diagnostics

MDPI

Article

Measuring Spinal Mobility Using an Inertial Measurement Unit System: a Reliability study in Axial Spondyloarthritis

Megan O'Grady [1]*, Tom O'Dwyer [2], James Connolly [3], Joan Condell [4], Karla Munoz Esquivel [4], Finbar D. O'Shea [5], Philip Gardiner [6], Fiona Wilson [1]

[1] Discipline of Physiotherapy, Trinity College Dublin, Ireland; ogradyme@tcd.ie, wilsonf@tcd.ie
[2] Independent Researcher, D08 W9RT Dublin, Ireland; odwyertk@tcd.ie
[3] Letterkenny Institute of Technology, Letterkenny, F92 FC93 Donegal, Ireland; james.connolly@lyit.ie
[4] Ulster University, Derry BT48 7JL, UK; j.condell@ulster.ac.uk (J.C.), kc.munoz-esquivel@ulster.ac.uk (K.M.E.)
[5] Rheumatology Department, St James's Hospital, D08 NHY1 Dublin, Ireland; FOShea@stjames.ie
[6] Western Health and Social Care Trust, Londonderry BT47 6SB, UK; Philip.Gardiner@westerntrust.hscni.net
* Correspondence: ogradyme@tcd.ie

Received: date; Accepted: date; Published: date

Abstract: The objectives of this study were to evaluate the reliability of wearable inertial motion unit (IMU) sensors in measuring spinal range of motion under supervised and unsupervised conditions in both laboratory and ambulatory settings. A secondary aim of the study was to evaluate the reliability of composite IMU metrology scores (JMU-ASMI (Amb)). Forty people with axSpA participated in this clinical measurement study. Participant spinal mobility was assessed by conventional metrology (Bath Ankylosing Spondylitis Metrology Index, linear version – $BASMI_{Lin}$) and by a wireless IMU sensor-based system which measured lumbar flexion-extension, lateral flexion and rotation. Each sensor-based movement test was converted to a normalized index and used to calculate JMU-ASMI (Amb) scores. Test-retest reliability was evaluated using intra-class correlation coefficients (ICC). There was good to excellent agreement for all spinal range of movements (ICC > 0.85) and JMU-ASMI (Amb) scores (ICC > 0.87) across all conditions. Correlations between JMU-ASMI (Amb) scores and conventional metrology were strong (Pearson correlation ≥ 0.85). An IMU sensor-based system is a reliable way of measuring spinal lumbar mobility in axSpA under supervised and unsupervised conditions. While not a replacement for established clinical measures, composite JMU-ASMI (Amb) scores may be reliably used as a proxy measure of spinal mobility.

Deleted: IMU-ASMI

Deleted: IMU-ASMI

Deleted: IMU-ASMI

Deleted: IMU-ASMI

Deleted: IMU-ASMI

Keywords: axial spondyloarthritis; spinal mobility; inertial measurement unit; reliability

1. Introduction

Axial spondyloarthritis (axSpA) is a complex chronic inflammatory disease predominantly affecting the axial skeleton [1]. In the early stages of the disease, restriction in spinal mobility is mainly due to reversible inflammation in and around the spine, but in later stages the restriction becomes permanent due to structural bony damage [2-3]. Monitoring of individuals with axSpA should center on aspects of the disease that cause symptoms or functional disability [4], and which are subject to change as the disease progresses or treatment is introduced, such as decreased spinal mobility [5-6].

Spinal mobility has been recognized as an important outcome in the management of axSpA and has been included in the Assessment of Spondyloarthritis International Society (ASAS) core set for

clinical assessment in axSpA [6-7]. The Bath Ankylosing Spondylitis Metrology Index (BASMI) is a well-established method of measuring spinal movement in axSpA [8]. While the BASMI is a low-cost tool with minimal training and equipment required, it cannot be performed independently, limiting its utility outside of the clinical setting. It also lacks the sensitivity to change required to monitor disease progression [9-11]. As a result of these concerns, the BASMI failed to achieve approval by the ASAS group as a core outcome measure in axSpA [5]. There is a rapidly growing role for telemedicine as a tool to improve care for individuals with rheumatic disease, although there is a recognition that limitations in technology need to be understood and addressed to achieve standards of care consistent with existing in-person services [12]. There is therefore a need for a reliable and sensitive measure of spinal mobility to be used in studies of drug and physical interventions in axSpA.

Video-based optoelectronic systems are often thought of as the laboratory gold standard for human motion analysis [13-14]. These systems can provide complex descriptions of body segment motion but only measure movement over a short period of time. Their capture area can be limited by cameras, body markers and (environmental) equipment positioning, and they create artificial environments for movement assessment. Due to the high cost of equipment and training, they are therefore not feasible for many research centers or real-world testing.

In recent years, wearable inertial motion unit (IMU) sensor systems have advanced to the point of offering a viable method of clinically measuring spinal mobility [13-14] and analyzing spinal posture [15]. An IMU sensor typically incorporates a tri-axial accelerometer, gyroscope and magnetometer, and several can be incorporated unobtrusively as part of a wearable sensor system. Validity and reliability of such systems in the measurement of lumbar spine mobility has been established in healthy populations [16-17]. The validity and reliability of an IMU sensor-based system for evaluating cervical and lumbar spinal mobility in individuals with axSpA were recently established under supervised conditions [18-19]. If the full range of spinal mobility can be reliably measured in unsupervised ambulatory settings using IMU sensor-based systems, this tool could provide a viable method of reducing variability in measurement of spinal mobility, be sensitive to small changes in mobility over time, and be an important step towards digital self-management. It could be used by individuals with axSpA and clinicians involved in their care to reliably monitor signs remotely, providing clinicians with a "real-life" assessment of current disease state.

This study is the third in a project [18, 19] investigating wearable IMU sensors and composite metrology scores in individuals with axSpA, with a focus on reliability in the ambulatory setting. The primary aim of this study was to assess the reliability of spinal IMU sensors in measuring spinal mobility of individuals with axSpA. The objectives were to evaluate the reliability of spinal IMU sensors in measuring spinal range of motion 1) under supervised and unsupervised conditions in the exercise laboratory, and 2) under unsupervised conditions in an ambulatory setting. A secondary aim of the study was to evaluate the reliability of calculated composite IMU metrology scores (IMU-ASMI (Amb)) and to establish correlations with BASMI. The reader is advised to refer to Gardiner et al. [19], Aranda-Valera et al. [18], and to the supplementary material for a detailed explanation of the IMU-ASMI (Amb) score.

Deleted: IMU-ASMI

2. Materials and Methods

2.1. Study Design

This was a clinical measurement study with a specific focus on reliability. The study was approved by the local research ethics committee [REC Reference: 2017-10 List 37 (20)].

2.2. Participant Eligibility & Recruitment

Inclusion criteria for the study were as follows: diagnosis of axSpA (by ASAS criteria) made at least six months prior to recruitment to the study, age between 18 years and 80 years old and the ability to read and understand the English language. Exclusion criteria were severe joint or spinal pain at the time of the study, prior total hip arthroplasty or severely restricted hip movement, history of previous vertebral fracture, history of previous spinal surgery, severe scoliosis, spinal deformity

or complete segmental fusion of the spine, pregnancy or being unable to mobilize without assistance or mobility aid.

Participant selection was through convenience sampling. Potential participants attending a dedicated hospital-based axSpA clinic were informed of the study by a gatekeeper who was not part of the research team. Notice of the study was also circulated via the social media channels of the Ankylosing Spondylitis Association of Ireland and Arthritis Ireland, and sent to individuals who were on a register having expressed interest in taking part in research projects. Interested persons contacted the study team and were screened for eligibility over the phone or via email. Participant diagnosis was confirmed by letter from the participant's rheumatologist or general practitioner.

