

Original Article

# Reliability of assisted indentation in measuring lumbar spinal stiffness

Tasha R. Stanton<sup>1</sup>, Gregory N. Kawchuk\*

University of Alberta, Department of Physical Therapy, Faculty of Rehabilitation Medicine, 3-44 Corbett Hall, Common Spinal Disorders Lab, Edmonton, Alberta T6G 2G4, Canada

Received 27 June 2007; received in revised form 22 January 2008; accepted 30 January 2008

## Abstract

The reliability of manual methods to assess spinal stiffness is modest at best. In response, instrumentation has been developed which may be reliable, but is often difficult to use in clinical settings. The purpose of this study was to determine the intra-rater reliability of assisted indentation (AI), a smaller, less automated technique of measuring spinal stiffness *in vivo*. Twenty-three asymptomatic subjects were included in the study. The AI device was placed over the 4th lumbar spinous process in each prone, resting subject. Ten indentations were performed at approximately 2-min intervals while load and displacement data were collected simultaneously. From these data, two outcome variables were calculated: Global Stiffness (GS; slope of the force–displacement data) and Mean Maximal Stiffness (MMS; peak force/peak displacement). Intra-class correlation coefficient values for 10 consecutive measures of GS and MMS were 0.93 and 0.91, respectively. A repeated measures analysis of variance (ANOVA) did not demonstrate significant differences between any indentation trials from the same subject. Measurement of spinal stiffness using AI demonstrated excellent intra-rater reliability. These data, in addition to specific features of AI (small, transportable, relatively low cost, ease of operation) suggest that AI may be of benefit within clinical environments.

© 2008 Elsevier Ltd. All rights reserved.

**Keywords:** Indentation; Assisted indentation; Reliability; Spinal stiffness; Posteroanterior compression

## 1. Background and purpose

The manual assessment of low back stiffness remains a key tenet for many professionals who diagnose and treat low back pain. Most often, the clinical assessment of spinal stiffness involves a manual pressure test where a clinician uses their hands to apply pressure in a posteroanterior (PA) direction to the spinous process of interest. During the application of PA pressure, the clinician appreciates the resulting tissue response and

forms a subjective impression of spinal stiffness. The resulting impression formed by the clinician during the pressure test is then used to judge if the spine is too compliant (hypermobility), too stiff (hypomobility), or within normal limits (Maitland et al., 2001). These judgments often provide a basis for individual treatment programs and have also been shown to be important in predicting therapeutic success when stabilization exercise programs are prescribed (Hicks et al., 2005).

Unfortunately, the PA pressure test is based on human performance, interpretation and communication. As a result, the PA pressure test is highly variable in many respects including the magnitude of applied peak force, (Latimer et al., 1998) the direction of force application (Caling and Lee, 2001) and in the

\* Corresponding author. Tel.: +1 780 492 6891; fax: +1 780 492 4492.

E-mail address: [greg.kawchuk@ualberta.ca](mailto:greg.kawchuk@ualberta.ca) (G.N. Kawchuk).

<sup>1</sup> University of Sydney, School of Physiotherapy, Faculty of Health Sciences, East Street, Lidcombe, NSW 2141, Australia.

identification of a specific spinous process (Harlick et al., 2007) as a PA pressure target. In addition, the level of human sensitivity in detecting alterations in stiffness is limited. It has been estimated that the discrimination threshold for stiffness is of the order of 11% when using a pisiform grip to evaluate stiffness in the range of 6–11 N/mm (Maher and Adams, 1995). As a result, clinicians may be unable to perceive significant changes in spinal stiffness that occur below this threshold.

Given the above, it is not surprising that stiffness values obtained from manual assessment of spinal stiffness vary considerably between clinicians (Snodgrass et al., 2006). Specifically, studies of between-clinician agreement have shown that the reliability of stiffness assessment remains poor with intra-class correlation coefficients (ICC) ranging between 0.03 and 0.55 (Fleiss, 1986; Maher and Adams, 1994; Binkley et al., 1995).

In response to the poor reliability (Fleiss, 1986; Maher and Adams, 1994; Binkley et al., 1995), large variability (Snodgrass et al., 2006), and limits of human perception (Maher and Adams, 1995) associated with manual assessment of spinal stiffness, mechanical instruments have been designed to measure the applied loads and resulting tissue deformations that occur during manual PA testing. These devices include the Spinal Physiotherapy Simulator (SPS) (Lee and Svensson, 1990), Lee and Evans' stiffness assessment device (Lee and Evans, 1992), Stiffness Assessment Machine (SAM) (Latimer et al., 1996a–c), Spinal Posteroanterior Mobilizer (SPAM) (Edmondston et al., 1998), and the Rigid Frame Indentor (Kawchuk and Herzog, 1996). While the reliability of the majority of these instruments is high, these devices are designed primarily for research applications. As a result, many features of these devices such as their size, expense, and complex operation preclude their use in clinical settings.

