International Spondyloarthritis Interobserver Reliability Exercise — The INSPIRE Study: I. Assessment of Spinal Measures

DAFNA D. GLADMAN, ROBERT D. INMAN, RICHARD J. COOK, DESIRÉE van der HEIJDE, ROBERT B.M. LANDEWÉ, JURGEN BRAUN, JOHN C. DAVIS, PHILIP MEASE, JOACHIM BRANDT, RUBEN BURGOS VARGAS, VINOD CHANDRAN, PHILIP HELLIWELL, ARTHUR KAVANAUGH, FINBAR D. O'SHEA, MUHAMMAD A. KHAN, NICOLO PIPITONE, PROTON RAHMAN, JOHN D. REVEILLE, MILLICENT A. STONE, WILLIAM TAYLOR, DOUGLAS J. VEALE, and WALTER P. MAKSYMOWYCH

ABSTRACT. ***Objective.*** To determine whether the axial measures used in primary ankylosing spondylitis (AS) were reproducible for both AS and psoriatic arthritis (PsA) with axial disease.

Methods. A group of 20 rheumatologists from 11 countries with expertise in spondyloarthritis (SpA) met for a combined physical examination exercise to assess 10 patients with PsA with axial involvement (9 men, 1 woman, mean age 52 yrs, mean disease duration 17 yrs) and 9 AS patients (7 men, 2 women, mean age 38 yrs, mean disease duration 16 yrs). A modified Latin-square design was used. Measures included were occiput to wall, tragus to wall, cervical rotation, chest expansion, lateral spinal bending, modified Schober, and hip mobility. Data were analyzed using intraclass correlation coefficients (ICC) adjusted for order of measurements.

Results. The majority of the variance was contributed by the patients. There was no order effect. Observer effect was noted especially for chest expansion for both AS and PsA patients, and for the modified Schober in PsA. The ICC demonstrated very good to excellent agreement for most measures for both AS and PsA. Chest expansion provided only moderate agreement for AS and PsA.

Conclusion. Overall, measures of spinal mobility used in primary AS perform well with respect to interobserver reliability, and are equally reproducible when applied to PsA patients with axial involvement. Thus, these measures should now be evaluated in therapeutic trials of patients with PsA to determine sensitivity to change and concordance with other measures of structural damage. (First Release July 1 2007; J Rheumatol 2007;34:1733–9)

Key Indexing Terms:
SPONDYLOARTHRITIS ASSESSMENT RELIABILITY SPINAL

Ankylosing spondylitis (AS) is an inflammatory joint disease affecting primarily the joints of the spine and the sacroiliac joints[1]. Arthritis in the peripheral joints occurs in 10%–35% of patients with AS[2-5]. Psoriatic arthritis (PsA) is an inflammatory joint disease associated with psoriasis[6]. Up to 50% of patients with PsA have inflammatory back disease, which

From the Division of Rheumatology, University of Toronto, Toronto, Ontario, Canada.

Supported by the SPARCC NRI grant from The Arthritis Society, Canada, the Krembil Foundation, and unrestricted funds from Abbott Canada, Amgen/Wyeth Canada, Schering Canada, Pfizer Canada, and Wyeth Global. R.J. Cook is Canada Research Chair in Statistical Methods for Health Research. W.P. Maksymowych is a Senior Scholar of the Alberta Heritage Foundation for Medical Research. V. Chandran is an Ogryzlo Fellow of The Arthritis Society, Canada, and Arthritis Centre of Excellence Fellow, University Health Network, Toronto. F.D. O'Shea is supported by an Irish Society for Rheumatology Traveling Fellowship. M.A. Stone is supported by the Cumming AS Visiting Professorship Award.

D.D. Gladman, MD, FRCPC; R.D. Inman, MD; V. Chandran, MD; F.D. O'Shea, MD, Division of Rheumatology, University of Toronto, Toronto, Ontario, Canada; R.J. Cook, MD, Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada; D. van der Heijde, MD, PhD, Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands; R.B.M. Landewé, MD, PhD, Department of Rheumatology, University Hospital Maastricht, Maastricht, The Netherlands; J. Braun, MD, Rheumazentrum, Ruhrgebiet, Herne, Germany; J.C. Davis, MD, Department of Rheumatology, Allergy and Immunology, University of California, San Francisco, California, USA; P. Mease, MD, Swedish Medical Center, Seattle, Washington, USA; J. Brandt, MD, Rheumatologische, Praxisgemeinschaft, Berlin, Germany; R. Burgos Vargas, MD, Rheumatology Unit, Hospital General de Mexico, Mexico City, Mexico; P. Helliwell, MD, PhD, University of Leeds, Leeds, United Kingdom; A. Kavanaugh, MD, Department of Rheumatology, Allergy and Immunology, University of California, San Diego, California, USA; M.A. Khan, MD, Department of Medicine, Case Western Reserve University, Cleveland, Ohio, USA; N. Pipitone, MD, U.O. di Reumatologia, Arcispedale S. Maria Nuova, Reggio Emilia, Italy; P. Rahman, MD, Department of Rheumatology, Memorial University, St. Johns, Newfoundland, Canada; J.D. Reveille, MD, University of Texas Health Sciences Centre, Houston, Texas, USA; M.A. Stone, MD, Department of Rheumatology, University of Toronto, Toronto, Royal National Hospital for Rheumatic Diseases, Bath, United Kingdom; W. Taylor, MBChB, FRACP, FAFRM, Rehabilitation Teaching and Research Unit, Wellington School of Medicine and Health Sciences, University of Otago, Wellington, New Zealand; D.J. Veale, MD, Rheumatology, St. Vincent's University Hospital, Dublin, Ireland; W.P. Maksymowych, FRCPC, Department of Rheumatology, University of Alberta, Edmonton, Alberta, Canada.