2.3. Data Collection and Baseline Assessments

Socio-demographic (age, sex and employment status) and anthropometric (weight, height and BMI) data were collected. Condition-specific data (time since onset of symptoms, time since diagnosis, medications and HLA-B27 status) were self-reported by participants. A battery of clinical questionnaires were self-completed by participants. These were: the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) [20], the Bath Ankylosing Spondylitis Functional Index (BASFI) [21], the Bath Ankylosing Spondylitis Global Score (BAS-G) [22], and the Ankylosing Spondylitis Quality of Life questionnaire (ASQoL) [23].

The ViMove™ wireless sensor kit (DorsaVi™, Melbourne, Australia) was used as an IMU sensor-based system to measure spinal range of movement. Members of the research team attended a half-day training course to ensure that sensor application and movement tests were carried out according to the manufacturer's standardized protocols. The ViMove system uses two IMU sensors to provide an absolute orientation estimation (roll, pitch, and yaw) and calculate the relative orientation in three planes (sagittal, frontal and transverse) by combining the measurements of both sensors. The sensors connect and transmit IMU data using radio frequency to a pocket recording device at a frequency of 20 Hz, from which data can be downloaded or viewed directly from a laptop (see Figure 1). The ViMove sensor setup was previously validated against both the Fastrak and Vicon motion sensor systems [17, 24]. Aranda-Valera et al [18] recently established the validity of the sensor setup in evaluating spinal mobility in an individual with axSpA using an optical motion capture system as a reference.

Deleted: using Bluetooth

(a) (b)

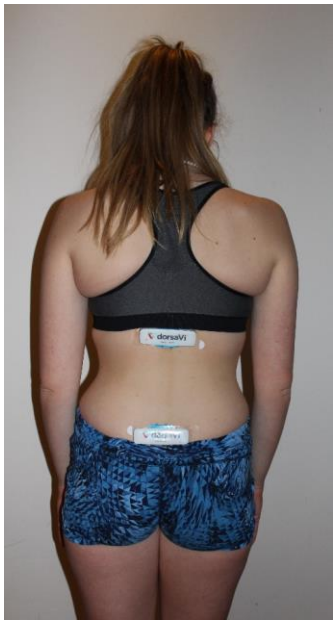

Figure 1. (**a**) ViMove sensor location; (**b**) Pocket recording device used by participant

2.4. Assessment Schedule

Eligible participants attended the test center for assessment on two consecutive days. A research physiotherapist (MOG) trained in assessing individuals with axSpA carried out clinical tests. Both assessments were at approximately the same time each day. The phase between the two appointments in the laboratory was a community-based ambulatory phase, during which participants were unsupervised. Table 1 summarizes the testing schedule.

Table 1. Study assessment schedule

	Day 1 - Laboratory	Ambulatory Phase	Day 2 - Laboratory
Baseline data collection	✓	--	--
$BASMI_{Lin}$ & chest expansion	✓	--	✓
Pain NRS & Fatigue NRS	✓	✓	✓
Spinal movement tests	Supervised & Unsupervised	Unsupervised	Supervised & Unsupervised

$BASMI_{Lin}$: Bath Ankylosing Spondylitis Metrology Index (linear version); NRS: numeric rating scale

On Day 1, socio-demographic data were recorded and anthropometric measurements were completed. Following a five minute warm-up (treadmill walking or stationary exercise bike depending on participant preference), chest expansion and spinal mobility using the linear versions of the Bath Ankylosing Spondylitis Metrology Index ($BASMI_{Lin}$) were recorded following ASAS guidelines [6, 25-26].

The two sensors were then attached to the participant according to the manufacturer's guidelines. The lower (sacral) sensor was positioned using a line drawn between the posterior superior iliac spines, and the upper lumbar (trunk) sensor was positioned above this line using DorsaVi™ designed height-specific templates to ensure the accurate positioning of the upper sensor over the T12 vertebra (see Fig. 1). Both sensors were mounted into a baseplate attached to an adhesive strip, which was placed directly on the skin. Calibration of the system was performed in relaxed standing (as per manufacturer's standardized protocol) and angles were recorded at the zero position for each IMU sensor to set the baseline position. Each sensor then calculated orientation angles with respect to this calibrated starting position.

Using standardized instructions, the assessor verbally guided the participants through a sequence of spinal movements: flexion, extension, lateral flexion (left then right), and rotation (left then right). Each movement was repeated three times before moving to the next movement (Condition 1: laboratory, supervised). Participants were then instructed to repeat the same sequence of spinal movements without supervision (Condition 2: laboratory, unsupervised). Participants followed either an instructional video (an example is included in the supplementary material) or written instructions (depending on preference); the same standardized instructions were used as during the supervised tests. The assessor left the room until all movements in the sequence were completed. Participants were instructed to press an 'event' button on the wireless pocket recorder when they were about to begin each movement, and again when they had completed the movement.

Participants left the exercise laboratory with the two IMU sensors *in situ*. During this ambulatory phase, participants repeated the spinal movement sequence at home by following video or written instructions and pressing the 'event' button on the wireless recorder (Condition 3: Ambulatory, unsupervised). The following day, participants returned to the exercise laboratory. The $BASMI_{Lin}$ and the same spinal movements were repeated under supervised and unsupervised conditions. As test sessions were at different times of day, the diurnal variation in symptoms was monitored by participants recording their levels of pain and fatigue on a numerical rating scale (Pain NRS from '0 - No pain' to '10 - Most severe' and Fatigue NRS from '0 - None' to '10 - Very severe') prior to and after completing the spinal movements [27].

2.5. Data Management

2.5.1. Sensor Data Output

Data was downloaded from sensors after each phase of testing using Microsoft Excel for Windows version 2009 (Microsoft Corporation, Washington, USA). The start and end of movement tests were identified using 'event' markers, and minimum and maximum degrees of movement were generated within each set of event markers. The data analyst visually inspected each movement test, and adjusted the start and end of the movement window if needed, to ensure they coincided with actual spinal movement. Each movement was repeated three times, and the maximum degree of movement was computed from the available repetitions (see Section 3.2). The mean of these degree of movement values was used in subsequent calculations. Output for rotation movements was only available under supervised conditions owing to technical limitations with the system. Output was designated as Trunk (from the upper sensor, the orientation angle from the upper lumbar sensor to the ground; represents lumbar and pelvic movement) and Lumbar (the angle between the upper lumbar sensor and the sacral sensor; represents lumbar movement). The 'full-arc' range of movement for a given spinal movement test was calculated. The reliability of full-arc movements has been shown to be higher than measurements from midline [19].

Minimum, maximum and range data were independently validated by examining the raw IMU sensor data for each test. A random selection of n=5 participant data samples (12.5% of all samples) were analysed using Microsoft Excel. The event markers corresponding with the start and end of each spinal movement test were again visually analysed by an independent reviewer (JC), and Excel-generated data values for each movement were numerically compared with the corresponding values generated by the ViMove software for each movement test. Results showed that there were no discrepancies between data generated by both methods for day 1 and day 2 of laboratory data. There was a comparison variation of 1.09 degrees within all ambulatory data samples. This was considered an acceptable amount of variation.

2.5.2. Calculation of composite metrology score (IMU-ASMI (Amb))

Normalized scales permit rapid evaluation of mobility, without the need for clinicians to know normal ranges of movement. Each sensor-based movement test (Flexion-Extension, Lateral Flexion L+R, and Rotation L+R) was converted into a normalized index using a formula based on that used to calculate $BASMI_{Lin}$ [26]. The 10th and 90th percentile ranges for each sensor-based movement test were obtained from research cohorts associated with this research group (Cordoba healthy controls, Altnagelvin AxSpA cohort). Normalized scores were calculated as follows: $((90th\ centile - A) / (90th\ centile - 10th\ centile)) / 10)$; A = range of motion in degrees). If A ≥ 90th centile, the normalized score = 0; if A ≤ 10th centile, the normalized score = 10. Composite IMU-ASMI (Amb) scores were calculated for the lumbopelvic region (Trunk-ASMI) and the lumbar region (Lumbar-ASMI). Trunk-ASMI and Lumbar-ASMI were calculated as the mean of the normalized scores of the lumbopelvic region and lumbar region, respectively. The reliability of regional composite indices has been shown to be superior to that of individual components [19]. The reader is advised to refer to the supplementary material for a detailed explanation of the IMU-ASMI composite metrology score.