To exploit the increased performance of mechanical devices in assessing spinal stiffness yet avoid the limitations common to these research-based devices, a new stiffness assessment technique is proposed. This technique, assisted indentation (AI), uses manual load application with the addition of instrumentation designed to assist the operator and improve reliability and accuracy. Given recent findings that indicate stiffness may be a variable which helps predict outcome success (Childs et al., 2004; Hicks et al., 2005), there may be a future clinical need for a device which can measure stiffness accurately and reliably. Although the accuracy of AI has been shown to be excellent (absolute maximal difference of 0.22 mm compared to gold standard) (Kawchuk et al., 2006), the reliability of AI has yet to be determined. Therefore, the purpose of this study was to measure the *in vivo*, with-in operator reliability of AI measurements of spinal stiffness. It was hypothesized that AI reliability would be excellent (ICC greater than 0.75) (Fleiss, 1986).

## 2. Methods

### 2.1. Subjects

Following approval from the University of Alberta Health Research Ethics Board, 23 consenting subjects were recruited from the University of Alberta and surrounding area over a 1-month period. This sample size was calculated *a priori* using a power of 80% and a level of significance of  $p = 0.05$ .

#### 2.1.1. Inclusion criteria

Study subjects included asymptomatic males and females between the ages of 18 and 30 with no history of low back pain within the last year as well as no current low back pain.

#### 2.1.2. Exclusion criteria

Subjects were excluded from this study if they reported back pain and/or medical conditions that could affect the safety of measurement of spinal stiffness using AI and/or intolerance to screening procedures designed to identify those persons sensitive to direct spinous process loading. Please refer to Table 1 for a detailed list of exclusion criteria.

### 2.2. Research design

This study quantified (1) single operator reliability of AI measures in a sample human population and (2) repeatability of AI measures generated by a single operator within single subjects.

### 2.3. Instrumentation

A description of the device used to perform AI has been published previously (Fig. 1) (Kawchuk et al., 2006). In brief, the AI equipment is made up of an outer

Table 1  
Exclusion criteria.

Injury related	Disease processes (Maitland, 1986)	Subject factors
Current low back pain	Osteoporosis	Pregnancy (unsure or confirmed)
Low back pain within the last year	Osteoarthritis	Medications affecting muscle function (e.g. steroids)
Previous back surgery	Rheumatoid arthritis	Medications affecting pain recognition (e.g. pain medications)
Lower extremity injury within the last year	Ankylosing Spondylitis	Unable to tolerate indentation
	Known malignancy	
	Known spondylolisthesis	
	Multiple sclerosis	
	Severe scoliosis	

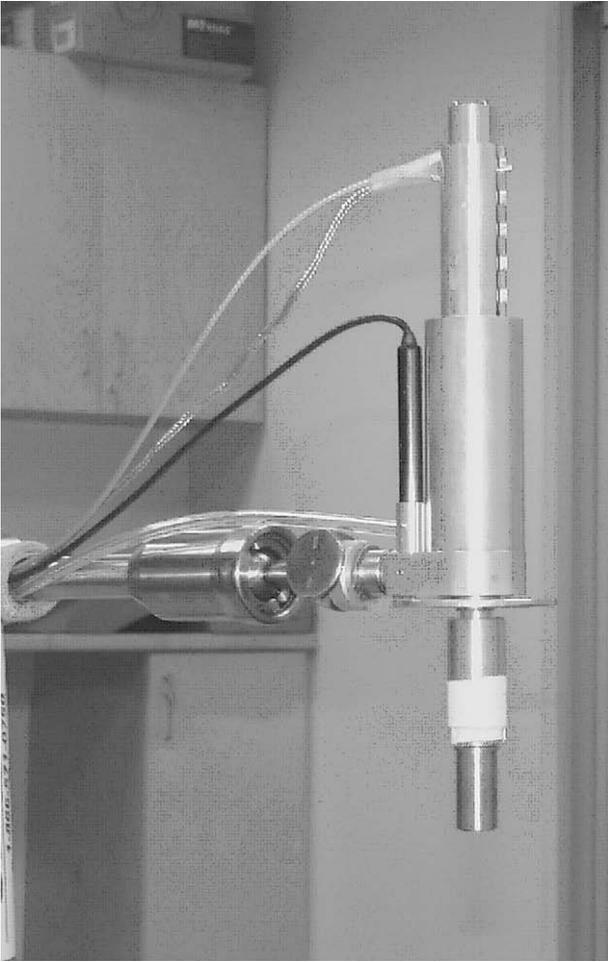


Fig. 1. Assisted Indentor.

frame that is supported by an external support arm (Tenet Medical Engineering, Calgary, Alberta, Canada). The use of this rigid arm creates a stationary reference point. These structures suspend an inner probe that is moved manually to apply an external force to the anatomical target of interest. By using a ceramic air-bearing to hold the indenting probe, near frictionless movement of the inner probe with respect to the outer frame can be achieved thereby reducing artifacts due to movement of the frame during indentation loading. To measure applied force, a compressive-tension load cell (Entran, Fairfield, NJ) is connected in-series with the probe. The displacement of the probe is measured by a linear variable differential transformer (LVDT) (Honeywell International Inc., Morristown, NJ) attached between the probe and the outer housing. Because the displacement of the indenter is initiated by a manual process, but restricted by mechanical boundaries, this form of indentation is called “Assisted Indentation”. Signals from the load cell and the LVDT were conditioned appropriately and collected by customized LABview software (National Instruments, Austin, TX) at a collection rate of 200 Hz.

