Preliminary study: reliability of the spinal wheel. A novel device to measure spinal postures applied to sitting and standing

Liba Sheeran · Valerie Sparkes · Monica Busse · Robert van Deursen

Received: 19 May 2009 / Revised: 5 November 2009 / Accepted: 28 November 2009 / Published online: 15 December 2009
© Springer-Verlag 2009

Abstract Postural re-education is an integral part of physiotherapy management in patients with back pain. Although the link between posture and back pain is largely unknown, postural re-education is performed on the premise of optimizing spinal alignment to minimize stresses on the passive structures of the spine, to facilitate optimal muscular support and thus to prevent possible damage and further pain. A reliable and clinically meaningful measurement of spinal postures to monitor such interventions remains challenging. This study evaluated within-day (intra-tester, inter-tester) and between-day (test–retest) reliability of a novel spinal wheel device measuring thoracic and lumbar postures during sitting and standing. 17 healthy volunteers (age 39.5 ± 5.4 , BMI 25 ± 9.2 ; 9 males) were measured three times, by three testers, on two separate occasions (1 week apart). The angular change between C7 and T12 and between T12 and S1 provided thoracic and lumbar curvatures, respectively. Intra-class correlation coefficient (ICC) with 95% confidence intervals and typical error were calculated. Excellent reliability was demonstrated with intra-tester ICCs between 0.947 and 0.980 and typical error between 1.7° and 3.7° and inter-tester ICCs between 0.949 and 0.986 and typical error between 2.0° and 4.7° . Test–retest reliability was high with ICCs 0.719–0.908 and typical error between 4.0° and 7.4° . In conclusion, the spinal wheel demonstrated excellent within-day and high between-day reliability. The device may be used in conjunction with 2D camcorder to provide

clinically useful visual evaluation of postures for assessment, intervention monitoring, and feedback during postural re-education.

Keywords Back pain · Thoracic kyphosis · Lumbar lordosis · Non-invasive measurement · Reliability

Introduction

Postural re-education is an integral part of physiotherapy management in the treatment of patients with back pain. Although the link between posture and back pain is largely unknown, it has been speculated that end-of-range sitting and standing postures, that were shown to be associated with altered muscle activation patterns [1–5], may result in excessive loading of the passive spinal structures, eventual trauma and injury and could potentially pose a risk to the development of back pain [4, 6]. Postural re-education is performed on the premise of optimizing spinal alignment to minimize stresses on the spinal structures and to improve trunk muscle efficiency in protecting the spine [7–10]. Objective evaluation of postures has traditionally been an important element of clinical assessment and evaluation of the effectiveness of postural re-training. In addition, objective evaluation of postural patterns is used clinically to sub-classify patients into specific motor control impairment sub-groups to guide intervention [10–12].

Several methods that quantify spinal postures have been utilized, ranging from simple observation and tape measure techniques to computerized motion analysis systems. Analysis of biplanar radiographs is generally considered the gold standard for measuring spinal curvatures and range of movement [13–15], but relatively high cost and patients' exposure to repeated radiation is a disadvantage

L. Sheeran (✉) · V. Sparkes · M. Busse · R. van Deursen
Department of Physiotherapy, School of Healthcare Studies,
Cardiff University, Ty Dewi Sant, Heath Park,
Cardiff CF14 4XN, UK
e-mail: libasheeran@googlemail.com

[13, 16, 17] which precludes its use as a routine measurement in clinical practice.

More clinically friendly, inexpensive techniques include goniometers [16, 18–21], inclinometers [17, 22–27] and flexicurves [28–31]. Goniometers and inclinometers are secured or held directly over previously palpated vertebrae, producing a measure that is limited to a single thoracic, lumbar, or sacral region of the spine [16, 17, 19, 22, 25]. The lack of examiner's experience in using the device causing a slippage or misplacement of the device and the poor ability to identify bony landmarks were found to be the main source of error [14, 16, 19, 25]. Flexicurve measurement, where a trace of the whole spine in the sagittal plane is measured [28, 29, 32], produced contrasting results in terms of reliability [30, 31].

Other skin-surface instruments (e.g. Pen-pointer device and Spinal Mouse) that quantify a sagittal curvature by guiding the device along the spinal column to measure spinal curvatures have also been used. The pen-pointer device (Zebris, Germany) is placed over the spinal processes to record three-dimensional (3-D) position using ultrasonic transmitters fitted in the probe. To the authors' knowledge, the reliability of this device has not been demonstrated. The Spinal Mouse (Idiag, Voletswil, Switzerland) is a device that contains accelerometers, which measure distance and inclination as it is guided along the length of the spine [33]. A study which evaluated the reliability of the Spinal Mouse demonstrated moderate to good between-day and inter-examiner reliability when measuring lumbar and thoracic curvatures [33].