2.6. Statistical Methods

Descriptive data are presented as frequencies and percentages for categorical variables, and continuous data were presented as mean and standard deviation, or median and interquartile range, as appropriate.

Test-retest reliability, compared across laboratory conditions (supervised and unsupervised) and ambulatory conditions, was evaluated using intra-class correlation coefficient (ICC) and standard error of measurement (SEM). Two-way, mixed-effects, single rater, absolute agreement model for ICCs were used. ICC interpretation was as follows: <0.5 = poor, 0.5 to 0.75 = moderate, 0.75 and 0.9 = good, > 0.90 = excellent [28]. The SEM was calculated as follows: $SEM=SD \times \sqrt{(1-ICC)}$, with SD representing the pooled (two measurements) SD of the measure. Agreement between movement tests under each condition was evaluated using Bland-Altman analysis. The mean bias and the limits

Deleted: For each movement, the mean of the

Deleted: .

Deleted:

Deleted: representing both lumbar and pelvic movement

Deleted: calculated based on the relative position of the upper and lower sensors

Deleted: Minimum and maximum

Deleted: This was caused by the incorrect identification of minimum and maximum degrees of movement for one participant by the ViMove software.

Deleted: IMU-ASMI

Deleted: IMU-ASMI

of agreement (LoA) were calculated to provide an estimate of the interval in which 95% of the differences between both test conditions are.

Correlations between $BASMI_{lin}$ and IMU-ASMI (Amb) scores under laboratory and ambulatory conditions were evaluated by Pearson correlation, which were interpreted as follows: values between 0.1 and 0.69 denoted weak to moderate correlation, values above 0.7 were regarded as a strong correlation [29]. Friedman's test, with post-hoc Wilcoxon Signed-rank tests, were used to evaluate the change in pain and fatigue NRS scores across test sessions. SPSS for Windows version 26 (IBM, New York, USA), MedCalc version 19.5.1 (MedCalc Software, Ostend, Belgium) and Microsoft Excel for Windows version 2013 (Microsoft Corporation, Washington, USA) were used for analysis.

Deleted: L

Deleted: IMU-ASMI

Deleted: 2009

3. Results

3.1. Recruitment and Participant Characteristics

Forty eligible participants were recruited to the study and completed the protocol between April 2018 and December 2018. Figure 2 illustrates the participant recruitment to the study. Twenty-five participants were male and 15 were female. Mean age of participants was 48.0 years (range 27 to 76) and mean symptom duration was 23.6 years (range 3 to 52). A range of disease severity is seen in the scores across clinical measures. Sixty-five percent of the participants were taking anti-TNFα medication. Participant baseline characteristics are summarized in Table 2.

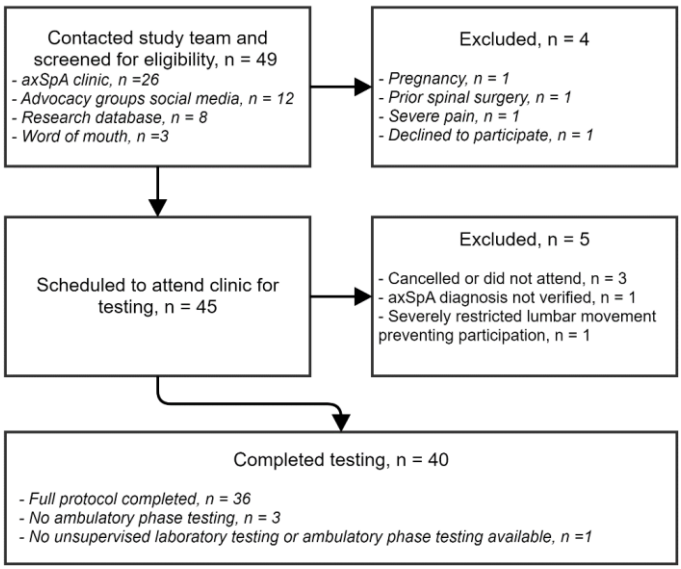

Figure 2. Participant recruitment

Table 2. Descriptive characteristics of study participants

Variable	AxSpA Cohort (n = 40)
Age, years	48.0 (12.9); [27 – 76]
Sex (male/female), *n*	25 / 15
Symptom duration, *years*	23.6 (13.7); [3 – 52]
Time since diagnosis, *years*	9.0 (26.5)*; [0-43]
BMI, kg/m^2	28.4 (7.5)*; [20.0 -37.7]

Employed, *n* (%)	23 (57.5)
BASMI$_{Lin}$[†]	3.8 (1.8); [1.2 – 7.9]
Lateral lumbar flexion	4.9 (2.5); [0 – 9.0]
Tragus to wall distance	2.4 (2.0); [0.5 – 7.6]
Modified Schober's test	5.4 (2.5); [0 – 9.7]
Intermalleolar distance	3.1 (1.9) ;[0 – 7.0]
Cervical rotation	3.3 (2.2); [0.3 – 9.5]
Chest expansion, *cm*	2.47 (2.2); [0.6 – 13.2]
Pharmacology, *n* (%)	
Anti-TNFα	26 (65)
NSAIDs	4 (10)
Analgesia	4 (10)
None	6 (15)
HLA-B27 status, *n* (%)	
Positive	17 (42.5)
Negative	7 (17.5)
Unknown	16 (40)
BAS-G, (scale 0-10)	3.4 (2.1) [0 - 7]
BASDAI[‡], (scale 0-10)	3.4 (2.0) [0 - 7.9]
BASFI, (scale 0-10)	3.4 (2.4) [0 - 8.4]

Results are presented as mean (SD); [min-max] unless otherwise stated.

[*] Median (IRQ)

[†] BASMI$_{Lin}$ values from initial assessment. BASMI$_{Lin}$ component results are item values on a 1-10 scale. The higher the BASMI$_{Lin}$ score, the more severe the individual's limitation of movement.

[‡] BASDAI not completed by n=1 participant

Abbreviations – BAS-G: Bath Ankylosing Spondylitis Global Score; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; BASMI$_{Lin}$: Bath Ankylosing Spondylitis Metrology Index (linear version); BMI: body mass index.

3.2. Protocol Fidelity

Thirty-six participants completed the study as per full protocol. One participant completed supervised testing but did not participate in the unsupervised laboratory testing or the ambulatory phase of testing due to a flare-up of leg pain. Technical issues affected three ambulatory test sessions; one sensor malfunctioned, one sensor fell off and was incorrectly re-positioned by the participant, and one recorder had insufficient battery for data collection. In all of these cases, the data was lost. During the unsupervised conditions, five participants performed an incorrect number of movement repetitions and one participant did not perform the movements bilaterally. Two participants did not consistently use the event button to record the start and end of a movement; this made identification of the tests difficult for the data analyst, as their movement was restricted and no clear movement sequence could be identified from the raw data.

3.3. Spinal Mobility Data

The 'full-arc' ROM of each measurement using the IMU sensors are summarized in Table 3. The normalized indices for each measurement, the BASMI$_{Lin}$ and the composite IMU-ASMI (Amb) scores are summarized in Table 4.

Table 3. Range of movement of participants measured by IMU sensors

Method	Movement	Day 1 – Supervised[*]	Day 1 – Unsupervised[†]	Ambulatory[‡]	Day 2 – Supervised[*]	Day 2 – Unsupervised[†]
Trunk IMU	Flexion-Extension	125.7 (27.1)	121.0 (27.2)	120.1 (27.4)	123.7 (25.6)	121.4 (26.4)

Deleted: ASQoL[§], (scale 0-18) ...