#### 2.4. Calibration

Calibration of the assisted indentation device was achieved using masses of known magnitude applied to the load cell and spacers of known dimensions applied to the LVDT. After each application of increasing calibration mass or dimension, force and displacement signals were collected then plotted against the known mass or dimension. These data were then modeled with a linear data fitting technique. In each case, the  $r^2$  value of the line of best fit was greater than 0.90. The resulting equation of the line of best fit was then used to determine the units of measure for the output voltage of each transducer. Calibration was completed prior to subject testing.

#### 2.5. Spinal stiffness measurement

In each prone subject, the AI device was placed perpendicular to the L4 spinous process with a contact load of less than 1 N (Fig. 2). The subject was then instructed to breathe out comfortably then to hold his/her breath for the duration of the indentation (approximately 5 s) (Kawchuk and Fauvel, 2001). During indentation, the indentation probe was advanced manually (approximately 2 mm/s) into the spine until a force threshold of 100 N was read from a visual indicator. This level of force application was considered to be safe as forces up to 200 N have been used within an asymptomatic human population (Latimer et al., 1998) and forces up to 105 N within a symptomatic human population (Latimer et al., 1996b) without any adverse



Fig. 2. Placement of the Assisted Indentor on a subject.

effects reported. When the 100 N threshold was reached, the indenter position was maintained at this load for approximately 1 s after which the indenter was removed from contacting the subject. To decrease variability in the rate of indentation loading, the equipment operator viewed a computerized bar graph which increased in size at a rate of 2 mm/s. Next to this graph, a second bar graph displayed the actual displacement of the AI device. With these two displays, the operator could continually adjust their performance to match the desired indentation rate.

### 2.6. Study procedure

Once informed consent was attained from the subjects, a verbal history questionnaire was completed to ensure that subjects met the inclusion criteria and did not possess any factors that would cause exclusion from the study. Following the questionnaire, each subject's height and weight were recorded, and Body Mass Index (BMI; kg/m<sup>2</sup>) was calculated (Astrand et al., 2003).

With the subject lying in prone on a plinth, the subject's spine was palpated by the researcher and the L4 spinous process identified. Although identification of spinous processes in the lumbar spine has demonstrated moderate accuracy with use of preferred palpation procedures (47% were on the level intended) (Harlick et al., 2007), a standardized procedure was utilized in this study to reduce this error. Specifically, the horizontal line between the iliac crests was used to identify the L4/5 interspace and the vertebrae above were determined to be L4 (if this line between the iliac crests gave a spinous process, this was identified as L4) (Grieves, 1984). The L4 vertebra was chosen as the site of indentation as this has been shown to be a commonly symptomatic area in patients with low back pain (Maitland et al., 2001).

The skin over the presumed L4 spinous was then marked using a pen to provide a visual guide for placing the indenter. The indenter was then placed over the ink marking and a series of five consecutive indentations were provided to familiarize subjects with the indentation process. Once the familiarization indentations were completed, 10 consecutive spinal stiffness measurements (indentations) were collected, each separated by a time period of approximately 2 min. During times between indentations, subjects were instructed to remain in a resting prone position and to remain stationary and relaxed. Each subject was examined at one time period. Indentations were performed by one researcher (T.S.) who had logged approximately 100 h of using the indentation device prior to data collected for this experiment.

During the indentation process, all subjects held an analog trigger to indicate if their level of discomfort

during indentation increased from baseline. If the subject wanted indentation to cease for any reason they were instructed to squeeze the trigger fully which produced an audible alarm alerting the researcher to remove the indenter. If this situation occurred, the researcher re-positioned the indenter and indentation was attempted again. Re-positioning of the indenter was allowed a maximum of two times after which further indications of painful indentation excluded the subject from further participation.

### 2.7. Analysis of spinal stiffness measurement

Indentation data (force and displacement) were used to calculate the spinal stiffness at the indentation site. Stiffness was quantified in two ways: (1) Global Stiffness (GS); and (2) Mean Maximal Stiffness (MMS). GS, calculated as the slope of the force–displacement curve between 30 N and maximal force, represents the stiffness of the underlying tissues during the indentation itself. It is assumed that the relationship between force and displacement is linear between 30 and 100 N given previous work. (Latimer et al., 1996b). MMS, the second variable representing stiffness, was computed by taking the average stiffness value (N/mm) over the time period where the maximal indentation force has been held for a period of approximately 1 s. The MMS variable is therefore a ratio between the applied maximal force and the resultant maximal displacement of the underlying tissues (Fig. 3).