Computerized movement analysis instruments (e.g. 3Space Fastrack, Isotrack, piezoresistive accelerometer) measure curvature and spinal movement using sensors attached to the skin at positions corresponding to the vertebrae, determined by palpation [34–37]. Various angles such as sacral tilt, lower lumbar and upper lumbar angles are calculated to determine the spinal curvature [2, 38]. The advantage of such devices is that the spinal curvature can be monitored and recorded continuously during movement [33]. Such systems have been found to be highly accurate and reliable [34, 35, 39–41], with reported measurement errors of 0.42 degrees for piezoresistive accelerometer [35] and errors ranging between 0.15° and 0.96° for Fastrack systems [41, 42]. Palpation of bony landmarks is however suggested to attribute to a significant source of error during repeated spinal measurements [43, 44]. The accuracy of these systems is also affected by the skin movement under the sensor that is unrelated to the vertebral movement, and the sensor's ability to maintain firm apposition during the length of the trial [33, 40]. Another disadvantage is that only a fraction of the spinal position, limited to few vertebrae is measured [33, 42, 45, 46]. Although these systems could have their place when

investigating specific spinal pathologies, the clinical usefulness of such instruments may be limited when evaluating postures of patients with NSLBP who may demonstrate postural changes affecting the whole spine. In addition, a major disadvantage of these systems is the limited clinical applicability related to the complexity of the set-up and lengthy data acquisition that lacks visual feedback, often required by clinicians to monitor and modify an intervention.

A novel device, the spinal wheel is a free-wheeling device with a reflective marker in its centre that is guided along the spine. The continuous motion of the spinal wheel can be detected either by a single video-camera or by a 3-D kinematic system which provides an exact trace signal of the curvature. The advantage of the spinal wheel is that it could provide a quick visual feedback to evaluate and modify intervention in clinical setting.

The aim of this study is to examine the within-day intra-tester, inter-tester and between-day test–retest reliability of the spinal wheel in measuring sagittal thoracic and lumbar curvatures. Although, some data exist on measuring spinal curvature in standing, there is limited literature on measuring sitting postures. Therefore, this study evaluated spinal curvatures of healthy participants in their usual standing and sitting postures.

Methods

Participants

Seventeen healthy subjects (8 females) were recruited to this study. A history of back pain requiring physiotherapy or other intervention within past 1 year and spinal surgery were the exclusion criteria. The study was approved by the institution's Ethical Committee and all subjects were required to provide informed consent.

Experimental protocol

A repeated measures design was followed where all the subjects' spinal curvatures were measured three times by three testers on two occasions (1 week apart). The spinal wheel is a custom-built, free-wheeling plastic wheel device 10 cm in diameter with a reflective marker in its exact centre (Fig. 1), which was guided by the researcher along the midline from the starting point of S1 vertebra up to the end point of the occiput (see Fig. 2). The continuous motion of the wheel was detected by a 3-D kinematic system (Vicon 512, Oxford Metrics Limited, Oxford, UK) providing an exact trace signal of the spinal posture. Prior to data collection, the kinematic system was calibrated according to the manufacturer's guidelines. The C7, T12

Fig. 1 Spinal wheel

Fig. 2 Method of using spinal wheel

and S1 bony landmarks were identified and reflective markers were securely attached in a sitting position to be measured to minimize the skin movement [47]. Bony landmarks were identified and marked by the main researcher (experienced physiotherapist) and double checked by an independent senior physiotherapist. Location of the C7 spinous process was found by the subject flexing the head and the most prominent spinous process at the base of the neck was palpated. The marker was affixed in the identified area while subject's head returned to the erect state. S1 process was located by counting one down from L5 spinous process, which was identified by finding L4-5 interspace with hands placed horizontally on the highest point of the iliac crests as described by Vergara et al. [47]. T12 spinous process was identified by counting from S1 [48].

Subjects were coded and randomized to establish the order of the testers measuring the spine. Prior to data collection, each tester (all of whom were senior

physiotherapists of similar experience level) performed 10 practice trials to familiarize them with the use of the device. Subjects were asked to sit on a stool (fixed height) and fold their arms across their chest and maintain the same posture while the measurements were being taken. Each tester then guided the spinal wheel along the spinal groove. The same procedure was repeated in relaxed standing position. All three testers collected three measurements in both positions (sitting and standing) for every subject. All subjects were invited 1 week later and the experimental protocol was repeated (including identification and marking of the bony landmarks) to obtain the between-day measurements. Due to logistical reasons, the second visit data collection was performed by a single tester only.

Data processing

Each trial was visually checked and labeled with a marker set defined within the Vicon Workstation system. Angular values between 18 reference points (corresponding to 18 inter-vertebral spaces between vertebrae C7 and S1) relative to vertical axis were calculated in a Matlab 7.0. Positive angles were defined in flexion direction. For this reliability study, a kyphosis and lordosis angle was calculated. The sum of the angular changes between each segment in the C7-T12 region (C7-T1, T1-T2, T2-T3, T3-T4, T4-T5, T5-T6, T6-T7, T7-T8, T8-T9, T9-T10, T10-T11, T11-T12) provided the angular value of thoracic kyphosis. The sum of the angular changes between each segment in the L1-S1 region (L1-L2, L2-L3, L3-L4, L4-L5, L5-S1) provided the angular value of lumbar lordosis. The exact spinal curve was constructed in a diagram form, allowing for visual comparison of the repeated measures (Fig. 3).