Address reprint requests to Dr. D.D. Gladman, Toronto Western Hospital, University Health Network, 399 Bathurst Street, 1E410B, Toronto, Ontario M5T 2S8, Canada. E-mail: dafna.gladman@utoronto.ca

Accepted for publication April 20, 2007.

resembles AS in that it presents with sacroiliitis and syndesmophyte formation. Both AS and PsA are classified as forms of spondyloarthritis (SpA) as they share not only the spinal involvement but also several other extraarticular features such as enthesitis, as well as the association with HLA-B*27. Clinical and genetic differences between AS and PsA have been documented[7]. Radiological differences in spinal disease have also been documented[7,8]. In general, patients with AS have more severe spondylitis, while patients with PsA have more severe peripheral arthritis. Patients with PsA do not have as much pain and limitation of movement in the spine as do patients with AS[7], which may be related to the less severe structural change observed in patients with PsA. While the peripheral joint assessment in PsA has been shown to be reliable both in a single clinic study[9] and in the SpA Research Consortium Canada (SPARCC) study[10], systematic evaluation of spinal measurements in PsA has not been performed[11].

At OMERACT 7 (Outcome Measures in Rheumatology; Asilomar, California, May 2004) the PsA workshop identified domains to be included in clinical trials and observational cohort studies in PsA[12]. In addition to the assessment of peripheral joint disease, participants recognized the importance of spinal involvement in PsA. Thus, spinal assessment was included among the proposed domains to be included in clinical trials and longitudinal cohort studies in PsA. However, the instruments to assess this domain were not identified. Indeed, the randomized controlled trials in PsA to date have included peripheral arthritis[13-18], but the assessment of spinal involvement, dactylitis, and enthesitis has not been incorporated into most clinical trials for lack of validated tools.

Several measures have been included in the assessment of spinal disease in AS. The Bath Ankylosing Spondylitis Metrology Index (BASMI) was developed to clinically assess patients with AS[19]. This index includes 5 clinical measures: tragus to wall distance, measured as the distance between the tragus of the ear and the wall; lumbar flexion, measured by a modification of the Macrae and Wright modification of the Schober test[20]; cervical rotation, measured by the ability of the patient to rotate the neck from side to side; lumbar side-flexion, measured as the difference between the distance from the third finger and the floor when the patient stands straight and when the patient bends sideways with the knees straight; and intermalleolar distance, measured as the distance between the 2 malleoli when the patient lies down with the hips fully abducted. A more recent index, the Edmonton AS Metrology Index (EDASMI)[21], measures cervical rotation using a tape measure, chest expansion at the xiphisternum, lateral bending of the spine, by placing a pen mark on the thigh at the extremes of movement, and internal rotation of the hip, by measurement of intermalleolar distance with the patient seated separating the ankles as far as possible. The ASsessment in Ankylosing Spondylitis Study Group (ASAS) has defined the core outcome measures for AS[22]. For spinal mobility these include chest expansion, modified Schober test (line across superior posterior iliac spines with another mark 10 cm above it), occiput to wall distance, and lateral spinal bending or BASMI. The latter 2 were added in an update of the core set[23]. Several other measures have been developed to record mobility in the sagittal, frontal, and coronal planes of the spine, with varying degrees of practicability and reproducibility (as reviewed[24,25]).

A PsA module was included at OMERACT 8 (May 2006), which aimed to identify a core set of domains as well as to define the instruments to be used to assess the domains. The purpose of this study was to begin the validation of spinal assessment among patients with PsA with spinal involvement, using a group of AS patients for comparison. Specifically, we sought to determine whether the axial measures used in primary AS were reproducible for both AS and PsA patients with axial disease.

MATERIALS AND METHODS

Patient selection. Nine patients with AS with varying degrees of mobility were selected from the AS clinic at the Toronto Western Hospital, Toronto, Canada. All met the modified New York criteria for the classification of AS[26]. They included 7 men and 2 women with a mean age of 38 years and mean disease duration 16 years. All patients with AS had evidence of spondylitis. Ten patients with PsA and radiological evidence of spinal disease including at least grade 2 sacroiliitis, 7 of whom also had spondylitis, were selected from the PsA Clinic in the same institution to reflect variable extent of joint inflammation and spinal mobility. These included 9 men and 1 woman with a mean age 52 years and mean disease duration 17 years.