### 2.8. Statistical analysis

For data analysis purposes, all five of the familiarization trials were discarded (Latimer et al., 1996b,c). In addition, the first trial (stiffness measurement during rest) of the 10 experimental indentations was discarded as this trial has been shown to be highly variable (Latimer et al., 1996b,c) while stiffness measurements from subsequent trials (after the first trial) have demonstrated stability (Latimer et al., 1996b,c).

To assess intra-rater reliability of the researcher/instrument in measuring spinal stiffness, the intra-class correlation coefficient (3,1) was calculated (Shrout and Fleiss, 1979).

To describe repeatability, inter-trial inconsistency values for stiffness variables were calculated by taking the difference between two consecutive indentations expressed as a percentage of the average of the same two indentations.

Finally, to further explore repeatability and investigate the possibility that a gradual change in stiffness values may occur with successive indentations, a condition that may not be reflected in ICC values, a repeated measures analysis of variance (ANOVA) with a Bonferroni correction was performed.

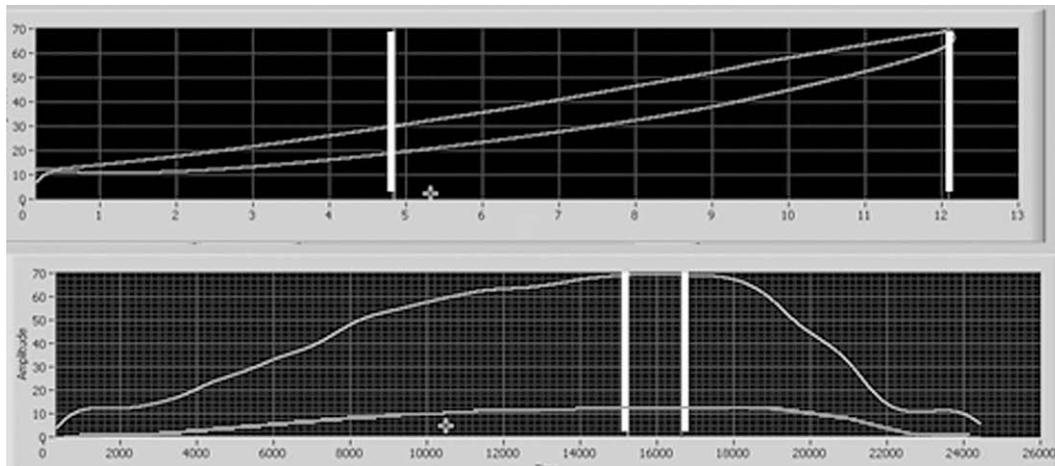


Fig. 3. Graphical example of stiffness measurement output. The top graph represents the load–displacement curve of a single AI (force on *y*-axis, displacement on *x*-axis) with vertical white bars representing the section over which slope was taken (GS). The bottom graph shows the indentation profile (numerical scale for force and displacement on *y*-axis, time on *x*-axis). In the bottom graph, the upper trace is the applied force while the bottom trace is the resultant displacement. The vertical white bars represent the section over which force was divided by displacement (MMS).

### 3. Results

A total of 30 subjects were recruited to participate in this project with three excluded due to previous back or lower extremity injury within the last year, two excluded for exceeding the age limit, and two excluded prior to formal testing (did not pass the indentation screening procedure in that they reported discomfort with indentation even after the indenter was re-positioned twice). This resulted in 12 male and 11 female subjects who participated in this study ( $n = 23$ ) (see Table 2 for subject demographics).

In this experiment, the reliability of the stiffness measures was described by the ICC which was calculated to be 0.91 for GS and 0.93 for MMS. Additionally, an estimate of the consistency in stiffness measures was obtained by calculating the inter-trial inconsistency value which was 6.23% ( $\pm 4.52\%$ ) for the GS and 7.71% ( $\pm 5.33\%$ ) for MMS (see Figs. 4 and 5 for individual subject representation of inter-trial inconsistency values).

The repeated measures ANOVA did not reveal significant differences between any indentation trials for either GS or MMS ( $p = 0.09$ – $1.00$  and  $p = 1.00$  for all comparisons, respectively). See Figs. 6 and 7 for the graphical representation of the change in stiffness values over time.

Table 2  
Mean ( $\pm$ standard deviation) of subject demographic characteristics.

