the BASMI this modification of the Schober test had criterion validity against 20 other items included in the index (r = 0.92). Interobserver as well as intraobserver reliability was demonstrated (r = 0.96 and 0.99, respectively). Viitanen, *et al*[16] used the Macrae and Wright modification of the Schober test and found that it provided excellent inter- and intrarater reliability (ICC 0.96 and 0.94, respectively), and correlated highly with the radiographic changes in the spine (ICC 0.71). They also demonstrated the modified Schober to be sensitive to change following intensive physiotherapy (effect size 0.24). However, another study showed that the Schober test did not change following this therapy[17]. Moreover, although infliximab was proven to be effective in AS, the Schober test did not distinguish between patients treated with active drug or with placebo[18]. It should be noted that the Schober test used in this trial was based on the BASMI method.

Feasibility: The Schober test and its modification are easy to perform, and scores are easy to record and interpret, with no special equipment required; minimal training is required and patient and physician acceptance is good. Thus, this test passes the OMERACT filter of truth, discrimination (only for the original modification), and feasibility.

Lumbar flexion. Forward flexion of the spine using finger to floor distance was included among the measures studied by Viitanen, *et al*[8]. They found that it had good reliability and sensitivity, but did not correlate with radiographic changes. Heikkilä, *et al*[17] demonstrated that finger to floor distance improved following physiotherapy in patients with AS.

Lumbar flexion/extension. A method to simultaneously assess the range of flexion and extension of the spine was described by Miller, *et al*[19]. This test uses a line across the dimples of Venus as the first landmark, from which three 10-cm segments are marked with the back fully flexed. The change in the dis-

Table 2. Reliability, sensitivity to change and feasibility in instruments used to assess spinal disease in ankylosing spondylitis.

Item	Interobserver	Intraobserver	Sensitivity to Change	"Truth"	Feasibility
Cervical spine					
Lateral flexion					
Inclinometer	0.90	0.96	NA	0.57 vs xray	Special tool
Tape measure 1	0.97	0.98	NA	0.60 vs xray	Tape measure
Rotation					
Myrin	0.98	0.96	NA	0.41 vs xray	Special tool
Inclinometer	0.97	0.98	NA	0.57 vs xray	Pen, tape measure
Tape measure 1	0.80, 0.89	0.82	0.19	—	Pen, tape measure
Tape measure 2	0.98	0.99	0.21	—	Special tool
Action goniometer					
Forward flexion					
Chin-chest	0.92	0.95	NA	Poor vs xray	Tape measure
Occiput to wall	0.92	NA	ES = 0.25	0.49 vs xray	Tape measure
Tragus to wall	0.99	0.99	ES = 0.1	0.92 vs xray	Tape measure
Lumbar spine					
Forward flexion					
Modified	0.96	0.94		0.97 vs xray	Pen, tape measure
Schober	0.96	0.99	ES = 0.24	0.97 vs xray	Pen, tape measure
Finger to floor	0.98	0.98	NA	NA	Pen, tape measure
Lateral bending					
Thoracic spine					
Chest expansion	0.85	0.95	ES = 0.42	0.35	Tape measure
Intermalleolar distance	0.99	0.98	ES = 0.15	NA	Tape measure
Hip internal rotation	0.98	0.88	ES = 0.06	NA	Tape measure
Composite indices					
BASMI	0.96	0.99	ES = 0.18	0.83	7 min to perform
EDASMI	0.94	0.98	ES = 0.27	0.75	5-7 min to perform

ES: effect size. NA: not applicable.

tance between these points is noted with the back fully extended. It was compared with finger to floor distance, modified Schober test, and a goniometric method, showing excellent correlation and low interobserver error. This was the most sensitive method for detecting loss of spinal mobility, with the upper segment being most sensitive. Thus, it may be more relevant to patients with early spondylitis than the Schober test.