Fig. 3 Graphical representation of spinal curvature measurement (three trials) in a single subject

Statistical analysis

Intra-class correlation coefficients (ICC) with 95% confidence intervals (CI) were calculated using the Statistical Package for the Social Sciences (SPSS) Version 14. Intra-tester and test–retest reliability were assessed with a two-way mixed model with consistency while inter-tester reliability was evaluated using a two-way mixed effects model with agreement [49]. An ICC level > 0.81 was considered to demonstrate excellent reliability, ICC > 0.61 – 0.80 substantial reliability, ICC of 0.41 – 0.60 moderate reliability and ICC < 0.40 demonstrated slight reliability [50]. In order to determine within-subject variation, intra- and inter-tester typical error between the three sets of measurements was calculated from the square root of the “mean squared error” term obtained from one-way analysis of variance (ANOVA) output [51–53]. Test–retest typical error and its 95% confidence intervals were obtained from the numerical difference between the two mean measures on day one and day two [52].

Power calculation

The sample size ($n = 17$) was sufficient to characterize reliability where an ICC score of > 0.6 was considered acceptable and an ICC of < 0.2 was deemed unacceptable, with repeated measurements by three testers (power of 80% and alpha of 0.05) [50, 54].

Results

Participants

The recruited sample consisted of 8 females [40 ± 10.8 years; height 163 ± 10 cm; weight 65 ± 6.3 kg; body mass index (BMI) 24.5 ± 6.8 kg/m²] and 9 males (37 ± 10.4 years; height 177 ± 3 cm; weight 80 ± 12.3 kg; BMI

25.5 ± 6.1 kg/m²). There was no significant difference between the BMI between females and males ($p = 0.870$) with both sexes demonstrating BMI within normal limits.

Within-day measures

Intra-tester mean scores and the reliability measures are shown in Table 1. Excellent intra-tester reliability was achieved with ICCs ranging from mean score of 0.947 (95% CI 0.819–0.990) for lumbar lordosis (LL) in sitting to 0.980 (0.928–0.996) for thoracic kyphosis (TK) in standing (Table 1). Typical error between the intra-tester measures ranged from 1.7° for standing TK to 3.7° for sitting LL. Inter-tester mean scores and reliability measures (Table 2) also demonstrated excellent reliability with ICC scores ranging between 0.949 (95% CI 0.845–0.987) for sitting LL and 0.986 (95% CI 0.958–0.997) for standing TK. The inter-tester typical errors were slightly higher ranging between 2.0° for standing TK and 4.7° for sitting LL (Table 2). Thoracic measures consistently demonstrated higher intra- and inter-tester correlations with mean ICC score of 0.982 with intra- and inter-tester errors ranging between 1.7° and 2.3°, compared to the lumbar measures with mean ICC score of 0.953 and errors ranging between 2.1° and 4.7° (Tables 1, 2).

Between-day measures

The test–retest mean scores and reliability measures are shown in Table 3. High reliability was demonstrated with ICC levels ranging from 0.719 (95%CI 0.685–0.924) for standing LL to 0.908 (95% CI 0.738–0.968) for sitting TK (Table 3). Thoracic measures again demonstrated higher reliability with mean ICC scores between 0.833 and 0.908 and errors between 4.0° and 5.5°, while lumbar curvature measures demonstrated mean ICC scores of 0.883 and 0.719 with errors of 7.4° and 7.0° for sitting and standing lumbar lordosis, respectively (Table 3).

Table 1 Within-day, intra-tester mean scores (one tester; three scores) and reliability measures of thoracic and lumbar curvatures in sitting and standing

	Intra-tester reliability			
	Sitting		Standing	
	Thoracic	Lumbar	Thoracic	Lumbar
Score 1; Mean (SD)	31.5 (11.1)	−13.8 (12.5)	31.2 (10.3)	−34.0 (11.7)
Score 2; Mean (SD)	34.2 (11.2)	−10.8 (18.8)	34.2 (11.8)	−35.9 (12.4)
Score 2; Mean (SD)	36.4 (9.93)	−9.5 (19.5)	34.3 (11.8)	−37.2 (13.6)
Typical error (°)	1.8	3.7	1.7	2.1
ICC (95% CI)	0.978 (0.921–0.996)	0.947 (0.819–0.990)	0.980 (0.928–0.996)	0.962 (0.868–0.993)

SD standard deviation, ICC intraclass correlation coefficient, CI confidence interval

Table 2 Within-day, inter-tester mean scores (three testers; three scores) and reliability measures of thoracic and lumbar curvatures in sitting and standing