Observers. A group of 20 rheumatologists from 11 countries with expertise in SpA, including 10 with expertise in AS and 10 with expertise in PsA, participated in this study. These were members of the ASAS group, GRAPPA (Group for Research and Assessment of Psoriasis and Psoriatic Arthritis), SPARCC, and SPARTAN (SpondyloArthritis Research and Treatment Network).

Clinical assessments. Patients and observers were divided into 2 groups, each including both AS and PsA patients. Within each group all patients were evaluated by the same 10 observers. The following clinical measurements were performed by each observer on each patient.

Cervical mobility. Occiput to wall distance measured from the occiput to the wall with the patient standing straight back and heels against the wall, looking straight forward with the chin at the usual carrying level and with a maximum effort to touch the head to the wall. Tragus to wall distance measured in the same position as above, both using a tape measure to the nearest 0.1 cm[27]. For each measurement the better of 2 attempts was recorded. Cervical rotation, measured with a goniometer in a modification of the BASMI method, averaging the 2 sides[19]. In the BASMI method a gravity goniometer was used. In the current study the patients were examined sitting on a chair, observer standing behind, with the goniometer anchored against the examiner's chest. The subject was asked to turn the head to the right and then the left and the measurements were recorded. The average of the 2 measurements was recorded. This method was selected to facilitate the measure in settings where a gravity goniometer was not available. In addition, a tape measure method was introduced[28]. With the patient sitting on the examination couch a pen mark was made in the suprasternal notch. Looking directly ahead, the patient was asked to rotate the head as far as possible towards the right shoulder. The distance between the pen mark in the suprasternal notch and the tragus of the right ear was recorded in centimeters using a tape measure (D1). The subject was then asked to rotate the head as far as possible towards the left shoulder. The distance between the pen mark in the suprasternal notch and the tragus of

the right ear was recorded with the tape measure (D2). Total cervical rotation was recorded as the difference between the 2 measurements (D1 – D2). This method has been used in the development and validation of the EDASMI[21].

Chest expansion. Chest expansion was measured as the difference in centimeters to the nearest 0.1 cm between full expiration and full inspiration. Repeat measurements were taken, one at the level of the fourth intercostal space, and the other at the level of the xiphisternum.

Spinal mobility. Spinal forward flexion was measured by a modification of the modified Schober test[20]. With the patient standing straight, a mark was made at the upper level of the dimples of Venus, and another mark 10 cm above it. The patient was then asked to bend forward with the knees fully extended and the distance between the 2 marks minus 10 was recorded. Lateral bending was performed in the following 3 ways. In each case the subject was asked to stand straight with the back, heels, buttocks, and shoulders against the wall, hands by the side, all the fingers straight against the thigh, and looking directly forward. The subject was then asked to slide the hand down the right thigh as far as possible without lifting the left foot or flexing the knees, and maintaining a straight posture with heels, buttocks, and shoulders against the wall. (1) The BASMI[19] method in which the difference between the distance from the tip of the third finger and the floor at rest and after maximal lateral bending is measured with a tape measure. (2) The Domjan method[29], where a pen mark is made on the thigh adjacent to the tip of the middle finger. A second pen mark is made adjacent to the tip of the middle finger at maximum lateral flexion. The difference between the 2 marks is measured using a tape measure. The maneuver is repeated on the left side and the mean of the 2 measurements recorded as the final score. (3) A modification of the Domjan method, denoted the INSPIRE method, was introduced at the meeting. A mark was made on the right side when the patient was bending maximally to the left, in addition to the mark made on the right when the patient was bending maximally to the right, and the distance between lateral flexion to the left and right was recorded as the distance between the 2 marks on the right side. This reflects both right and left lateral bending.

Hip mobility. Intermalleolar distance recorded with the patient supine as the distance between the 2 medial malleoli when the patient spreads the legs apart as far as they can[19]. For internal rotation of the hip, the subject was seated on an examining couch with the knees and hips flexed, the knees together against the couch clasping a piece of card paper. The subject was asked to move the ankles apart as far as possible without releasing the card paper between the knees. The distance between the medial malleoli was recorded using a tape measure[21].

BASMI. Both the original BASMI[19] and the BASMI-10, which provides a new scoring system based on a 0–10 scale[30], were calculated based on the clinical assessments.

EDASMI. The original EDASMI[21] and a modification replacing the cervical rotation by goniometer were tested[31].