Lateral bending. Several methods have been developed to measure lateral bending of the lumbar spine. Moll and Wright used a measurement of the change in the distance of 2 marks inked on the skin of the lateral trunk between upright position and lateral flexion. The upper mark was placed at the point where a horizontal line through the xiphisternum crossed the coronal line, and the second mark where a horizontal line through the highest point of the iliac crest crossed the coronal line[15]. Although the authors demonstrated differences between patients with AS and normal controls, the method was felt to be cumbersome. Subsequent methods used the distance between the tip of the third finger and the floor when the patient stands upright, heels, buttocks, and shoulders against the wall, and bends sideways without lifting the opposite foot off the ground. This has been measured either by tape measure secured on the wall, or by making a mark on the thigh and leg[10,16]. The first method has been included in the BASMI[10], and the latter method in the Edmonton AS Metrology Index (EDASMI)[9]. Both methods provided excellent reliability. Both methods are relatively easy to perform (requiring the observer to bend to the ground several times) using only a tape measure and require minimal training.

Chest expansion. This is measured by the difference in chest circumference between full expiration and inspiration at the fourth intercostal space[8,16]. Chest expansion did not correlate with radiographic changes (r = 0.38). However, chest expansion provided interobserver as well as intraobserver reliability (ICC 0.85 and 0.95, respectively). Sensitivity to change was noted for chest expansion, with an effect size of 0.42. Although chest expansion did not correlate with radiographic change, it did provide reliability and discrimination. This measure is included in the BASMI and the ASAS response criteria.

Intermalleolar distance. Intermalleolar distance measures abduction of the hips. Keeping the knees straight and the legs in contact with the resting surface the patient is asked to take the legs as far apart as possible, and the distance between the medial malleoli is measured. This measure had criterion validity as part of the BASMI compared to 20 clinical measurements (r = 0.92)[10]. Intermalleolar distance provided excellent inter- and intraobserver variability (r = 0.98 and 0.99, respectively). Along with the 4 other measures of BASMI it takes a total of 7 minutes to complete. Thus this measure passes the OMERACT filter requirement, but is cumbersome to perform in daily practice.

COMPOSITE METROLOGY SCORING SYSTEMS OF SPINAL MOBILITY

BASMI[10]. This index was based on a study of 20 measurements performed on 43 patients. From this total metrology exercise, 5 simple clinical measurements were defined that most accurately reflected axial status; they included cervical rotation (goniometer), tragus to wall distance, lateral lumbar flexion, modified Schober's, and intermalleolar distance. The BASMI was tested against the 20 clinical measurements in these patients and an additional 54 patients, with excellent agreement. Interobserver reliability was excellent when 3 physiotherapists examined the same patients (r = 0.96 and 0.99, respectively). It should be noted, however, that the BASMI scores are based on a 0–2 score scale for each of the items, with 0 representing normal mobility, 1 mild to moderate reduction, and 2 severe reduction. A total score of 10 may be achieved if the mobility is restricted severely in all 5 measurements. The BASMI showed sensitivity to change with significant differences noted in treatment groups in the infliximab trial[19]. However, the Cohen effect size calculated was only 0.29.

EDASMI[9]. The EDASMI includes 4 measures: cervical rotation (by tape measure), chest expansion, lateral lumbar flexion, and hip internal rotation. The score is based on percentiles, 0 representing higher than 80th percentile, 1 the 60th–80th percentile, 2 the 40th–60th percentile, 3 the 20th–40th percentile, and 4 lower than 20th percentile, allowing a total score of 16. Interobserver and intraobserver reliability were measured in 44 patients by a nurse clinician and a rheumatologist, with excellent agreement (ICC 0.94–0.98 for the total EDASMI and for the individual components). Interestingly, both the EDASMI and BASMI correlated very well with measures of structural damage, and poorly with the Bath AS Disease Activity Index. A standardized response mean of 0.44 was noted using the EDASMI in patients undergoing new therapies.

SUMMARY

A number of spinal mobility measurements have been developed and validated in patients with AS. Several pass the OMERACT filter of truth, discrimination, and feasibility. It is now important to include these measures in therapeutic trials both individually and within indices to determine which combination may be the most useful for (1) routine clinical use, (2) outcome and prognostic studies, and (3) evaluation of new therapies. Moreover, it is important to test these measurements in patients with spondylitis other than ankylosing spondylitis.