	Male ( $n = 12$ )	Female ( $n = 11$ )
Age (years)	26.17 ( $\pm 3.10$ )	24.45 ( $\pm 3.21$ )
Height (m)	1.79 ( $\pm 0.065$ )	1.63 ( $\pm 0.052$ )
Weight (kg)	76.23 ( $\pm 9.64$ )	58.41 ( $\pm 8.28$ )
BMI ( $\text{kg}/\text{m}^2$ )	23.85 ( $\pm 2.39$ )	21.59 ( $\pm 2.21$ )

### 4. Discussion

Data from this study support the hypothesis that AI has excellent reliability ( $\text{ICC} \geq 0.75$ ) (Fleiss, 1986). Specifically, AI exhibited excellent intra-rater reliability for all outcome variables used to quantify L4 stiffness. Furthermore, the average inter-trial inconsistency remained below 8% for all stiffness variables.

Compared to the manual testing of stiffness, ICC values found for the AI technique were much higher (Table 3). Overall, reliability values for the evaluation of spinal stiffness using the manual PA pressure test have been found to be poor (Matyas and Bach, 1985; Maher and Adams, 1994; Binkley et al., 1995). Matyas and Bach (1985) first found poor reliability of manual PA stiffness assessment when they reported Pearson's  $r$  ranging from 0.09 to 0.46. Unfortunately, these reliability results using Pearson's  $r$  cannot be compared directly to the current study. Later studies also noted poor reliability with ICC (1,1) values ranging from 0.03 to 0.37 (Maher and Adams, 1994; Binkley et al., 1995). With improvements to the testing protocol and delineation of stiffness into ranges, reliability increased to a fair level (Fleiss, 1986) with an ICC value reported to be 0.55 (range 0.50–0.62) (Maher et al., 1998). The ICC value of the PA pressure test increased further when an 11-point stiffness rating scale was employed and more rigorously controlled testing protocol were used ( $\text{ICC} = 0.77$ ) (Maher et al., 1998). Although improvements in the reliability of the manual assessment of spinal stiffness have been demonstrated, these improvements occur only under standardized and artificial conditions that are not typically employed in the clinical environment.

It may be argued that any form of instrumented stiffness assessment, such as AI, is also not typical of the clinical procedures (i.e. PA testing) due to increased

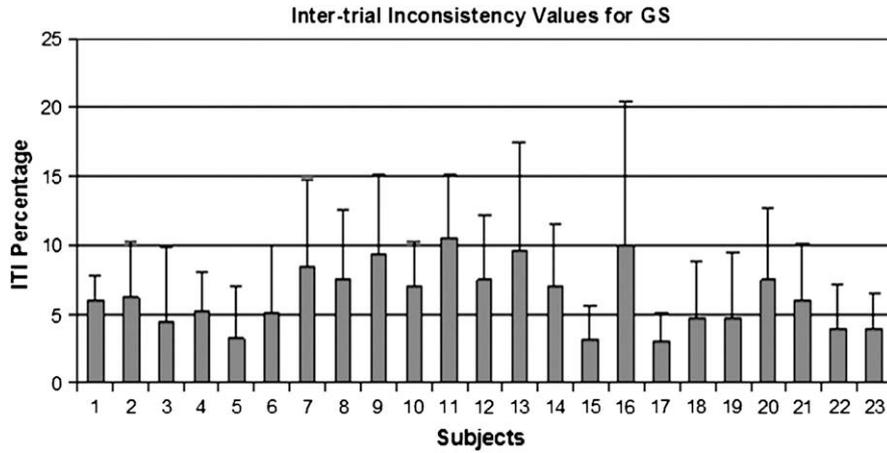


Fig. 4. Inter-trial inconsistency values (mean ± standard deviation) for GS estimates of L4 stiffness values.

size of the instrumentation and necessary operator training. However, if the desire is to objectively quantify stiffness in a reproducible way, then changing clinical practice to involve use of scales to delineate stiffness levels or involve use of an instrument becomes important. If changes to clinical practice are mandated and/or desirable, using a method of stiffness assessment with the combination of high reliability values and minimally clinically invasiveness is paramount. With this in mind, AI may become a viable option for clinical stiffness testing due to its excellent reliability values as well as a design that allows for ease of use by a single operator in a small footprint, low cost device (~\$10,000 Canadian dollars) that does not require advanced mechanization such as motors, pulleys or pistons.

The observation that AI exhibits greater reliability than manual assessment of spinal stiffness was expected for three reasons. First, AI measures several variables in an objective manner, increasing the reliability of spinal stiffness assessment. Specifically, use of technology to quantify force and displacement data (load cell and

a LVDT, respectively), in addition to customized computer programming, allows consistency of force application and real-time visualization of results. Second, AI reduces variability in factors shown to alter spinal stiffness measures including visual occlusion (Maher and Adams, 1996), peak force (Latimer et al., 1998), frequency of PA loading (Lee and Svensson, 1990; Lee and Liversidge, 1994), direction of force application (Caling and Lee, 2001), and force angulation (Kawchuk and Herzog, 1996). Finally, we elected to employ stiffness variables which considered regions of data that were larger than those used in previous studies. This approach was chosen because the most clinically important region of a load–displacement graph remains unknown. While there is some evidence to suggest that stiffness may play a role in predicting outcomes of specific treatments (Childs et al., 2004; Hicks et al., 2005), an understanding of the physiologic basis of spinal stiffness, or its alteration due to pathology or treatment, remains elusive.