	Inter-tester reliability			
	Sitting		Standing	
	Thoracic	Lumbar	Thoracic	Lumbar
Tester 1; Mean (SD)	32.8 (11.3)	−15.4 (13.5)	32.4 (11.3)	−36.3 (12.8)
Tester 2; Mean (SD)	34.0 (11.7)	−16.6 (13.4)	32.7 (11.6)	−35.4 (12.0)
Tester 3; Mean (SD)	36.1 (11.8)	−14.3 (13.4)	33.0 (11.2)	−33.0 (11.5)
Typical error (°)	2.3	4.7	2.0	3.1
ICC (95% CI)	0.985 (0.953–0.996)	0.949 (0.845–0.987)	0.986 (0.958–0.997)	0.956 (0.828–0.989)

SD standard deviation, ICC intraclass correlation coefficient, CI confidence interval

Table 3 Between-day, test–retest mean scores (one tester; three scores) and reliability measures of thoracic and lumbar curvatures in sitting and standing

	Test–retest reliability			
	Sitting		Standing	
	Thoracic	Lumbar	Thoracic	Lumbar
Day 1; Mean (SD)	32.8 (11.3)	−15.4 (13.6)	32.4 (11.4)	−36.3 (12.9)
Day 2; Mean (SD)	34.3 (9.1)	−17.9 (15.7)	33.9 (9.1)	−34.9 (8.0)
Typical error° (95% CI)	4.0 (2.2–5.9)	7.4 (4.0–10.9)	5.5 (3.2–7.8)	7.0 (4.5–10.0)
ICC (95% CI)	0.908 (0.738–0.968)	0.883 (0.509–0.959)	0.833 (0.666–0.944)	0.719 (0.685–0.924)

SD standard deviation, ICC intraclass correlation coefficient, CI confidence interval

Discussion

The main objective of this study was to evaluate reliability of a novel, non-invasive spinal wheel device in measuring thoracic and lumbar postures in a sample of healthy volunteers. The results demonstrate that the reliability of measuring thoracic kyphosis and lumbar lordosis was high to excellent within and between testers and within and between days. The within-day typical error ranged between 1.7° and 2.3° for thoracic kyphosis and between 2.0° and 4.7° for lumbar lordosis. The between-day errors ranged between 4.0° and 5.5° for thoracic kyphosis and between 7.0° and 7.4° for lumbar lordosis. Normative values of thoracic kyphosis and lumbar lordosis in sitting and standing in a sample of healthy subjects were also provided.

In order to establish the level of agreement of the mean values produced by the spinal wheel with other spinal measures (convergent validity), the normative data of the thoracic kyphosis and lumbar lordosis from this study were compared to findings from other studies. The large variation in the sitting postures measured in other studies (upright, neutral, relaxed, slouched), however, was considered to pose a potential problem when making any comparisons of the sitting spinal curvatures [2, 55]. The standing postures were thought to yield more comparable data. The 33° mean standing thoracic kyphosis with

standard deviation (SD 11°) and 35° mean standing lordosis (SD 11.9°) measures obtained by the spinal wheel in this study are comparable with the 33°–36° kyphosis and the 28°–32° lordosis measures in previous studies [33, 56–59]. Ng et al. [17] demonstrated lower values of lumbar lordosis (24°–28°) which could be related to measuring only male subjects with arguably less pronounced standing lordosis [17].

The reliability of the spinal wheel in this study is comparable with radiographic techniques that reported intra- and inter-observer ICC of 0.81–0.99 for lumbar lordosis [59, 60] and 0.88–0.94 for thoracic kyphosis [59, 61] and inclinometer measures of lumbar lordosis with ICC score of 0.95 [17]. The absence of the measurement error analysis in these studies, however, makes direct comparisons of reliability difficult. Comparable inter-examiner error of 4.3° was produced by a similar skin-surface Spinal Mouse device when measuring a spinal posture in standing [33]. Spinal curvature flexicurve measurements demonstrated good intra- and inter-tester reliability with errors of 3°–4° in upper and lower lumbar spine [31]. In addition, the intra- and inter-tester typical error in this study (between 1.7 and 4.7°) has been considered to be acceptable [62, 63].

Despite that a difference between spinal measures on different days (test–retest reliability) is of clinical

importance to establish a change over time, research comparing between-day measures is limited [62]. Results from one study examining between-day radiographic measures of thoracic and lumbar curvatures demonstrated comparable ICC values (0.84–0.91), although again, measurement error was not reported [59]. Other skin-surface device, the Spinal Mouse, demonstrated somewhat lower between-day mean ICC value of 0.64 in spinal flexion measurement, compared to the spinal wheel [33].