JOHN C. DAVIS Jr, MD, MPH,
Associate Professor of Medicine,
Director, Clinical Trials Center,
Division of Rheumatology,
University of California,
San Francisco, California, USA;
DAFNA D. GLADMAN, MD, FRCPC,
Professor of Medicine,
University of Toronto,
Senior Scientist, Toronto Western Research Institute,
Deputy Director, Centre for Prognosis Studies in the Rheumatic Diseases,
Toronto, Ontario, Canada

Address reprint requests to Dr. D. Gladman, Toronto Western Hospital, 1E-410B, 399 Bathurst Street, Toronto, Ontario M5T 2S8.
E-mail: dafna.gladman@utoronto.ca

REFERENCES

1. Rigby AS, Silman AJ. Outcome assessment in clinical trials of ankylosing spondylitis. Br J Rheumatol 1991;30:321-2.
2. 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.
3. Amor B, Santos RS, Nahal R, Listrat V, Dougados M. Predictive factors for the longterm outcome of spondyloarthropathies. J Rheumatol 1994;21:1883-7.
4. Braun J, Baraliakos X, Brandt J, Sieper J. Therapy of ankylosing spondylitis. Part II: biological therapies in the spondyloarthritides. Scand J Rheumatol 2005;34:178-90.
5. van der Heijde DMFM, Calin A, Dougados M, Khan MA, van der Linden S, Bellamy N. Selection of instruments in the core set for DC-ART, SMARD, physical therapy, and clinical record keeping in ankylosing spondylitis. Progress report of the ASAS Working Group. Assessments in Ankylosing Spondylitis. J Rheumatol 1999;26:951-4.
6. Pile KD, Laurent MR, Salmond CE, Best MJ, Pyle EA, Moloney RO. Clinical assessment of ankylosing spondylitis: a study of observer variation in spinal measurements. Br J Rheumatol 1991;30:29-34.
7. Boers M, Brooks P, Strand CV, et al. The OMERACT filter for outcome measures in rheumatology. J Rheumatol 1998;25:198-9.
8. Viitanen JV, Kautiainen H, Suni J, Kokko ML, Lehtinen K. The relative value of spinal and thoracic mobility measurements in ankylosing spondylitis. Scand J Rheumatol 1995;24:94-97.
9. Maksymowych WP, Mallon C, Richardson R, et al. Development and validation of the Edmonton Ankylosing Spondylitis Metrology Index. Arthritis Rheum 2006;55:575-82.
10. Jenkinson TR, Mallorie PA, Whitelock HC, Kennedy LG, Garrett SL, Calin A. Defining spinal mobility in ankylosing spondylitis (AS). The Bath AS Metrology Index (BASMI). J Rheumatol 1994;21:1694-8.
11. Heuft-Dorenbosch L, Vossse D, Landewé R, et al. Measurement of spinal mobility in ankylosing spondylitis: comparison of occiput-to-wall and tragus-to-wall. J Rheumatol 2004;31:1779-84.
12. Schober P. The lumbar vertebral column and backache. Munch Med Wschr 1937;84:336.
13. Macrae IF, Wright V. Measurement of back movement. Ann Rheum Dis 1969;28:584-9.
14. Moll JMH, Wright V. Normal range of spinal mobility. An objective clinical study. Ann Rheum Dis 1971;30:381-6.

15. Moll JMH, Wright V. The pattern of chest and spinal mobility in ankylosing spondylitis. An objective clinical study of 106 patients. Rheumatol Rehabil 1973;12:115-34.
16. Viitanen JV, Heikkilä S, Kokko M-L, Kautiainen H. Clinical assessment of spinal mobility measurements in ankylosing spondylitis: a compact set for follow-up and trials? Clin Rheumatol 2000;19:131-7.
17. Heikkila S, Viitanen JV, Kautiainen H, Kauppi M. Sensitivity to change of mobility tests; effect of short term intensive physiotherapy and exercise on spinal, hip, and shoulder measurements in spondyloarthropathy. J Rheumatol 2000;27:1251-6.
18. van der Heijde DMFM, Dijkmans B, Geusens P, et al. Efficacy and safety of infliximab in patients with ankylosing spondylitis: results of a randomized, placebo-controlled trial (ASSERT). Arthritis Rheum 2005;52:582-91.
19. Miller MH, Lee P, Smythe HA, Goldsmith CH. Measurements of spinal mobility in the sagittal plane: new skin contraction technique compared with established methods. J Rheumatol 1984;11:507-11.