Deleted: [§] Cases with more than three missing responses cannot be allocated a total score therefore data from n=2 participants were excluded.¶

Deleted: ASQoL: Ankylosing Spondylitis Quality of Life questionnaire;

Deleted: Thirty-three participants completed self-report pain and fatigue NRS during all three test sessions (Day 1, Ambulatory, Day 2). There was a statistically significant difference in fatigue NRS scores depending on test session, $\chi^2(2) = 8.6154$, $p < 0.001$. Post-hoc analysis with Wilcoxon signed-rank test was conducted with a Bonferroni-adjusted significance level set at $p < 0.017$. Median (IQR) fatigue NRS scores for Day 1, Ambulatory and Day 2 sessions were 3.0 (1.8 to 6.0), 4.0 (2.0 to 6.3), and 3.0 (1.0 to 5.0), respectively. There was a statistically significant reduction in fatigue score on Day 2 compared to the ambulatory session ($Z = 3.0567$, $p = 0.0022$). No statistically significant differences in fatigue NRS scores were observed between Day 1 and ambulatory ($Z = -1.20$, $p = 0.23$) or Day 2 sessions (between $Z = 1.20$, $p = 0.1192$). No statistically significant effect of test session on pain NRS scores was observed, $\chi^2(2) = 0.1538$, $p = 0.86$. Median (IQR) pain NRS scores for Day 1, Ambulatory and Day 2 sessions were 2.0 (1.0 to 3.3), 2.0 (0.8 to 3.0), and 2.0 (0.0 to 3.0), respectively. ¶

Deleted: IMU-ASMI

	Lateral flexion L+R	46.8 (19.7)	45.4 (17.8)	43.9 (18.2)	47.1 (19.8)	45.5 (19.0)
	Rotation L+R	42.1 (22.2)	-	-	42.2 (22.4)	-
Lumbar region IMU	Flexion-Extension	60.9 (27.0)	58.6 (26.2)	57.8 (25.4)	58.2 (26.4)	56.6 (24.9)
	Lateral flexion L+R	35.4 (19.1)	34.3 (19.2)	33.6 (19.1)	35.4 (19.2)	34.4 (19.0)
	Rotation L+R	27.1 (16.1)	-	-	26.8 (15.9)	-

Figures presented as mean (SD). All movements in degrees (°). *n = 40; †n = 39; ‡n = 36

Trunk IMU: the orientation angle from the upper lumbar sensor to the ground; represents lumbar and pelvic movement. Lumbar region IMU: the angle between the upper lumbar sensor and the sacral sensor; represents lumbar movement. Output for rotation movements was only available under supervised conditions owing to technical limitations with the system.

Table 4 Normalized indices for each IMU movement and composite IMU-ASMI (Amb) score per IMU region

Deleted: IMU-ASMI

Method	Movement	Day 1 – Supervised*	Day 1 – Unsupervised†	Ambulatory‡	Day 2 – Supervised*	Day 2 – Unsupervised†
Trunk IMU	Flexion-Extension	2.2 (2.0)	2.6 (2.0)	2.7 (2.0)	2.3 (2.0)	2.5 (2.1)
	Lateral flexion L+R	4.0 (2.7)	4.2 (2.4)	4.4 (2.5)	4.0 (2.7)	4.2 (2.6)
	Rotation L+R	3.4 (3.2)	-	-	3.4 (3.2)	-
	Trunk-ASMI (Amb)	3.2 (2.3)	3.4 (2.1)	3.5 (2.1)	3.2 (2.3)	3.3 (2.2)
Lumbar region IMU	Flexion-Extension	2.5 (2.9)	2.6 (3.0)	2.7 (2.9)	2.7 (3.1)	2.8 (3.0)
	Lateral flexion L+R	4.1 (3.3)	4.3 (3.3)	4.4 (3.3)	4.1 (3.3)	4.3 (3.2)
	Rotation L+R	3.1 (3.3)	-	-	3.0 (3.2)	-
	Lumbar-ASMI (Amb)	3.2 (2.8)	3.5 (3.0)	3.5 (3.0)	3.2 (2.9)	3.5 (3.0)
$BASMI_{Lin}$	Total score	3.8 (1.8)	-	-	3.8 (1.8)	-

Figures presented as mean (SD). *n = 40; †n = 39; ‡n = 36

Trunk IMU: the orientation angle from the upper lumbar sensor to the ground; represents lumbar and pelvic movement. Lumbar region IMU: the angle between the upper lumbar sensor and the sacral sensor; represents lumbar movement. Output for rotation movements was only available under supervised conditions owing to technical limitations with the system.

Abbreviations – BASMILin: Bath Ankylosing Spondylitis Metrology Index (linear version)

3.4. Reliability and Agreement of IMU Movements

The test-retest reliability results for IMU movement tests performed in the laboratory are summarized in Table 5. Both the Trunk IMU and Lumbar region IMU showed good to excellent agreement for all movements. The SEM ranged from 5.12° to 9.02° for Flexion + Extension, 2.12° to 4.86° for Lateral flexion, and 5.98° to 8.19° for Rotation (see Supplemental Table 1). Test-retest reliability and agreement of IMU movement tests performed on different days are available in Supplemental Table 2.

The reliability and agreement analyses of IMU movement tests performed under laboratory and ambulatory conditions are summarized in Table 6. Both the Trunk IMU and Lumbar region IMU showed good to excellent agreement for all movements. The SEM ranged from 4.67° to 8.54° for Flexion + Extension and 2.17° to 5.39° for lateral flexion.

Table 5. Test-retest reliability and agreement of full-arc movement measurements and composite ASMI (Amb) scores under supervised and unsupervised conditions in the laboratory

	Supervised Day 1 v Supervised Day 2*					Unsupervised Day 1 v Unsupervised Day 2†					Supervised Day 1 v Unsupervised Day 1†					Supervised Day 2 v Unsupervised Day 2†				
	ICC [95% CI]	SEM	95% LOA Bias	95% LOA Lwr	95% LOA Upr	ICC [95% CI]	SEM	95% LOA Bias	95% LOA Lwr	95% LOA Upr	ICC [95% CI]	SEM	95% LOA Bias	95% LOA Lwr	95% LOA Upr	ICC [95% CI]	SEM	95% LOA Bias	95% LOA Lwr	95% LOA Upr
Trunk IMU																				
Flexion + Extension	**0.93** [0.88-0.96	6.98	2.0	-16.7	20.6	**0.93** [0.88-0.96]	7.20	-0.3	-19.5	18.8	**0.96** [0.84-0.98]	5.45	4.7	-7.6	17.1	**0.94** [0.90-0.97]	6.00	2.4	-14.1	19.0
Lateral flexion L+R	**0.96** [0.93-0.98]	3.95	-0.3	-11.4	10.8	**0.95** [0.91-0.97]	3.99	-0.0	-11.5	11.4	**0.94** [0.89-0.97]	4.86	2.5	-18.6	23.7	**0.97** [0.94-0.98]	3.45	2.8	-13.4	19.0
Trunk-ASMI (Amb)	**0.94** [0.89-0.97]	0.56	-0.0	-1.6	1.5	**0.96** [0.94-0.98]	0.42	0.0	-1.15	1.22	**0.91** [0.84-0.95]	0.68	-0.2	-2.0	1.5	**0.93** [0.86-0.96]	0.62	-0.1	-1.9	1.6
Lumbar region IMU																				
Flexion + Extension	0.89 [0.80-0.94]	9.02	2.7	-21.8	27.2	0.89 [0.80-0.94]	8.70	2.0	-21.4	25.4	**0.96** [0.93-0.98]	5.12	3.8	-28.4	35.9	**0.98** [0.97-0.99]	3.57	0.9	-8.7	10.5
Lateral flexion L+R	**0.98** [0.96-0.99]	2.84	-0.1	-8.1	7.9	**0.97** [0.95-0.99]	3.32	-0.1	-8.6	8.5	**0.98** [0.97-0.99]	2.45	0.7	-6.1	7.5	**0.99** [0.98-0.99]	2.12	0.9	-4.7	6.6
Lumbar-ASMI (Amb)	**0.95** [0.90-0.97]	0.63	-0.0	-1.9	1.8	**0.96** [0.93-0.98]	0.60	-0.1	-1.64	1.50	**0.96** [0.92-0.98]	0.57	-0.3	-1.9	1.4	**0.96** [0.92-0.98]	0.57	-0.3	-1.9	1.4

All ICC results were statistically significant, $p<0.001$. Bold denotes ICC >0.9. *n = 40; †n = 39

Abbreviations - ICC: Intraclass correlation coefficient; SEM: standard error of measurement (deg); 95% LOA: 95% limits of agreements (deg)

Deleted: -

Table 6. Test-retest reliability and agreement of full-arc movement measurements and composite -ASMI (Amb) scores under laboratory and ambulatory conditions.