With respect to other studies, the reliability values for AI, although slightly lower, are comparable to those

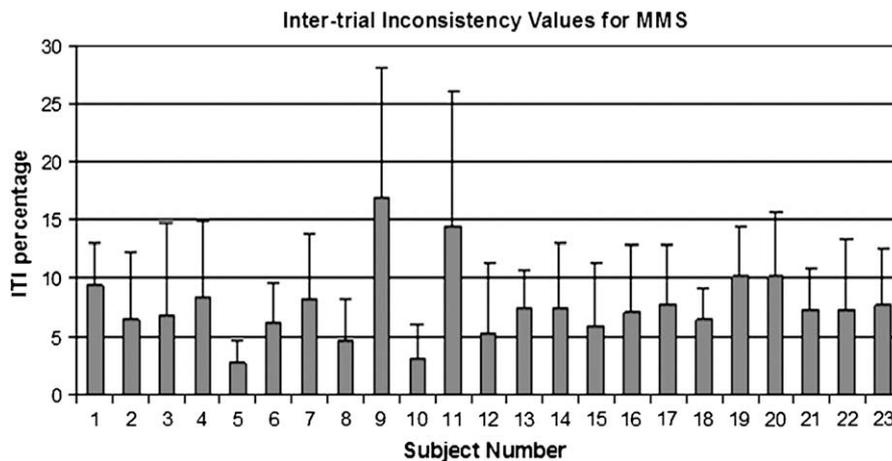


Fig. 5. Inter-trial inconsistency values (mean ± standard deviation) for MMS estimates of L4 stiffness values.

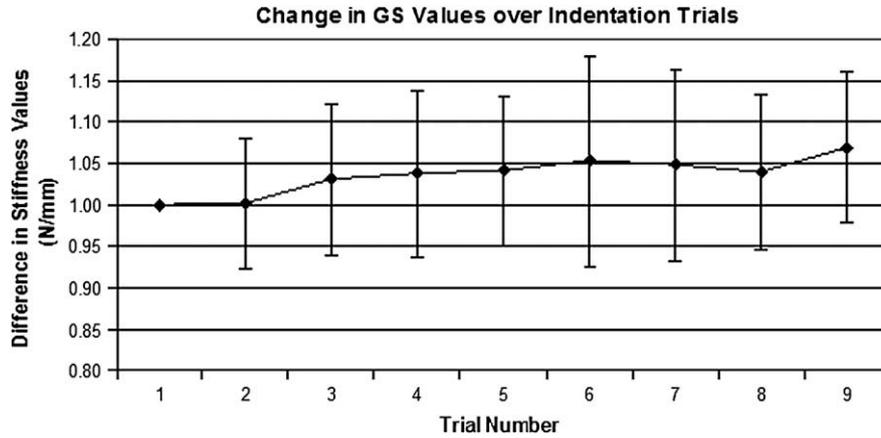


Fig. 6. Change in GS values over time for all subjects. GS values (mean  $\pm$  standard deviation) normalized to indentation trial 2.

found for mechanical indentation devices (Table 3). Intra-class correlation coefficient values have been reported to be over 0.90 for almost all mechanical indentation instruments. Specifically, the SPAM was found to have an ICC value of 0.979 at L5 (Edmondston et al., 1998), Lee and Evans' stiffness assessment device had an ICC value of 0.99 for L3/4 and 0.95 for L4/5 (Lee and Evans, 1992), SAM had an ICC value of 0.96 for lumbar vertebrae (Latimer et al., 1996a–c), and Rigid Frame Indentation at 0.99–1.00 for varying experimental conditions (Kawchuk and Herzog, 1996). Interestingly, the reliability of AI was higher than that of the SPS which found an ICC value of 0.88 at L3 (Lee and Svensson, 1990). That mechanical indentation devices have higher reliability values (overall) than AI is expected. While the rate of indentation studying the AI procedure is standardized using a visual cue (graphic display of force data), slight variations in the rate of indentation were likely to occur. These variations may alter the resulting measures as (1) the target tissues are viscoelastic and may exhibit rate dependant behaviors (White and Panjabi, 1990) and (2) variations in the data may influence stiffness analysis techniques such as GS which is based on a linear approximation of shape.

While these variations were not of sufficient magnitude to create poor reliability, they may account for the slightly lower reliability values that occur with AI compared to other automated techniques.