Thoracic curvature measures demonstrated consistently better intra-, inter- and test–retest reliability compared to lumbar curvature measures with the highest errors in sitting lumbar curvatures. This observation was also made by Leroux et al. [58] who reported ICC of 0.94 for kyphosis and 0.91 for lordosis with errors of 5° and 6°, respectively, when comparing radiographic and non-invasive stereographic techniques [58]. Relatively large lumbar spine movement in sagittal plane, with rather small contributions from rather more rigid thoracic region [64], could be associated with larger measurement error as measurement errors tend to increase with increased magnitude of the measured variable [65]. In addition, the subcutaneous fat that typically accumulates in the around the waistline and lumbar region (relative to the thoracic region) [66], could have attributed to the differences in the accuracy between the regions. Indeed, subcutaneous fat was suggested to introduce a source error when using any hand-held, skin-surface device to measure joint range [33, 67]. In this study, all subjects were active and fit individuals with BMI within normal limits and no significant difference in the BMI between the sexes. Although gender differences may exist in the relative distribution of subcutaneous fat in the body regardless of BMI level [66], no difference was observed in the consistency of the measures obtained from either sex in this study. Nevertheless, findings in this study have to be interpreted with caution when related to measures from obese subjects. Future research that compares the reliability of the spinal wheel in subjects of varied BMI is therefore warranted.

Factors that could contribute to the measurement error in this study include the accuracy of the tool, the tester variability, as well as subject's variation of in the sitting and standing postures. With regards to the accuracy of the actual device, preliminary tests that were part of the device development, demonstrated minimal variation (coefficient of variation of 7%) and mean intra-tester error not exceeding 0.3° when measuring a static, rigid, spine-like shape. Such accuracy is comparable with the most accurate measurement techniques [35, 41, 42].

The tester variability, such as the level of experience in using the device, could introduce possible error. The speed at which the device is moved along the spine and the exerted pressure on the device and the path selected by

each assessor were suggested as possible sources of error in other skin-surface devices [33, 44]. In this study, the moving marker was detected by Vicon 512 which is a highly reliable 3-D kinematic system capable of monitoring speeds of up to 50 Hz. It is therefore unlikely that a speed of a moving marker would be a source of error. The variation in the pressure applied on the device while measuring the spinal curvature, could have also introduced the measurement error. Although, the amount of pressure applied was not monitored within- or between-testers, all three testers were experienced and skilled physiotherapists and performed a number of practice trials with specific instructions to maintain the same pressure while measuring the curvature. Nevertheless, it has to be stressed that relative level of skill and experience as well as the amount of practice trials was shown to effect accuracy of spinal measures [25]. Thus, the findings in this study should be interpreted with caution when relating to measures taken by testers of varied skill set and experience.

In this study, the spinal levels C7, T12 and S1 were marked to obtain thoracic and lumbar curvature measures. Accurate identification and marking of bony landmarks in the spine and pelvis has been argued to pose a substantial risk to the between-day reliability of any kinematic measurement [25, 43, 44, 47]. Harlick et al. [43] reported 47% accuracy of a manipulative physiotherapist in identifying spinous processes in the lumbar spine, and suggested that marker placement error could reduce the between-day reliability [43]. Therefore, guidelines outlined in previous literature concerned with reliable identification and placement of spinal and pelvic markers were followed in this study [47].

Human performance variability from one measurement to another was reported as a significant element that may seemingly distort the accuracy of any tested tool [25]. Previous literature showed that healthy subjects tend to slump into lumbar flexion during prolonged sitting and also demonstrate repositioning errors in sitting and standing postures [41, 45, 68, 69]. Maffey-Ward et al. [68] found a between-day error of 2.6° during sitting while Swinkels and Dolan [41] showed between-day errors of $5.27 \pm 3.47^\circ$ during lumbar spine sagittal movement. Variations in subjects' posture in this study, where subjects were required to maintain the same position for about 2 min while three testers obtained required measurements might have also contributed to reduced inter-rater and test–retest reliability. To avoid a systematic error, the order of the testers was randomised to stop one assessor always taking the last measure when the subject could have slumped. It is important to stress that this by no means reflects on the accuracy of the test device itself but rather highlights the utmost importance of standardizing the experimental procedure including the tested positions. This may be of

particular importance when measuring spinal curvatures of patients with low back pain who may have reduced ability to maintain any spinal posture due to poor muscle endurance [1], thus introducing further potential for measurement error. Therefore, studies assessing reliability of the spinal wheel in patient population need to consider possible ways of standardizing the postures that are measured.

Concerning clinical applicability, although the link between posture and back pain is not clearly established, it has been consistently demonstrated that patients with spinal pain disorders have a difficulty in accurately identifying neutral (mid-range) spinal postures [39, 45]. This was suggested to contribute to increased strain on the spinal tissues, potentially causing further damage and injury [6, 70] and also to influence the trunk muscle activity role in spinal control and stability [2, 5], which may perpetuate the spinal pain disorder and contribute to more disability [2, 4–6]. Establishing an accurate measure of a baseline posture that patients report to be painful may be invaluable with respect to evaluating the efficacy of postural re-education interventions. In addition, as the spinal wheel uses two-dimensional (2-D) coordinates to calculate sagittal spinal curvature, the device may be used in conjunction with a standard 2-D camcorder, thus is easily transferrable into clinical setting. Such measurement could provide a visual evaluation of patients' postures to provide feedback during postural re-education. In addition, an objective measure of spinal postures would be useful in further exploration of sub-classification systems that use visual observation of posture to sub-classify patients into different motor impairment subsets [12].