Deleted: -

	Supervised Day 1 v Ambulatory					Unsupervised Day 1 v Ambulatory					Supervised Day 2 v Ambulatory					Unsupervised Day 2 v Ambulatory				
	ICC [95% CI]	SEM	95% LOA Bias	95% LOA Lwr	95% LOA Upr	ICC [95% CI]	SEM	95% LOA Bias	95% LOA Lwr	95% LOA Upr	ICC [95% CI]	SEM	95% LOA Bias	95% LOA Lwr	95% LOA Upr	ICC [95% CI]	SEM	95% LOA Bias	95% LOA Lwr	95% LOA Upr
Trunk IMU																				
Flexion + Extension	**0.94** [0.58-0.98]	6.56	6.9	-6.1	19.9	**0.97** [0.94-0.98]	4.67	2.3	-10.0	14.6	0.89 [0.78-0.94]	8.54	4.7	-18.6	28.0	**0.92** [0.85-96]	7.52	2.4	-18.5	23.3
Lateral flexion L+R	**0.94** [0.84-0.97]	4.95	7.3	-19.4	34.1	**0.96** [0.93-0.98]	3.31	4.8	-18.1	27.7	**0.93** [0.81-0.97]	5.39	7.6	-18.8	34.0	**0.97** [0.93-0.98]	3.45	4.8	-18.8	28.4
Trunk-ASMI (Amb)	**0.91** [0.80-0.96]	0.68	-0.4	-2.1	1.3	**0.97** [0.94-0.99]	0.36	-0.2	-1.1	0.6	0.89 [0.78-0.94]	0.77	-0.3	-2.4	1.7	**0.97** [0.93-0.98]	0.39	-0.2	-1.2	0.7
Lumbar region IMU																				
Flexion + Extension	**0.93** [0.86-0.96]	7.18	3.6	-14.7	21.8	**0.92** [0.85-0.96]	7.52	2.1	-18.2	22.5	**0.93** [0.87-0.96]	6.82	0.5	-18.3	19.2	**0.93** [0.88-0.97]	6.38	-0.1	-18.2	17.9
Lateral flexion L+R	**0.98** [0.94-0.99]	2.73	2.0	-4.6	8.7	**0.98** [0.97-0.99]	2.33	1.2	-4.9	7.4	**0.97** [0.93-0.98]	3.43	2.1	-6.8	10.9	**0.98** [0.96-0.99]	2.61	1.1	-6.1	8.2
Lumbar-ASMI (Amb)	**0.94** [0.88-0.97]	0.69	-0.4	-2.2	1.4	**0.97** [0.93-0.98]	0.52	-0.2	-1.7	1.3	**0.95** [0.90-0.97]	0.64	-0.3	-2.0	1.5	**0.98** [0.97-0.99]	0.43	-0.1	-1.1	1.0

All ICC results were statistically significant, $p<0.001$. n = 36

Abbreviations - ICC: Intraclass correlation coefficient; SEM: standard error of measurement (deg); 95% LOA: 95% limits of agreements (deg)

3.5. Reliability and Agreement of JMU-ASMI (Amb) indices

The reliability and agreement analyses of JMU-ASMI (Amb) scores are summarized in Table 5 and Table 6. Both the Trunk-ASMI (Amb) and Lumbar-ASMI (Amb) showed strong agreement under laboratory and ambulatory conditions. The SEM ranged from 0.36 to 0.77 for Trunk-ASMI (Amb) and 0.43 to 0.69 for Lumbar-ASMI (Amb). The JMU-ASMI (Amb) scores showed good correlation with $BASMI_{Lin}$ under all test conditions (see Table 7). Pearson correlations were ≥ 0.85.

Table 7 Pearson Correlations between $BASMI_{Lin}$ and JMU-ASMI (Amb) scores under laboratory and ambulatory conditions

Method	Test	$BASMI_{Lin}$ Day 1	$BASMI_{Lin}$ Day 2
Trunk-ASMI (Amb)	Supervised Day 1	0.85	0.87
	Supervised Day 2	0.85	0.88
	Unsupervised Day 1	0.85	0.88
	Unsupervised Day 2	0.86	0.91
	Ambulatory	0.88	0.91
Lumbar-ASMI (Amb)	Supervised Day 1	0.86	0.88
	Supervised Day 2	0.86	0.86
	Unsupervised Day 1	0.86	0.88
	Unsupervised Day 2	0.86	0.90
	Ambulatory	0.87	0.89

Abbreviations – $BASMI_{Lin}$: Bath Ankylosing Spondylitis Metrology Index (linear version)

3.6 Pain and Fatigue Monitoring

Thirty-three participants completed self-report pain and fatigue NRS during all three test sessions (Day 1, Ambulatory, Day 2). There was a statistically significant difference in fatigue NRS scores depending on test session, $\chi^2(2) = 8.6154$, $p < 0.001$. Post-hoc analysis with Wilcoxon signed-rank test was conducted with a Bonferroni-adjusted significance level set at $p < 0.017$. Median (IQR) fatigue NRS scores for Day 1, Ambulatory and Day 2 sessions were 3.0 (1.8 to 6.0), 4.0 (2.0 to 6.3), and 3.0 (1.0 to 5.0), respectively. There was a statistically significant reduction in fatigue score on Day 2 compared to the ambulatory session ($Z = 3.0567$, $p = 0.0022$). No statistically significant differences in fatigue NRS scores were observed between Day 1 and ambulatory ($Z = -1.20$, $p = 0.23$) or Day 2 sessions (between $Z = 1.20$, $p = 0.1192$). No statistically significant effect of test session on pain NRS scores was observed, $\chi^2 (2) = 0.1538$, $p = 0.86$. Median (IQR) pain NRS scores for Day 1, Ambulatory and Day 2 sessions were 2.0 (1.0 to 3.3), 2.0 (0.8 to 3.0), and 2.0 (0.0 to 3.0), respectively.

4. Discussion

This study demonstrates the reliability of an IMU sensor-based system for measuring spinal range of motion of individuals with axSpA. The IMU sensor-based system showed good to excellent test-retest reliability under supervised and unsupervised conditions in the laboratory setting, and unsupervised in the home setting.

Composite JMU-ASMI (Amb) scores were calculated for the lumbopelvic region (Trunk-ASMI) and the lumbar region (Lumbar-ASMI) based on methods used in previous studies [19, 26]. In this study, rotation movement data was only included in the supervised JMU-ASMI (Amb) scores due to technical limitations within the system. As rotation has a smaller range of movement in the lumbar spine than movement in the other two planes, this limitation was hypothesized to have been of negligible practical consequence [30-32]. Reliability of composite JMU-ASMI (Amb) scores was excellent across supervised test conditions, with ICCs for JMU-ASMI (Amb) scores under supervised

Deleted: IMU-ASMI
Deleted: IMU-ASMI
Deleted: IMU-ASMI
Deleted: IMU-ASMI
Formatted: MDPI_3.1_text
Deleted: IMU-ASMI
Deleted: IMU-ASMI
Deleted: IMU-ASMI
Deleted: IMU-ASMI

conditions similar to previously reported scores in a similar cohort [19]. Reliability of composite JMU-ASMI (Amb) scores was also found to be excellent in unsupervised ambulatory settings; an unsupervised JMU-ASMI (Amb) score could therefore function as a reliable surrogate for a supervised IMU-ASMI score.