Design. A Latin-square design was proposed for two 10 by 10 grids. This method allows determination of the observer, patient, and order effects at the same time. A modification of the Latin-square design was used since one patient with AS did not attend the session, leaving 9 patients with AS and 10 with PsA, while all 20 observers attended. Before beginning the clinical assessments the participants discussed the measures to be performed and reviewed the method of assessment for each measure, with a demonstration by one of the participants (WPM) of the spinal measures on a healthy subject, but there was not a formal training session.

Statistical analysis. Variance components analyses were conducted for continuous measurements based on ANOVA models with random observer, random patient, and both including and excluding fixed order effects. Estimates of intraclass correlation coefficients (ICC) and associated 95% confidence intervals were obtained[32]. Sackett, *et al*[33] suggest that values of kappa in the intervals (0.80, 1.00) represent excellent agreement beyond chance, (0.60, 0.80) substantial agreement, (0.40, 0.60) moderate agreement, (0.20, 0.40) fair agreement, and (0.0, 0.20) poor agreement beyond chance. For the purpose of interpreting our results we adopted this same classification for the ICC.

RESULTS

The measurements in patients with AS and PsA are presented in Table 1. The median and range for all observers for all patients are shown. The results of the reliability of the measurements are presented in Tables 2 to 4. Table 2 provides ICC values for the various spinal measures for patients with AS and PsA. Tables 3 and 4 provide information on the sources of variance in these patient groups.

Cervical mobility. As shown in Table 2, both occiput to wall and tragus to wall distance provided excellent agreement among the observers for patients with AS. Tragus to wall was not quite as reproducible in patients with PsA as the occiput to wall, with only substantial agreement for tragus to wall and excellent agreement for occiput to wall.

Cervical rotation using the goniometer provided moderate agreement in patients with AS, and excellent agreement in PsA. The tape measure method used in the EDASMI[21] was not reproducible in patients with AS, but provided substantial agreement for patients with PsA in the INSPIRE study.

Chest expansion. Chest expansion at the level of the xiphisternum was more reliable than at the fourth intercostals in AS, providing moderate agreement, but in PsA both measurements gave similar substantial agreement.

Lumbar mobility. The modified Schober test provided excellent agreement in patients with AS and substantial agreement in PsA. Lateral bending methods all gave excellent agreement in both AS and PsA.

Hip mobility. Intermalleolar distance similarly gave substantial agreement in AS and excellent agreement in PsA. Internal rotation of the hip gave excellent agreement in both AS and PsA.

BASMI and EDASMI. We determined the reliability of the BASMI and EDASMI. As shown in Table 2 the ICC for the original BASMI was 0.92 for the AS group and 0.89 for the PsA group, whereas the original EDASMI provided ICC of 0.66 for the AS group and 0.85 for the PsA patients. This is likely related to the differences in neck measurements. The BASMI is based on the goniometer method, whereas the EDASMI is based on the tape measure method, which in our study did not provide good agreement for AS. Indeed, when we repeated the analysis for EDASMI using a goniometer reading, the ICC increased to 0.82 for AS and 0.91 for PsA. We further analyzed the use of the modification of the BASMI on a 10-point scale. This provided a lower ICC for AS of 0.59, whereas the ICC for PsA remained essentially unchanged at 0.90.

Variance. Table 3 provides the degree of variance noted among patients with AS. As can be seen, the major source of variation was the patients. The observers contributed to the variance in occiput to wall distance, cervical rotation by goniometer, and chest expansion at the fourth intercostals space. The order of examination did not contribute to the variation.

Table 1. Spinal measurements in patients with AS and PsA in the INSPIRE study.

Measurement, median (range)	AS	PsA
Occiput to wall, cm	6.9 (0.0, 15.1)	5.0 (0.0, 16.3)
Tragus to wall, cm	17.4 (11.2, 25.6)	14.8 (11.8, 25.5)
Cervical rotation – average, degrees	51.0 (34.2, 76.6)	57.2 (7.7, 81.9)
Cervical rotation – left, degrees	52.8 (38.0, 78.2)	53.2 (7.7, 84.2)
Cervical rotation – right, degrees	47.9 (30.3, 75.2)	56.9 (7.7, 79.6)
Cervical rotation, cm	4.1 (0.9, 5.0)	1.7 (0.3, 4.2)
Chest expansion, at xiphisternum	3.4 (1.1, 5.6)	4.0 (3.1, 9.2)
Chest expansion, 4th intercostal space	3.2 (1.7, 4.4)	3.6 (1.7, 7.6)
Modified Schober, cm	2.8 (0.6, 5.5)	3.9 (1.2, 6.0)
Lateral flexion – BASMI, cm	13.3 (4.0, 16.8)	14.4 (5.9, 22.7)
Lateral flexion – Domjan method, cm	12.2 (3.6, 17.3)	14.2 (5.1, 22.9)
Lateral flexion – INSPIRE, cm	25.2 (8.0, 34.1)	28.4 (10.2, 42.9)
Intermalleolar distance, cm	101.0 (49.2, 132.8)	101.6 (58.0, 126.2)
Internal rotation hip, cm	38.0 (17.8, 48.2)	30.8 (21.2, 54.9)

Table 2. ICC for spinal measurements by disease.