Conclusion

The spinal wheel device demonstrated excellent within-day, intra- and inter-tester reliability and substantial between-day, test–retest reliability when measuring thoracic and lumbar postures in healthy individuals during sitting and standing. Within-day, intra- and inter-tester typical error did not exceed 5°, which has been deemed to be acceptable. Between-day measures demonstrated slightly higher errors ranging between 4.0° and 7.4°. The varying degree of error in the inter-tester measures may suggest the need for training of the testers in using the device while higher errors in the between-day measures may be attributed to marker placement errors and also subjects' variability in adopting sitting and standing postures on different days, potentially warranting reservations in the interpretation of particularly the between-day measures. Given that on the premise of optimizing spinal alignment to minimize stresses on its structures, postural re-education is an integral part of physiotherapy management in patients with back pain, the

spinal wheel could potentially have its uses in measuring spinal postures to monitor the effectiveness and give guidance in progression of postural re-education interventions in clinical practice.

Acknowledgments The authors thank to Mark Thomas at the Departments of Medical Physics, University Hospital of Wales, Cardiff for his competent and prompt help with the development of the spinal wheel device. Funding was received from SiliconCoach. The study was approved by the Cardiff University Ethical Committee and all subjects were required to provide informed consent.

References

1. O'Sullivan PB, Mitchell T, Bulich P, Waller R, Holte J (2006) The relationship between posture and back muscle endurance in industrial workers with flexion-related low back pain. *Man Ther* 11:264–271
2. Dankaerts W, O'Sullivan P, Burnett A, Straker L (2006) Differences in sitting postures are associated with nonspecific chronic low back pain disorders when patients are subclassified. *Spine* 31:698–704
3. Dankaerts W, O'Sullivan P, Burnett A, Straker L (2006) Altered patterns of superficial trunk muscle activation during sitting in nonspecific chronic low back pain patients: importance of sub-classification. *Spine* 31:2017–2023
4. Panjabi MM (2003) Clinical spinal instability and low back pain. *J Electromyogr Kinesiol* 13:371–379
5. O'Sullivan PB, Grahamslaw KM, Kendell M, Lapenskie SC, Moller NE, Richards KV (2002) The effect of different standing and sitting postures on trunk muscle activity in a pain-free population. *Spine* 27:1238–1244
6. Panjabi MM (1992) The stabilizing system of the spine. Part II. Neutral zone and instability hypothesis. *J Spinal Disord* 5:390–396 discussion 397
7. Vezina MJ, Hubley-Kozey CL (2000) Muscle activation in therapeutic exercises to improve trunk stability. *Arch Phys Med Rehabil* 81:1370–1379
8. van Dieen JH, Cholewicki J, Radebold A, van Dieen JH, Cholewicki J, Radebold A (2003) Trunk muscle recruitment patterns in patients with low back pain enhance the stability of the lumbar spine. *Spine* 28:834–841
9. O'Sullivan PB, Twomey L, Allison GT (1998) Altered abdominal muscle recruitment in patients with chronic back pain following a specific exercise intervention. *J Orthop Sports Phys Ther* 27: 114–124
10. O'Sullivan PB, Phytty GD, Twomey LT, Allison GT (1997) Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis. *Spine* 22:2959–2967
11. Dankaerts W, O'Sullivan PB, Straker LM, Burnett AF, Skouen JS (2006) The inter-examiner reliability of a classification method for non-specific chronic low back pain patients with motor control impairment. *Man Ther* 11:28–39
12. O'Sullivan P (2006) Classification of lumbopelvic pain disorders—why is it essential for management? *Man Ther* 11:169–170
13. Ordway NR, Seymour R, Donelson RG, Hojnowski L, Lee E, Edwards TW (1997) Cervical sagittal range of motion analysis using three methods: cervical range of movement device, 3Space and Radiography. *Spine* 22:501–508
14. Portek I, Percy MJ, Reader GP, Mowat AG (1983) Correlation between radiographic and clinical measurement of lumbar spine movement. *Rheumatology* 22:197–205