The limits of agreement showed greater full-arc range of motion and lower JMU-ASMI (Amb) scores (better performance) when participants were supervised than when unsupervised, suggestive of a small systematic bias. Participants may have tried harder when under direct observation than when unsupervised, due to beliefs about researcher expectations and social desirability [33]. Participants may also be more likely to perform movements slightly 'off-plane' or with less accuracy when performing the movements unsupervised, resulting in reduced range of motion being recorded. Circadian rhythm of symptoms in axSpA may have influenced the performance of spinal movement tasks, however, pain and fatigue symptoms were shown to be largely stable across test sessions.

A secondary aim of the study was to evaluate the reliability of an JMU-ASMI (Amb) score and to determine correlation with conventional metrology. Both supervised and unsupervised JMU-ASMI (Amb) scores showed strong correlations with $BASMI_{Lin}$ and may be a suitable proxy for conventional metrology when direct measurement by a clinician is not possible. A limitation of the JMU-ASMI (Amb) scores described is they do not include measures of standing posture, hip or neck range of motion. As a result, they should not be considered a substitute for conventional BASMI. Including these additional components would require additional IMU sensors, and longer set-up and test protocol time and was beyond the scope of this study. Despite this limitation, the JMU-ASMI (Amb) gives a comprehensive and accurate representation of spinal movement in degrees across three planes of movement

This study supports the concept that individuals with axSpA can use an IMU sensor-based system to monitor their spinal mobility reliably and accurately, without supervision at home or in non-clinical settings. While this would not replace supervised tests in a clinical setting, it offers clinicians a reliable method of remotely monitoring spinal mobility in individuals with axSpA. This is an important step in developing a system that will allow clinicians and researchers to track small changes in spinal mobility over time, and measure the impact of exercise programs, without necessitating frequent, in-person consultations. The increase in remote telehealth consultations, accelerated by the SARS-CoV-2 pandemic, is broadly supported by patients and clinicians [12, 34-36]. However, the inability to conduct a physical examination of spinal mobility has presented a persistent obstacle to the adoption of remote consultations [35, 37]. IMU sensor-based systems could provide a solution by facilitating reliable and accurate measurement of spinal metrology.

eHealth and mobile-based applications have been recognized as potential ways of improving remote monitoring. Mobile health (mHealth) can contribute to the empowerment of patients, who could manage their health more actively and live more independently thanks to self-assessment or remote monitoring solutions, and support healthcare professionals in treating patients more efficiently [38]. Most studies to date examining eHealth and mHealth in rheumatology have focused on rheumatoid arthritis [39-40]. Recent systematic reviews have identified approximately 35 apps currently available that offer symptom tracking, educational information and links to online communities for people with rheumatic and musculoskeletal disease [41-43]. However, only one app was specific to axSpA [44]. Future development of an IMU sensor-based system linked to a mobile application could enhance the utility and specificity of such an application in relation to axSpA, where monitoring of spinal range of motion is an important indicator of disease progression [45-46].

In addition to providing data to a clinician, the output of the IMU sensor-based system could support self-management interventions [47-50]. JMU-ASMI (Amb) scores, expressed on a scale of 0 to 10, are easy to interpret without knowledge of the normal ranges of each spinal movement. This could be used by people with axSpA as a motivational point to encourage adherence to exercise or pharmaceutical treatments and facilitate self-monitoring during maintenance phases or disease flares. It is a strength of the study that a broad cross-section of individuals with axSpA participated, ranging in demographic characteristics, clinical features and treatments. The majority of participants

Deleted: IMU-ASMI
Deleted: IMU-ASMI
Deleted: IMU-ASMI
Deleted: 0
Deleted: IMU-ASMI
Deleted: IMU-ASMI
Deleted: IMU-ASMI
Deleted: IMU-ASMI
Deleted: 1
Deleted: 3
Deleted: 2
Deleted: 4
Deleted: 5
Deleted: 6
Deleted: 37
Deleted: 38
Deleted: 0
Deleted: 1
Deleted: 2
Deleted: 3
Deleted: 4
Deleted: 47
Deleted: IMU-ASMI

completed the simple test movements unsupervised by following simple standardized instructions, either by video or written instructions; just ten percent of participants failed to complete the study protocol in full, demonstrating the feasibility.

One participant did not complete the ambulatory phase testing as a sensor detached from the skin, and they were unable to reposition it themselves. This illustrates a limitation of the current sensor setup, with participants unable to self-attach sensors with the degree of precision required. In this study sensors were attached by a trained member of the research team, however, if this is to be adopted as a self- or remote-monitoring tool, an alternate way of attaching sensors is required. This study used a two-sensor setup, however, a single sensor set-up warrants further investigation; results from the Trunk IMU (positioned at L1 vertebra) showed equivalent reliability to the results from the two sensor Lumbar region IMU setup. Finally, unlike conventional metrology measures, the sensor set-up used in this study does not include measures of cervical mobility. However, the correlation between mean –ASMI (Amb) versus mean $BASMI_{Lin}$ was 0.82, which suggests that the IMU–ASMI (Amb) is a clinically relevant measure, despite not including the cervical region. Furthermore, it has been previously shown that removing cervical mobility tests does not affect the reliability of the IMU-ASMI [19].

Currently, standard clinical tests of spinal movement in individuals with axSpA focus on movements in a single plane. IMU sensor systems offer the potential for measuring multi-planar spinal movements that would be closer to 'real-world' movements; as well as providing the additional benefit of monitoring and attaining data over longer periods than that of a clinic-based assessment. Future research should seek to establish the validity and reliability of an IMU sensor-based system to measure spinal mobility during functional movement tests. More performance-based tests may be of more interest to both the clinician and the patient and may be a more objective measure of function instead of pure mobility.

5. Conclusions

This study has demonstrated that an IMU sensor-based system is a reliable way of measuring spinal mobility in axSpA under supervised and unsupervised conditions. While not a replacement for established clinical measures, composite IMU-ASMI (Amb) scores may be reliably used as a proxy measure of spinal mobility.

Deleted: IMU-ASMI

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Supplemental Table 1: Test-retest reliability and agreement of full-arc rotation measurements between IMU sensors under supervised conditions in the laboratory; Supplemental Table 2: Test-retest reliability and agreement of full-arc movement measurements and composite -ASMI scores under supervised and unsupervised conditions on different days in the laboratory, Supplementary Document: The $BASMI_{Lin}$ and IMU-ASMI composite indices explained in detail; Supplementary Video: Example of instructional video.

Author Contributions: Conceptualization, MO'G, TO'D, FDO'S, PG and FW; methodology, MO'G, TO'D, FDO'S, PG and FW; software, JC (Joan Condell); validation, JC (James Connolly) and PG; formal analysis, MO'G, TO'D, JC (Joan Condell) and PG; investigation, MO'G and KME; resources, JC (Joan Condell), FDO'S, PG and FW; data curation, MO'G, JC (James Connolly) and FDO'S; writing—original draft preparation, MO'G and TO'D; writing—review and editing, MO'G, TO'D, JC (James Connolly), JC (Joan Condell), KME, FDO'S, PG and FW; visualization, MO'G and TO'D; supervision, JC (Joan Condell), FDO'S, PG and FW; project administration, JC (Joan Condell), KME, FDO'S, PG and FW; funding acquisition, JC (Joan Condell) and PG. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Foundation for Research in Rheumatology (FOREUM). The funding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of the data; in the writing of the manuscript; or in the decision to publish the results. Sensors were purchased from DorsaVi™ under a research agreement. DorsaVi™ had no input in the analysis or interpretation of the data. This study was funded by FOREUM under a three year project grant entitled 'Measuring SpA' [available at https://foreum.org/pSpA_10_imuaxspa.cfm]. Foreum is a not-for-profit organization to promote research in rheumatology (charity registration number CHE 234.768.368)

Acknowledgments: We thank the Ankylosing Spondylitis Association of Ireland, Arthritis Ireland and Tochukwu Anachebe for their support in participant recruitment, James Hale for his assistance in preparing the instructional videos, and the technical support team at DorsaVi™ for their assistance. All authors belong to iMaxSpA Study Group.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Supplemental Table 1. Test-retest reliability and agreement of full-arc rotation measurements between IMU sensors under supervised conditions in the laboratory

	Supervised Day 1 v Supervised Day 2				
	ICC [95% CI]	SEM	95% LOA		
			Bias	Lwr	Upr
Rotation L+R					
Trunk IMU	0.86 [0.76-0.93]	8.19	-0.2	-23.2	22.9
Lumbar region IMU	0.86 [0.75-0.92]	5.98	0.3	-16.3	16.9

n = 40. All ICC results were statistically significant, p<0.001.