Measurement	AS	PsA with Axial Disease
Occiput to wall	0.81 (0.64, 0.94)	0.82 (0.66, 0.94)
Tragus to wall	0.80 (0.63, 0.94)	0.67 (0.46, 0.88)
Cervical rotation		
Goniometer	0.66 (0.43, 0.88)	0.94 (0.87, 0.98)
Tape measure	0.04 (–0.05, 0.32)	0.77 (0.59, 0.92)
Chest expansion		
Xiphisternum	0.57 (0.33, 0.84)	0.64 (0.42, 0.87)
4th intercostal	0.24 (0.07, 0.59)	0.70 (0.50, 0.89)
Modified Schober	0.85 (0.70, 0.95)	0.65 (0.43, 0.87)
Lateral spinal bending		
BASMI	0.83 (0.67, 0.95)	0.82 (0.66, 0.94)
Domjan	0.91 (0.82, 0.98)	0.85 (0.71, 0.95)
INSPIRE	0.92 (0.83, 0.98)	0.77 (0.58, 0.92)
Intermalleolar distance	0.78 (0.60, 0.93)	0.96 (0.91, 0.99)
Internal rotation hip	0.92 (0.82, 0.98)	0.92 (0.83, 0.97)
BASMI	0.92 (0.82, 0.98)	0.89 (0.77, 0.96)
BASMI-10	0.60 (0.36, 0.85)	0.90 (0.80, 0.97)
EDASMI tape measure	0.69 (0.47, 0.90)	0.85 (0.71, 0.95)
EDASMI goniometer	0.82 (0.65, 0.95)	0.91 (0.82, 0.97)

BASMI: Bath AS Metrology Index, EDASMI: Edmonton AS Metrology Index.

Table 4 provides the sources of variance in patients with PsA. Again, patients provided the major source of variance. Assessor effect was noted for occiput to wall distance, cervical rotation by the goniometer, and chest expansion at both the xiphisternum and fourth intercostals space. Order effect was noted only for the intermalleolar distance.

Using the BASMI and EDASMI to express the results of the assessments, again the major source of variation was the patients. However, there was some assessor contribution to variation with the original BASMI in PsA and with the EDASMI in both AS and PsA (Tables 3 and 4).

DISCUSSION

This is the first study to assess the reliability of the measurements used in the assessment of AS among patients with PsA with spinal involvement. Moreover, in this study an international group of assessors was gathered to assure that measurements are reproducible by experts from different countries, and thus with different education and healthcare systems. Some of the measures have been tested in AS, but with a small number of assessors, and generally with examiners from the same unit. The INSPIRE study thus provides evidence for the reliability of the measures in an international context, with 10 observers evaluating 2 groups of patients. We found that the measures provide substantial to excellent agreement among observers, with the major source of variation being the patients, and with little observer and order effect for most measures.

Looking at the specific measures, occiput to wall distance has been questioned because of the variability in technique and the concern about a floor effect since the normal is 0. The tragus to wall technique was chosen for the BASMI primarily because it was thought to be less susceptible to measurement error caused by flexion/extension movements of the head. It has previously been shown that the occiput to wall and tragus to wall measurements were both equally reliable[27]. A similar observation was made by the INSPIRE investigators. One concern among the INSPIRE investigators was that the tragus to wall does not have the immediately recognizable normal value (of zero), and it is very dependent on the patient's body structure, such that heavier individuals will have greater tragus to wall distance.

Cervical rotation has been included as a measure in the BASMI method[19,30] as an indication of cervical spine involvement with AS or PsA. In the BASMI-defined technique a gravity-action goniometer was used with the patient lying supine. In the INSPIRE study we used a modification in which the patient is seated, and a regular goniometer was used. This method provided substantial agreement in patients with AS,

Table 3. Variation in patients with AS.

Measure	% Total Variance (p value) Due to Patient	Assessor	Order
Occiput to wall	81.8 (< 0.001)	8.0 (0.003)	0.1 (0.98)
Tragus to wall	79.8 (< 0.001)	3.7 (0.66)	1.9 (0.13)
Cervical rotation			
Goniometer	68.1 (< 0.001)	14.3 (0.008)	0.1 (0.98)
Tape measure	13.7 (0.21)	20.9 (0.39)	3.1 (0.64)
Chest expansion			
Xiphisternum	58.5 (< 0.001)	26.2 (< 0.001)	2.5 (0.06)
4th intercostal space	29.3 (< 0.001)	40.1 (< 0.001)	1.5 (0.56)
Modified Schober	84.9 (< 0.001)	5.3 (0.04)	0.7 (0.38)
Lateral bending			
BASMI	83.3 (< 0.001)	3.9 (0.43)	0.4 (0.73)
Domjan method	91.4 (< 0.001)	2.3 (0.24)	0.5 (0.30)
INSPIRE	91.5 (< 0.001)	1.6 (0.56)	1.0 (0.07)
Intermalleolar distance	78.4 (< 0.001)	5.9 (0.19)	1.2 (0.31)
Internal rotation hip	91.5 (< 0.001)	2.0 (0.34)	0.6 (0.24)
BASMI	91.5 (< 0.001)	2.7 (0.09)	0.4 (0.34)
BASMI-10	60.9 (0.001)	11.3 (0.18)	1.1 (0.65)
EDASMI tape measure	69.4 (< 0.001)	11.2 (0.02)	1.9 (0.19)
EDASMI goniometer	82.3 (< 0.001)	7.7 (0.003)	0 . 5 (0.54)