15. Stokes IAF, Wilder DG, Frymoyer JW, Pope MH (1981) Assessment of patients with low-back-pain by biplanar radiographic measurement of intervertebral motion. *Spine* 6:233–240
16. Kachingwe AF, Phillips BJ (2005) Inter- and intrarater reliability of a back range of motion instrument. *Arch Phys Med Rehabil* 86:2347–2353
17. Ng JKF, Kippers V, Richardson CA, Parnianpour M (2001) Range of motion and lordosis of the lumbar spine: reliability of measurement and normative values. *Spine* 26:53–60
18. Fitzgerald GK, Wynveen KJ, Rheault W, Rothschild B (1983) Objective assessment with establishment of normal values for lumbar spinal range of motion. *Phys Ther* 63:1776–1781
19. Tousignant M, Morissette J, Murphy M (2002) Criterion validity study of lumbar goniometers BROM II and EDI-320 for range of motion of lumbar flexion of low back pain patients. *J Back Musculoskelet Rehabil* 16:159–167
20. Troup JD, Hood CA, Chapman AE (1968) Measurements of the sagittal mobility of the lumbar spine and hips. *Ann Phys Med* 9:308–321
21. Troup JD, Foreman TK, Baxter CE, Brown D (1987) The perception of back pain and the role of psychophysical tests of lifting capacity. *Spine* 12:645–657
22. Adams MA, Dolan P, Marx C, Hutton WC (1986) An electronic inclinometer technique for measuring lumbar curvature. *Clin Biomech* 1:130–134
23. Chen SPC, Samo DG, Chen EH, Crampton AR, Conrad KM, Egan L, Mitton J (1997) Reliability of three lumbar sagittal motion measurement methods: Surface inclinometers. *J Occup Environ Med* 39:217–223
24. Ensink FBM, Saur PMM, Frese K, Seeger D, Hildebrandt J (1996) Lumbar range of motion: Influence of time of day and individual factors on measurements. *Spine* 21:1339–1343
25. Mayer TG, Kondraske G, Beals SB, Gatchel RJ (1997) Spinal range of motion—accuracy and sources of error with inclinometric measurement. *Spine* 22:1976–1984
26. Miller SA, Mayer T, Cox R, Gatchel RJ (1992) Reliability problems associated with the modified Schober technique for true lumbar flexion measurement. *Spine* 17:345–348
27. Williams R, Binkley J, Bloch R, Goldsmith CH, Minuk T (1993) Reliability of the modified-modified Schober and Double inclinometer methods for measuring lumbar flexion and extension. *Phys Ther* 73:33–44
28. Burton AK (1986) Regional lumbar sagittal mobility—measurement by flexicurves. *Clin Biomech* 1:20–26
29. Burton AK, Tillotson KM, Troup JD (1989) Variation in lumbar sagittal mobility with low-back trouble. *Spine* 14:584–590
30. Lundon KM, Li AM, Bibershtein S (1998) Interrater and intrarater reliability in the measurement of kyphosis in postmenopausal women with osteoporosis. *Spine* 23:1978–1985
31. Tillotson KM, Burton AK (1991) Noninvasive measurement of lumbar sagittal mobility. An assessment of the flexicurve technique. *Spine* 16:29–33
32. Burton AK, Tillotson KM (1989) Is recurrent low back trouble associated with increased lumbar sagittal mobility? *J Biomed Eng* 11:245–248
33. Mannion A, Knecht K, Balaban G, Dvorak J, Grob D (2004) A new skin-surface device for measuring the curvature and global and segmental ranges of motion of the spine: reliability of measurements and comparison with data reviewed from the literature. *Eur Spine J* 13:122–136
34. An KN, Jacobsen MC, Berglund LJ, Chao EYS (1988) Application of a magnetic tracking device to kinesiological studies. *J Biomech* 21:613
35. Fahrenberg J, FF, Smeja M, Muller V (1997) Assessment of posture and motion by multichannel piezoresistive accelerometer recordings. *Pathophysiology* 34:607–612
36. Mannion A, Troke M (1999) A comparison of two motion analysis devices used in the measurement of lumbar spinal mobility. *Clin Biomech* 14:612–619
37. Swinkels A, Dolan P (2004) Spinal position sense and disease progression in ankylosing spondylitis—a longitudinal study. *Spine* 29:1240–1245
38. Dolan P, Greenfield K, Nelson RJ, Nelson IW (2000) Can exercise therapy improve the outcome of microdiscectomy? *Spine* 25:1523–1532
39. Newcomer KL, Laskowski ER, Yu B, Johnson JC, An KN (2000) Differences in repositioning error among patients with low back pain compared with control subjects. *Spine* 25:2488–2493
40. Pearcy MJ, Hindle RJ (1989) New method for the non-invasive 3-dimensional measurement of human back movement. *Clin Biomech* 4:73–79
41. Swinkels A, Dolan P (2000) Spinal position sense is independent of the magnitude of movement. *Spine* 25:98–104 discussion 105
42. Newcomer K, Laskowski ER, Yu B, Larson DR, An KN (2000) Repositioning error in low back pain. Comparing trunk repositioning error in subjects with chronic low back pain and control subjects. *Spine* 25:245–250
43. Harlick JC, Milosavljevic S, Milburn PD (2007) Palpation identification of spinous processes in the lumbar spine. *Manual Ther* 12:56–62
44. Mayer RS, Chen IH, Lavender SA, Trafimow JH, Andersson GBJ (1995) Variance in the measurement of sagittal lumbar spine range of motion among examiners, subjects, and instruments. *Spine* 20:1489–1493
45. O'Sullivan PB, Burnett A, Floyd AN, Gadsdon K, Logiudice J, Miller D, Quirke H (2003) Lumbar repositioning deficit in a specific low back pain population. *Spine* 28:1074–1079
46. Swinkels A, Dolan P (1998) Regional assessment of joint position sense in the spine. *Spine* 23:590–597
47. Vergara M, Page A, Sancho JL (2006) Analysis of lumbar flexion in sitting posture: Location of lumbar vertebrae with relation to easily identifiable skin marks. *Int J Ind Ergonom* 36:937–942
48. Kellis E, Adamou G, Tziliou G, Emmanouilidou M (2008) Reliability of spinal range of motion in healthy boys using a skin-surface device. *J Manipulative Physiol Ther* 31:570–576
49. Shrout PE, Fleiss JL (1979) Intraclass correlation: uses in assessing rater reliability. *Psychol Bull* 86:420–426
50. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
51. Batterham AM, George KP (2003) Reliability in evidence based clinical practice: a primer for allied health professions. *Phys Ther Sport* 4:122–128
52. Hopkins WG (2000) Measures of reliability in sports medicine and science. *Sports Med* 30:1–15
53. Stratford PW, Goldsmith CH (1997) Use of the standard error as a reliability index of interest: an applied example using elbow flexor strength data. *Phys Ther* 77(6):745
54. Walter SD, Eliasziw M, Donner A (1998) Sample size and optimal designs for reliability studies. *Stat Med* 17:101–110
55. Allison GT, Fukushima S (2003) Estimating three-dimensional spinal repositioning error: the impact of range posture and number of trials. *Spine* 28:2510–2516
56. Dolan KJ, Green A (2006) Lumbar spine reposition sense: the effect of a slouched posture. *Manual Ther* 11:202–207
57. Hultman G, Saraste H, Ohlson H (1992) Anthropometry, Spinal-canal width, and flexibility of the spine and hamstring muscles in 45–55-year-old men with and without low-back-pain. *J Spinal Disord* 5:245–253
58. Leroux MA, Zabjek K, Simard G, Badeaux J, Coillard C, Rivard CH (2000) A noninvasive anthropometric technique for measuring kyphosis and lordosis: an application for idiopathic scoliosis. *Spine* 25:1689–1694