Abbreviations - ICC: Intraclass correlation coefficient; SEM: standard error of measurement (deg); 95% LOA: 95% limits of agreements (deg)

Supplemental Table 2. Test-retest reliability and agreement of full-arc movement measurements and composite -ASMI scores under supervised and unsupervised conditions on different days in the laboratory

	Supervised Day 1 v Unsupervised Day 2					Unsupervised Day 1 v Supervised Day 2				
	ICC [95% CI]	SEM	95% LOA			ICC [95% CI]	SEM	95% LOA		
			Bias	Lwr	Upr			Bias	Lwr	Upr
Trunk IMU										
Flexion + Extension	**0.93** [0.84-0.96]	7.31	4.4	-14.1	22.9	**0.91** [0.84-0.95]	8.17	-2.8	-24.1	18.6
Lateral flexion L+R	**0.95** [0.91-0.98]	4.28	2.5	-17.9	23.0	**0.92** [0.85-0.96]	5.04	-2.8	-22.7	17.0
Trunk-ASMI	**0.92** [0.86-0.96]	0.64	-0.2	-1.9	1.6	0.88 [0.78-0.93]	0.72	0.2	-2.0	2.3
Lumbar region IMU										
Flexion + Extension	**0.90** [0.82-0.95]	8.37	3.4	-18.1	24.9	0.87 [0.76-0.93]	9.46	1.1	-25.7	27.8
Lateral flexion L+R	**0.98** [0.96-0.99]	2.92	0.6	-7.5	8.7	**0.96** [0.93-0.98]	3.84	-1.0	-11.3	9.3
Lumbar-ASMI	**0.94** [0.89-0.97]	0.69	-0.3	-2.2	1.5	**0.92** [0.85-0.96]	0.85	0.2	-2.1	2.5

n = 39. All ICC results were statistically significant, p<0.001. Bold denotes ICC >0.9.

Abbreviations - ICC: Intraclass correlation coefficient; SEM: standard error of measurement (deg); 95% LOA: 95% limits of agreements (deg)

References

1. Sieper, J.; Poddubnyy, D. Axial spondyloarthritis. *Lancet* **2017**, *390*, 73-84.
2. Baraliakos, X.; Heldmann, F.; Callhoff, J.; Listing, J.; Appelboom, T.; Brandt, J.; Van den Bosch, F.; Breban, M.; Burmester, G.; Dougados, M.; Emery, P.; Gaston, H.; Grunke, M.; Van Der Horst-Bruinsma, I. E.; Landewé, R.; Leirisalo-Repo, M.; Sieper, J.; De Vlam, K.; Pappas, D.; Kiltz, U.; Van Der Heijde, D.; Braun, J. Which spinal lesions are associated with new bone formation in patients with ankylosing spondylitis treated with anti-TNF agents? A long-term observation study using MRI and conventional radiography. *Ann Rheum Dis* **2014,** 73, 1819-1825.
3. Wanders, A.; Landewé, R.; Dougados, M.; Mielants, H.; Van der Linden, S.; Van der Heijde, D. Association between radiographic damage of the spine and spinal mobility for individuals with ankylosing spondylitis: can assessment of spinal mobility be a proxy for radiographic evaluation?. *Ann Rheum Dis* **2005**, *64*, 988-994.
4. Van der Heijde, D.; Ramiro, S.; Landewé, R.; Baraliakos, X.; Van den Bosch, F.; Sepriano, A.; Regel, A.; Ciurea, A.; Dagfinrud, H.; Dougados, M.; Van Gaalen, F.; Géher, P.; Van der Horst-Bruinsma, I.; Inman, R. D.; Jongkees, M.; Kiltz, U.; Kvien, T. K.; Machado, P.; Marzo-Ortega, H.; Molto, A.; Navarro-Compan, V.; Ozgocmen, S.; Pimentel-Santos, F. M.; Reveille, J.; Rudwaleit, M.; Sieper, J.; Sampaio-Barros, P.; Wiek, D.; Braun, J. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. *Ann Rheum Dis* **2017**, *76*, 978-991.
5. Van der Heijde, D.; Calin, A.; Dougados, M.; Khan, M.; 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*, 951-954.
6. Sieper, J.; Rudwaleit, M.; Baraliakos, X.; Brandt, J.; Braun, J.; Burgos-Vargas, R.; Dougados, M.; Hermann, K.G.; Landewé, R.; van der Heijde, D. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann Rheum Dis* **2009**, *68*, ii1-ii44.
7. Landewé, R.; Van Tubergen, A. Clinical Tools to Assess and Monitor Spondyloarthritis. *Curr Rheumatol Rep* **2015**, *17*, 522.
8. Jenkinson, T.; Mallorie, P.; Whitelock, H.; Kennedy, L.; Garrett, S.; Calin, A. Defining spinal mobility in ankylosing spondylitis (AS). The Bath AS Metrology Index. *J Rheumatol* **1994**, *21*, 1694-1698.
9. Jauregui, E.; Conner-Spady, B.; Russel, A.; Maksymowych, W. Clinimetric evaluation of the bath ankylosing spondylitis metrology index in a controlled trial of pamidronate therapy. *J Rheumatol* **2004**, *31*, 2422-2428.
10. Auleley, G.; Benbouazza, K.; Spoorenberg, A.; Collantes, E.; Hajja-Hassouni, N.; Van der Heijde, D.; Dougados, M. Evaluation of the smallest detectable difference in outcome or process variables in ankylosing spondylitis. *Arthrit Care Res* **2002**, *47*, 582-587.
11. Calvo-Gutiérrez, J.; Garrido-Castro, J.; González-Navas, C.; Castro-Villegas, C.; Ortega-Castro, R.; López-Medina, C.; Font-Ugalde, P.; Escudero-Contreras, A.; Collantes-Estévez, E. Inter-rater reliability of clinical mobility measures in ankylosing spondylitis. *BMC Musculoskelet Disord* **2016**, *17*, 382.
12. American College of Rheumatology. Position statement of Telemedicine. June 29, 2020, available online at https://www.rheumatology.org/Portals/0/Files/Telemedicine-Position-Statement.pdf [Accessed 7th November 2020]
13. Papi, E.; Koh, W.; McGregor, A. Wearable technology for spine movement assessment: a systematic review. *J Biomech* **2017**, *64*, 186-197.
14. Cuesta-Vargas, A.; Galán-Mercant, A.; Williams, J. The use of inertial sensors system for human motion analysis. *Phys Ther Rev* **2010**, *15*, 462-473.
15. Simpson, L.; Maharaj, M.; Mobbs, R. The role of wearables in spinal posture analysis: a systematic review. *BMC Musculoskelet Disord* **2019**, *20*, 55.
16. Ronchi, A.; Lech, M.; Taylor, N.; Cosic, I. A reliability study of the new back strain monitor based on clinical trials. *Annu Int Conf IEEE Eng Med Biol Soc* **2008**, *2008*, 693-696.
17. Charry, E.; Umer, M.; Taylor, S. Design and validation of an ambulatory inertial system for 3-D measurements of low back movements. In: *2011 Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing* **2011**, 58-63.
18. Aranda-Valera, I.; Cuesta-Vargas, A.; Garrido-Castro, J.; Gardiner, P.; López-Medina, C.; Machado, P.; Condell, J.; Connolly, J.; Williams, J.; Muñoz-Esquivel, K.; O'Dwyer, T.; Castro-Villegas, M.; González-