Table 4. Variation in patients with PsA.

Measure	% Total Variance (p value) Due to Patient	Assessor	Order
Occiput to wall	83.0 (< 0.001)	7.7 (0.001)	0.1 (0.97)
Tragus to wall	68.0 (< 0.001)	10.0 (0.04)	1.7 (0.24)
Cervical rotation			
Goniometer	94.0 (< 0.001)	2.8 (0.001)	3 (0.20)
Tape measure	78.9 (< 0.001)	4.0 (0.60)	0.2 (0.96)
Chest expansion			
Xiphisternum	66.2 (< 0.001)	17.3 (< 0.001)	2.2 (0.10)
4th intercostal space	72.1 (< 0.001)	15.7 (< 0.001)	0.2 (0.87)
Modified Schober	67.1 (< 0.001)	18.3 (< 0.001)	1.0 (0.36)
Lateral bending			
BASMI	82.0 (< 0.001)	2.6 (0.82)	1.1 (0.27)
Domjan method	84.4 (< 0.001)	5.1 (0.01)	1.4 (0.04)
INSPIRE	75.8 (< 0.001)	4.5 (0.46)	2.7 (0.04)
Intermalleolar distance	95.3 (< 0.001)	1.2 (0.04)	0.9 (< 0.001)
Internal rotation hip	91.3 (< 0.001)	2.2 (0.15)	0.8 (0.08)
BASMI	89.0 (< 0.001)	4.2 (0.004)	0.3 (0.59)
BASMI-10	90.3 (< 0.001)	2.0 (0.43)	0.3 (0.52)
EDASMI tape measure	85.7 (< 0.001)	3.6 (0.21)	0.3 (0.76)
EDASMI goniometer	91.5 (0.001)	3.0 (0.008)	0.5 (0.20)

and excellent agreement among patients with PsA. Two tape measure methods to assess cervical rotation have been studied previously. Viitanen, *et al*[34] used the distance between the chin and the coronoideus process of the clavicle at extreme cervical rotation and found it to be reliable. Another tape measure method included in the EDASMI[18] measures the difference in the distance between the sternal notch and the tragus of the right ear at right and left extreme rotation of the neck. It proved reliable in the EDASMI validation and in a subsequent study[28] of patients with AS. In INSPIRE, the tape measure method was not reliable for patients with AS, although in PsA patients it demonstrated substantial agreement. This may be because the patients with AS had much more severe restriction of their neck movement that prevented adequate measurement using the tape measure method. Indeed, no patient effect was observed with this method, although there was also no assessor or order effect. It appears that it was easier for observers to perform the goniometer method, although it requires an additional instrument, and for short observers it was necessary to use a stool. Alternatively, more training may be required for the tape measure method to make it reliable in studies that include a large number of observers.

Chest expansion was not included among the measurements of the BASMI. However, it has been recommended as an outcome measure by ASAS. Chest expansion measured at the fourth intercostals space was shown to be reproducible (ICC 0.85) by 2 physiotherapists in a previous study of 52 patients with AS[34], but was less reproducible (ICC 0.40) in the SPARCC study, which included patients with PsA, only half of whom had back involvement. In the current INSPIRE study, chest expansion at both the xiphisternum and fourth intercostals space provided substantial agreement in PsA, but only moderate agreement (ICC 0.57) in AS patients at the level of the xiphisternum. Chest expansion at the fourth intercostals space was not reliable in patients with AS. The reason for this discrepant result may include the severity of thoracic involvement in patients with AS included in this study, resulting in much reduced chest expansion. Indeed, the variation due to patients was much lower among patients with AS than patients with PsA. While it is not clear why there would be a larger observer effect, there may have been a lack of effort on the part of the patients. It seems that in patients with AS the xiphisternum provides a better level to measure chest expansion, whereas in PsA patients with spinal involvement either level produces similar results.