Further support for the comparability of AI to mechanical techniques is suggested by our repeated measures ANOVA results; no significant differences were present between any of the indentation trials for both GS and MMS measures. This observation suggests that all indentations using the AI found similar stiffness values regardless of the time at which the stiffness measure was taken. This finding strengthens the excellent reliability values by demonstrating consistency over time with the stiffness measurements. Further, these ANOVA data suggest that repeated AI trials do not affect viscoelasticity of the target tissue. This is likely due to tissues reaching a steady state of viscoelastic change following sufficient familiarization trials and experimental indentations and/or adequate time between all indentations such that between-trial tissue recovery was complete.

It should be noted that large differences in individual subject inter-trial inconsistency values were exhibited with some single subjects having inter-trial inconsistency values approaching 30% ( $\pm 1$  SD). This suggests that the

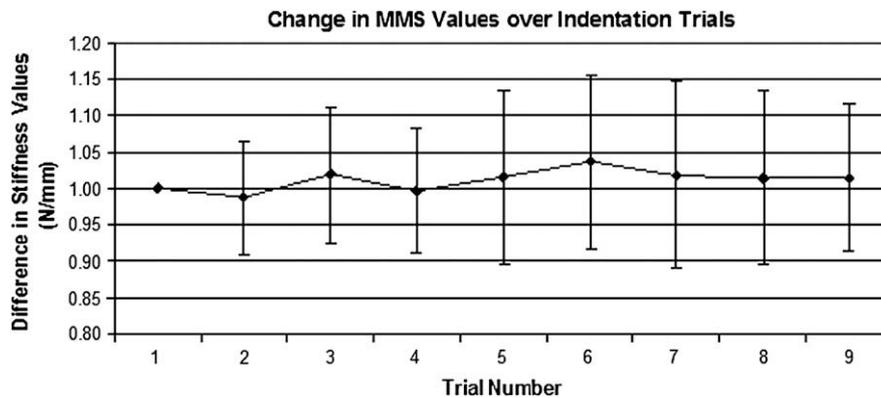


Fig. 7. Change in MMS values over time for all subjects. MMS values (mean  $\pm$  standard deviation) normalized to indentation trial 2.

Table 3  
Intra-class correlation values for three methods of stiffness assessment.

Method of stiffness assessment	ICC value
Manual	0.03–0.77
Mechanical	0.88–1.00
Assisted indentation	0.91–0.93

consistency of stiffness results obtained by AI may be specific to the individual and may be influenced by other factors not defined in this study. Possible factors that could explain measurement inconsistency with these few subjects may include inconsistent localization of the indentation contact point between trials or failure to control subject specific factors which influence stiffness (e.g. intra-abdominal pressure, muscle contraction, subject movement, etc.) (Kawchuk and Fauvel, 2001). In this situation, changes in measured spinal stiffness may occur as the indentation test may involve different anatomy. In addition, the subject's baseline stiffness could also be a confounding factor. Although a formal analysis was not performed, it was observed that those subjects with high baseline stiffness values for GS and MMS (stiff back) often had large changes in their stiffness values over time. Finally, variables such as plinth padding (Maher et al., 1999), subject positioning, (Edmondston et al., 1998), adipose tissue (Viner et al., 1997), and breathing (Beaumont et al., 1991) must be controlled within a single subject if stiffness measures are to be compared within the same subject over time.

Several limitations of this study are noted. First, only intra-operator reliability was measured in asymptomatic subjects. As a result, we cannot comment on intra-operator reliability in a symptomatic population nor inter-operator reliability. Second, our results apply specifically to a sample of patients with an average age of 25 years and average BMI of 22 kg/m<sup>2</sup>; generalization to those outside this group is unwarranted.

## 5. Conclusion

Measurement of spinal stiffness using AI demonstrated excellent intra-rater reliability. Due to the smaller and less cumbersome nature of AI compared to other mechanical instruments, AI may be viable technology for clinical use, however, further research is needed to quantify inter-rater reliability and to investigate the responsiveness of this instrument (sensitivity and specificity) to alterations in stiffness values.

## Acknowledgements

Funding for this project and for Tasha Stanton was provided by the Province of Alberta Graduate Scholarship, the Strathcona Physiotherapy Foundation and

NSERC. Support for Greg Kawchuk was supplied by the Canada Research Chairs program. We would like to express our sincere appreciation to Gian Jhangri and Dr. Trish Manns for their statistical assistance, Al Fleming and Sam Graziano for their technical support, and the members of the Common Spinal Disorder Lab at the University of Alberta for their feedback and support.