Navas, C.; Collantes-Estévez, E. Measuring spinal mobility using an inertial measurement unity system: a validation study in axial spondyloarthritis. *Diagnostics* **2020**, *10*, 426.
19. Gardiner, P.; Small, D.; Muñoz-Esquivel, K; Condell, J.; Cuesta-Vargas, A.; Williams, J.; Machado, P.; Garrido-Castro, J. Validity and reliability of a sensor-based electronic spinal mobility index for axial spondyloarthritis. *Rheumatology* **2020**, 59, 3415-3423.1-9.
20. Garrett, S.; Jenkinson, T.; Kennedy, G.; Whitelock, H.; Gaisford, P.; Calin, A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis disease activity index. *J Rheumatol* **1994**, *1*, 2286-2291.
21. Calin, A.; Garett, S.; Whitelock, H.; Kennedy, L.; O'Hea, J.; Mallorie, P.; Jenkinson, T. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. *J Rheumatol* **1994**, *21*, 2281-2285.
22. Jones, S.; Steiner, A.; Garrett, S.; Calin, A. The bath ankylosing spondylitis patient global score (BAS-G). *Rheumatology (Oxford)* **1996**, *35*, 66-71.
23. Doward, L.; Spoorenberg, A.; Cook, S.; Whalley, D.; Helliwell, P.; Kay, L.; McKenna, S.; Tennant, A.; van der Heijde, D.; Chamberlain, M. Development of the ASQoL: a quality of life instrument specific to ankylosing spondylitis. *Ann Rheum Dis* **2003**, *62*, 20-26.
24. Mjøsund, H. L.; Boyle, E.; Kjaer, P.; Mieritz, R. M.; Skallgård, T.; Kent, P. Clinically acceptable agreement between the ViMove wireless motion sensor system and the Vicon motion capture system when measuring lumbar region inclination motion in the sagittal and coronal planes. *BMC Musculoskelet Disord* **2017**, *18*, 124.
25. Jenkinson, T.; Mallorie, P.; Whitelock, H.; Kennedy, L.; Garrett, S.; Calin, A. Defining spinal mobility in ankylosing spondylitis (AS). The Bath AS Metrology Index. *J Rheumatol* **1994**, *21*, 1694-1698.
26. van der Heijde, D.; Landewé, R.; Feldkeller, E. Proposal of a linear definition of the Bath Ankylosing Spondylitis Metrology Index (BASMI) and comparison with the 2-step and 10-step definitions. *Ann Rheum Dis* **2008**, *67*, 489-493.
27. Spies, C.M.; Cutolo, M.; Straub, R.H.; Burmester, G-R.; Buttgereit, F. More night than day – circadian rhythms in polymyalgia rheumatica and ankylosing spondylitis. *J Rheumatol* **2010**, *37*, 894-899.
28. Koo, T.; Li, M. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* **2016**, *15*, 155-163.
29. Schober, P.; Boer, C.; Schwarte, L.A; Correlation coefficients: appropriate use and interpretation. *Anesth Analg* **2018**, *126*, 1763-1768.
30. Chow, D.H; Luk, K.D.; Leong, J.C.; Woo, C.W. Torsional stability of the lumbosacral junction. Significance of the iliolumbar ligament. *Spine* **1989**, *14*, 611-615.
31. White, A.A.; Panjabi, M.M. *Clinical biomechanics of the spine*, 2nd ed.; J.B. Lippincott, Philadelphia, United States of America, 1990.
32. Bogduk, N. Functional anatomyof the spina. *Handb Clin Neurol* **2016**, *136*, 675-688.
33. McCambridge, J.; Witton, J.; Elbourne, D.R. Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *J Clin Epidemiol* **2014**, *67*, 267–277.
34. Cavagna, L.; Zanframundo, G.; Codullo, V.; Grazia Pisu, M.; Caporali, R.; Montecucco, C. Telemedicine in rheumatology: a reliable approach beyond the pandemic. *Rheumatology* **2020**, *keaa554*.
35. Opinc, A.; Lukasik, Z.; Makowska, J. The attitude of Polish rheumatology patients towards telemedicine in the age of the COVID-19 pandemic. *Reumatologia* **2020**, *58*, 134–141.
36. López-Medina, C.; Escudero, A.; Collantes-Estevez, E. COVID-19 pandemic: an opportunity to assess the utility of telemedicine in patients with rheumatic diseases. *Ann Rheum Dis* **2020**.
37. Devadula, S.; Langbecker, D.; Vecchio, P.; Meiklejohn, J.; Benham, H. Tele-rheumatology to regional hospital outpatient clinics. Patient perspectives on a new model of care. *Telemed J E-Health* **2020**, 26.
38. European Commission. Green paper on mobile health ("mHealth"). **2014,** available online https://ec.europa.eu/digital-single-market/en/news/green-paper-mobile-health-mhealth. [Accessed 7th November 2020]
39. Magnol, M.; Berard, E.; Rempenault, C.; Castagne, B.; Pugibet, M.; Lukas, C.; Tournadre, A.; Vergne-Salle, P.; Barnetche, T.; Truchetet, M.E.; Ruyssen-Witrand, A. Use of eHealth by patients with rheumatoid arthritis: an observational, cross sectional, multicenter study. *Ann Rheum Dis* **2020**, *79*, 532-533.
40. Knitza, J.; Simon, D.; Lambrecht, A.; Raab, C.; Tascilar, K.; Hagen, M.; Kleyer, A.; Bayat, S.; Derungs, A.; Amft, O.; Schett, G.; Hueber, A.J. Mobile health usage, preferences, barriers, and eHealth literacy in Rheumatology: patient survey study. *JMIT Mhealth Uhealth* **2020**, *8*, 19661.

41. Luo, D.; Wang, P.; Lu, F.; Elias, J.; Sparks, J.A.; Lee, Y.C. Mobile apps for individuals with rheumatoid arthritis: a systematic review. *J Clin Rheumatol* **2019**, *25*, 133–141.
42. Grainger, R.; Townsley, H.; White, B.; Langlotz, T.; Taylor, W. J. Apps for people with rheumatoid arthritis to monitor their disease activity: a review of apps for best practice and quality. *JMIR Mhealth Uhealth* **2017**, *5*, e7.
43. Najm, A.; Gossec, L.; Weill, C.; Benoist, D.; Berenbaum, F.; Nikiphorou, E. Mobile Health Apps for Self-Management of Rheumatic and Musculoskeletal Diseases: Systematic Literature Review. *JMIR mHealth and uHealth* **2019**, 7, e14730.
44. Salaffi, F.; Gasparini, S.; Ciapetti, A.; Gutierrez, M.; Grassi, W. Usability of an innovative and interactive electronic system for collection of patient-reported data in axial spondyloarthritis: comparison with the traditional paper-administered format. *Rheumatology* **2013**, *52*, 2062–2070.
45. Azevedo, A.R.P.; de Sousa, H.M.L.; Monteiro, J.A.F; Lima, A.R.N.P. Future perspectives of smartphone applications for rheumatic diseases self-management. *Rheumatol Int* **2015**, *35*, 419-431.
46. Van Der Veer, S.; Austin, L.; Sanders, C.; Dixon, W. Using smartphones to improve remote monitoring of rheumatoid arthritis: completeness of patients' symptom reports. *Ann Rheum Dis* **2017**, *76*, 547.
47. Barlow, J.; Wright, C.; Sheasby, J.; Turner, A.; Hainsworth, J. Self-management approaches for people with chronic conditions: a review. *Patient Educ Couns* **2002**, *48*, 177-187.
48. Risling, T.; Martinez, J.; Young, J.; Thorp-Froslie, N. Evaluating patient empowerment in association with eHealth technology: Scoping review. *J Med Internet Res* **2017**, *19*, e329.
49. Bradway, M.; Årsand, E.; Grøttland, A.; Mobile health: empowering patients and driving change. *Trends Endocrinol Metab* **2015**, *26*, 114–117.
50. Holmes, M.M.; Stanescu, S.; Bishop, F.L. The use of measurement systems to support patient self-management of long-term conditions: an overview of opportunities and challenges. *Patient Relat Outcome Meas* **2019,** *10*, 385–394.