Forward flexion of the spine has been included in both the BASMI and the ASAS core set. According to the Macrae and Wright modification of the Schober test, one places an anchor at the L5–S1 level (dimples of Venus) from which marks at 10 cm above and 5 cm below are placed with the patient standing upright. This 15-cm segment should increase to at least 21 cm when the patient bends forward, although this increase diminishes with age. The instructions for the BASMI method suggest the first mark be put at the level of the iliac crest, which is the L4/L5 level. In the ASAS studies only the 10-cm segment was included. The Schober test was found to have an ICC > 0.70 in 2 studies of AS patients[35], but both methods provided very low ICC (ICC 0.27) in the SPARCC study of patients with PsA[10]. In the INSPIRE study, the Schober test provided excellent agreement in AS (ICC 0.85) and substantial agreement in PsA with spinal involvement (ICC 0.65). It is interesting that the Schober test functioned well in patients with inflammatory spinal disease, since in these patients the thoracolumbar junction is usually the first to be affected and

the lumbar spinal limitation is late. However, it should be noted that disease duration in these patients was 16–17 years. It remains to be determined whether the Schober test would function as well in patients with early disease.

Lateral bending of the spine has been included as a measure of spinal disease in the BASMI, and it is also included in the ASAS core set. Lateral bending reflects thoracic as well as lumbar spine mobility. It demonstrated excellent agreement in the SPARCC study (ICC 0.83). In INSPIRE, 3 methods to determine lateral bending were included, the original BASMI method, the Domjan method, and the new method developed by INSPIRE investigators. All 3 methods provided excellent agreement in both AS and PsA. The INSPIRE method may be preferred since it requires only one set of measurements on the part of the assessor, and divided into 2 provides the same information as the average of right and left lateral bending required for the BASMI and Domjan methods.

Intermalleolar distance is included among the BASMI measures. It proved extremely reliable in the SPARCC study, and in patients with PsA in the INSPIRE study. In AS the intermalleolar distance agreement was substantial. Again, the differences between patients with AS and PsA are noted, in that there was much more variance due to patient effect in the PsA group. Internal rotation of the hip as measured in this study also proved reliable in patients both with AS and PsA.

We further analyzed whether the BASMI and EDASMI were reliable in this study. Both indices were equally reliable in patients with PsA. In patients with AS the original BASMI produced higher ICC than EDASMI, likely due to the differences in the measurement of cervical rotation included in these indices. This result is different from the study reported by Maksymowych, *et al*[21], which showed similar ICC for both the BASMI and the EDASMI. In that study, the tape measure method of cervical rotation was found to be much more reliable than in the INSPIRE study. Indeed, when the goniometer measurements were substituted for the tape measure in the EDASMI, much higher ICC were obtained in AS, with equally good, that is, excellent results in PsA. The BASMI-10 was not as good in AS, with lower ICC and wider confidence intervals than the original BASMI, whereas in PsA both produced similar results. However, the new calculation of EDASMI produced better results than the one using a tape measure. Thus, both the clinical assessments and the indices derived from them work well in both AS and PsA.

The INSPIRE study thus confirms that measures of spinal mobility are reliable for AS. It further demonstrated that measures of spinal mobility that have been applied to primary AS perform well with respect to interobserver reliability when applied to PsA patients with axial involvement.

Thus, these measures may now be included in therapeutic trials of patients with PsA, to test their ability to respond to therapy and to complete their validation for these patients.

ACKNOWLEDGMENT

The authors acknowledge Ker-Ai Lee for helping with the statistical analysis; Catherine Schentag for organizing the study and database support; Anne MacKinnon for administrative support; Dr. Sergio Toloza, Fawnda Pellett, Ruby Syal, Sutharshini Shanmugarajah, and Sue Gilchrest for organizational support.