## References

- Astrand PO, Rodahl K, Dahl HA, Stromme SB. Textbook of work physiology: physiological bases of exercise. 4th ed. Champaign, IL: Human Kinetics; 2003.
- Beaumont A, McCrumb C, Lee M. Differences in the posteroanterior stiffness of the lumbar spine during tidal breathing and breath holding. In: Proceedings of the seventh biennial conference of the manipulative physiotherapists association of Australia, Sydney, New South Wales, Australia 1991, p. 244–51.
- Binkley JM, Stratford PW, Gill C. Intrarater reliability of lumbar accessory motion mobility testing. *Phys Ther* 1995;75:786–92.
- Caling B, Lee M. Effect of direction of applied mobilization force on the posteroanterior response in the lumbar spine. *J Manipulative Physiol Ther* 2001;24:71–8.
- Childs JD, Fritz JM, Flynn TW, Irrgang JJ, Johnson KK, Majkowski GR, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med* 2004;141:920–8.
- Edmondston SJ, Allison GT, Gregg CD, Purden SM, Svansson GR, Watson AE. Effect of position on the posteroanterior stiffness of the lumbar spine. *Man Ther* 1998;3:21–6.
- Fleiss JL. The design and analysis of clinical experiments. 1st ed. New York: Wiley; 1986.
- Grieves G. Mobilisation of the spine: notes on examination, assessment, and clinical method. 4th ed. Edinburgh; New York: Churchill Livingstone; 1984.
- Harlick JC, Milosavljevic S, Milburn PD. Palpation identification of spinous processes in the lumbar spine. *Man Ther* 2007;12:56–62.
- Hicks GE, Fritz JM, Delitto A, McGill SM. Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization program. *Arch Phys Med Rehabil* 2005;86:1753–62.
- Kawchuk G, Herzog W. A new technique of tissue stiffness (compliance) assessment: its reliability, accuracy and comparison with an existing method. *J Manipulative Physiol Ther* 1996;19:13–8.
- Kawchuk GN, Fauvel OR. Sources of variation in spinal indentation testing: indentation site relocation, intraabdominal pressure, subject movement, muscular response, and stiffness estimate. *J Manipulative Physiol Ther* 2001;24:84–91.
- Kawchuk G, Liddle T, Fauvel R. The accuracy of ultrasonic indentation: a comparison of three techniques. *J Manipulative Physiol Ther* 2006;29:126–33.
- Latimer J, Lee M, Adams R. The effects of high and low loading forces on measured values of lumbar stiffness. *J Manipulative Physiol Ther* 1998;21:157–63.
- Latimer J, Lee M, Adams R, Moran C. An investigation of the relationship between low back pain and lumbar posteroanterior stiffness. *J Manipulative Physiol Ther* 1996a;19:587–91.
- Latimer J, Lee M, Goodsell M, Maher C, Wilkinson B, Adams R. Instrumented measurement of spinal stiffness. *Man Ther* 1996b;1:204–9.
- Latimer J, Goodsell MM, Lee M, Maher CG, Wilkinson BN, Moran CC. Evaluation of a new device for measuring responses to posteroanterior forces in a patient population, Part I: Reliability testing. *Phys Ther* 1996c;76:158–65.

- Lee R, Evans J. Load–displacement–time characteristics of the spine under posteroanterior mobilisation. *Aust J Physiother* 1992;38: 115–23.
- Lee M, Liversidge K. Posteroanterior stiffness at three locations in the lumbar spine. *J Manipulative Physiol Ther* 1994;17:511–6.
- Lee M, Svensson NL. Measurement of stiffness during simulated spinal physiotherapy. *Clin Phys Physiol Meas* 1990;11:201–7.
- Maher CG, Adams R. Reliability of pain and stiffness assessments in clinical manual lumbar spine examination. *Phys Ther* 1994;74:801–9.
- Maher C, Adams R. A psychophysical evaluation of manual stiffness discrimination. *Aust J Physiother* 1995;41:161–7.
- Maher CG, Adams RD. Stiffness judgments are affected by visual occlusion. *J Manipulative Physiol Ther* 1996;19:250–6.
- Maher CG, Latimer J, Adams R. An investigation of the reliability and validity of posteroanterior spinal stiffness judgments using a reference-based protocol. *Phys Ther* 1998;78: 829–37.
- Maher CG, Latimer J, Holland MJ. Plinth padding confounds measures of posteroanterior stiffness. *Man Ther* 1999;14:145–50.
- Maitland GD. *Vertebral manipulation*. 5th ed. London: Butterworth–Heinemann; 1986.
- Maitland GD, Hengeveld E, Banks K, English K. *Maitland’s vertebral manipulation*. 6th ed. London: Butterworth–Heinemann; 2001.
- Matyas T, Bach TM. The reliability of selected techniques in clinical arthrometrics. *Aust J Physiother* 1985;31:175–99.
- Shrout PE, Fleiss JL. Intraclass correlation: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.
- Snodgrass SJ, Rivett DA, Robertson VJ. Manual forces applied during poster-to-anterior spinal mobilization: a review of the evidence. *J Manipulative Physiol Ther* 2006;29:316–29.
- Viner A, Lee M, Adams R. Posteroanterior stiffness in the lumbosacral spine: the correlation between adjacent vertebral levels. *Spine* 1997;22:2724–9 [discussion 2729–30].
- White AA, Panjabi MM. *Clinical biomechanics of the spine*. 2nd ed. Philadelphia: Lippincott; 1990.