59. Pinel-Giroux FM, Mac-Thiong JM, de Guise JA, Berthonnaud E, Labelle H (2006) Computerized assessment of sagittal curvatures of the spine: comparison between Cobb and tangent circles techniques. *J Spinal Disord Tech* 19:507–512
60. Polly DW Jr, Kilkelly FX, McHale KA, Asplund LM, Mulligan M, Chang AS (1996) Measurement of lumbar lordosis. Evaluation of intraobserver, interobserver, and technique variability. *Spine* 21:1530–1535 discussion 1535–6
61. Goh S, Price RI, Leedman PJ, Singer KP (2000) A comparison of three methods for measuring thoracic kyphosis: implications for clinical studies. *Rheumatology* 39:310–315
62. Carman DL, Browne RH, Birch JG (1990) Measurement of scoliosis and kyphosis radiographs— intraobserver and interobserver variation. *J Bone Joint Surg* 72A:328–333
63. Morrissy RT, Goldsmith GS, Hall EC, Kehl D, Cowie GH (1990) Measurement of the Cobb angle on radiographs of patients who have scoliosis—evaluation of intrinsic error. *J Bone Joint Surg* 72A:320–327
64. Dolan P (1993) Influence of lumbar and hip mobility on the bending stresses acting on the lumbar spine. *Clin Biomech* 8:185–192
65. Atkinson G, Nevill AM (1998) Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 26:217–238
66. Power ML, Schulkin J (2008) Sex differences in fat storage, fat metabolism, and the health risks from obesity: possible evolutionary origins. *Br J Nutr* 99:931–940
67. Kuo Y-L, Tully EA, Galea MP (2008) Skin movement errors in measurement of sagittal lumbar and hip angles in young and elderly subjects. *Gait Posture* 27:264–270
68. Maffey-Ward L, Jull G, Wellington L (1996) Toward a clinical test of lumbar spine kinesthesia. *J Orthop Sports Phys Ther* 24:354–358
69. O'Sullivan P, Dankaerts W, Burnett A, Chen D, Booth R, Carlsen C, Schultz A (2006) Evaluation of the flexion relaxation phenomenon of the trunk muscles in sitting. *Spine* 31:2009–2016
70. Panjabi MM (1992) The stabilizing system of the spine Part I Function, dysfunction, adaptation, and enhancement. *J Spinal Disord* 5:383–389 discussion 397