REFERENCES

1. Van der Linden S, van der Heijde D, Braun J. Ankylosing spondylitis. In: Harris ED, Budd RC, Firestein GS, et al, editors. Kelley's textbook of rheumatology. 7th ed. Philadelphia: W.B. Saunders; 2005:1125-41.
2. Helliwell PS, Wright V. Ankylosing spondylitis. In: Bellamy N, editor. Prognosis in the rheumatic diseases. London: Kluwer Academic Publishers; 1991:133-52.
3. Carette S, Graham D, Little H, Rubenstein J, Rosen P. The natural disease course of ankylosing spondylitis. Arthritis Rheum 1983;26:186.
4. Calin A. Ankylosing spondylitis. Clin Rheum Dis 1985;11:41-60.
5. Khan MA. Ankylosing spondylitis. In: Calin A, editor. Spondyloarthropathies. Orlando: Grune & Stratton; 1984:69-118.
6. Gladman DD. Psoriatic arthritis. In: Harris ED, Budd RC, Firestein GS, et al, editors. Kelley's textbook of rheumatology. 7th ed. Philadelphia: W.B. Saunders; 2004:1155-64.
7. Gladman DD, Brubacher B, Buskila D, Langevitz P, Farewell VT. Differences in the expression of spondyloarthropathy: a comparison between ankylosing spondylitis and psoriatic arthritis. Genetic and gender effects. Clin Invest Med 1993;16:1-7.
8. Helliwell PS, Hickling P, Wright V. Do the radiological changes of classic ankylosing spondylitis differ from the changes found in the spondylitis associated with inflammatory bowel disease, psoriasis, and reactive arthritis? Ann Rheum Dis 1998;57:135-40.
9. Gladman DD, Farewell V, Buskila D, et al. Reliability of measurements of active and damaged joints in psoriatic arthritis. J Rheumatol 1990;17:62-4.
10. Gladman DD, Cook RJ, Schentag C, et al. The clinical assessment of patients with psoriatic arthritis: results of a validation study of the SpondyloArthritis Research Consortium of Canada (SPARCC). J Rheumatol 2004;31:1126-31.
11. Gladman DD, Helliwell P, Mease PJ, Nash P, Ritchlin C, Taylor W. Assessment of patients with psoriatic arthritis. A review of currently available measures. Arthritis Rheum 2004;50:24-35.
12. Gladman DD, Mease P, Krueger G, et al. OMERACT 7 Workshop. Outcome measures in psoriatic arthritis. J Rheumatol 2005;32:2262-9.
13. Mease PJ, Goffe BS, Metz J, et al. Etanercept in the treatment of psoriatic arthritis and psoriasis: a randomised trial. Lancet 2000;356:385-90.
14. Kaltwasser JP, Nash P, Gladman D, et al, for the TOPAS Study Group. Efficacy and safety of leflunomide in the treatment of psoriatic arthritis and psoriasis. Arthritis Rheum 2004;50:1939-50.
15. Mease PJ, Kivitz AJ, Burch FX, et al. Etanercept treatment of psoriatic arthritis: safety, efficacy, and effect on disease progression. Arthritis Rheum 2004;50:2264-72.
16. Antoni C, Krueger GG, de Vlam K, et al. Infliximab improves signs and symptoms of psoriatic arthritis: results of the IMPACT 2 trial. Ann Rheum Dis 2005;64:1150-7.
17. Antoni CE, Kavanaugh A, Kirkham B, et al. Sustained benefits of infliximab therapy for dermatologic and articular manifestations of psoriatic arthritis: results from the Infliximab Multinational Psoriatic Arthritis Controlled Trial (IMPACT). Arthritis Rheum 2005;52:1227-36.
18. Mease PJ, Gladman DD, Ritchlin CT, et al, for the ADEPT Study Group. Adalimumab in the Treatment of Patients with Moderately

to Severely Active Psoriatic Arthritis: Results of the ADEPT Trial. Arthritis Rheum 2005;52:3279-89.
19. Jenkinson TR, Mallorie PA, Whitelock HC, Kennedy LG, Garrett SL, Calin A. Defining spinal mobility in ankylosing spondylitis (AS). The Bath AS Metrology Index. J Rheumatol 1994;21:1694-8.
20. Macrae IF, Wright V. Measurement of back movement. Ann Rheum Dis 1969;28:584-9.
21. Maksymowych WP, Mallon C, Richardson R, et al. Development and validation of the Edmonton Ankylosing Spondylitis Metrology Index (EDASMI). Arthritis Rheum 2006;55:575-82.
22. 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.
23. Zochling J, Braun J, van der Heijde D. Assessments in ankylosing spondylitis. Best Pract Res Clin Rheumatol 2006;20:521-37.
24. Helliwell PS, Moll JMH, Wright V. Measurement of spinal movement and function. In: Jayson MIV, editor. The lumbar spine and back pain. London: Churchill Livingstone; 1992:173-206.
25. Davis JC, Gladman DD. Spinal mobility measures in spondyloarthritis: Application of the OMERACT filter. J Rheumatol 2007;34:666-70.
26. 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.
27. Heuft-Dorenbosch L, Vosse 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.
28. Maksymowych WP, Mallon C, Richardson R, et al. Development and validation of a simple tape-based measurement tool for recording cervical rotation in patients with ankylosing spondylitis: comparison with a goniometer-based approach. J Rheumatol 2006;33:2243-8.
29. Domjan L, Nemes T, Balint GP, Toth Z, Gomor B. A simple method for measuring lateral flexion of the dorsolumbar spine. J Rheumatol 1990;17:663-5.
30. Jones SD, Porter J, Garrett SL, Kennedy LG, Whitelock H, Calin A. A new scoring system for the Bath Ankylosing Spondylitis Metrology Index (BASMI). J Rheumatol 1995;22:1609.
31. Maksymowych WP, Mallon C, Morrow S, Hutchings E. A comparison of the discriminant capacity of the BASMI (0-10) and EDASMI mobility instruments in AS patients receiving anti-TNFα treatment. Ann Rheum Dis 2007; (in press).
32. Burdick RK, Graybill FA. Confidence intervals on variance components. Statistics: a series of textbooks and monographs. New York: Marcel Dekker; 1992:127.
33. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Clinical epidemiology: a basic science for clinical medicine. 2nd ed. Toronto: Little, Brown and Company; 1991.
34. Viitanen JV, Heikkila 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.
35. 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.