Posteroanterior movements of the cervical spine: repeatability of force displacement curves

Abstract
The repeatability of instrumented assessments of posteroanterior (PA) movements has been reported previously for lumbar and thoracic spines, but only in relation to limited parameters of the movement. This study describes a device for measuring PA movements of the cervical spine and reports on repeatability of the entire force/displacement (FD) curves. Repeatability was assessed using coefficients of multiple determination (CMDs) and adjusted CMDs (where the mean offsets between the two curves are removed and the shape of the curve can be more directly assessed) for inter-rater intra-day (Inter-rater), intra-rater inter-day (Inter-day) and intra-rater intra-day (Intra-rater) repeated measurements. The mean CMD and mean adjusted CMD for intra-rater measurements (0.90 and 0.99 respectively) were significantly higher than for the other measurement intervals. Inter-rater and Inter-day mean CMDs were 0.76 and 0.73 and mean adjusted CMDs were 0.96 and 0.97. It is concluded that the maximum repeatability is achieved if the same operator reassesses the patient on the same day. It is hoped that the methodology described will form the basis for further research that will enable greater understanding of what characteristics of PA movements inform manual palpation and thereby enable improvement in both manual therapy treatment teaching efficiency of manual therapy skills.
Introduction

Musculoskeletal disorders of the spine are amongst the most common health problems in Australia (Bogduk, Bolton, Jull, & Bellamy, 2003). Spinal mobilisation is a common form of treatment for spinal musculoskeletal disorders and is based on a presumed relationship between symptoms and intervertebral mobility. Posteroanterior (PA) movements typically produced by the manual application of force to an individual vertebra either on or lateral to the midline are commonly used for assessment and treatment of spinal symptoms (Maitland, Hengeveld, Banks, & English, 2005, pp. 117-128).

During a PA movement, the clinician perceives the relationship between the force applied and the resulting displacement. The PA movement is intended to provide the clinician with information about source of the patient’s symptoms and is often represented by a force-displacement (FD) curve. Although PA movements were once thought to produce localised translational intervertebral movement, it is now clear from both in vivo (Caling & Lee, 2001; Kulig, Landel, & Powers, 2004; Lee & Evans, 1997; Lee, McGregor, Bull, & Wragg, 2005; McGregor, Wragg, Bull, & Gedroyc, 2005; McGregor, Wragg, & Gedroyc, 2001) and in vitro studies (Gal, Herzog, Kawchuk, P. Conway, & Zhang, 1997a; Gal, Herzog, Kawchuk, Conway, & Zhang, 1997b; Sran, Khan, Zhu, & Oxland, 2005) that in addition to segmental movement around a flexion/extension axis, other movement also occurs including regional spinal movement, soft tissue compression, and movement of muscle and connective tissue. There are two challenges to assessing the source of a patient’s symptoms by PA movements. Firstly, it is not known how the characteristics of PA movements are altered in the presence of symptoms. Secondly, the displacement produced by a PA movement is complex including deformation of soft tissues as well as the vertical movement of the vertebra. The vertical movement of the vertebra in turn not only involves the entire cervical spine, but is influenced by other factors including rocking of the head and compression of the padding of the plinth. The difficulty in interpreting the clinically relevant characteristics of PA movements is therefore in being able to extract a symptomatic structure’s as yet unknown influence from a PA movement’s already complex signal.

In spite of the apparent difficulties, clinicians are able to detect clinically useful information from manual assessments of PA movements. For example, the location of congenital fusion can be reliably detected by manual motion palpation (Humphreys, Delahaye, & Peterson, 2004). In a more clinically relevant study, patients who received treatment corresponding to findings on manual palpation had better outcomes than those receiving randomly allocated treatment (Fritz, Whitman, & Childs, 2005).

Several researchers have attempted to objectively assess PA movements using an indentor to apply a force over part of a vertebra and sensors to measure displacement and force (Edmondston, Allison, Gregg, Purden, Svanssson, et al., 1998; Kawchuk, Fauvel, & Dmowski, 2001; Latimer, Goodsel, et al., 1996; Lee & Svensson, 1990; Lee & Evans, 1992). Attempts at characterisation of PA movements from instrumented assessments relied primarily on single values of displacement or stiffness at specified force
levels. Lee and colleagues (Latimer et al., 1996; Lee & Svensson, 1990) considered that the FD curve could be considered as consisting of a ‘toe region’ where the slope was non-linear followed by a linear region of the curve over 20 N or 30 N. The slope of the linear portion of the FD curves, the overall displacement and the length of the toe region have all been used to characterise PA movements. Although the rationale for selection of these parameters is not clear, they have been widely used to describe PA movements (e.g. Allison et al., 2001; Chiradejnant, Maher, & Latimer, 2003; Kaigle, Ekstrom, Holm, Rostedt, & Hansson, 1998; Latimer et al., 1996; Lee & Evans, 1992; Shirley, Ellis, & Lee, 2002; Sran et al., 2005).

Differences in single measures of stiffness or displacement of PA movements have been found to be related to factors affecting the target intervertebral structures (Kawchuk et al., 2001; Latimer, Lee, Adams, & Moran, 1996; Sran et al., 2005) as well as factors affecting regional or extraspinal factors (Chansirinukor, Lee, & Latimer, 2003; Colloca & Keller, 2004; Edmondston et al., 1998; Kawchuk & Fauvel, 2001; Shirley, Hodges, Eriksson, & Gandevia, 2003). In spite of the apparent difficulties, clinicians are able to detect clinically useful information from manual assessment of PA movements. Furthermore instrumented assessments have not agreed with clinician’s interpretations from manual assessments (Latimer et al., 1996) and the single measures used to characterise PA movements do not appear to be the same as the parameters that inform manual motion palpation (Maher, Simmonds, & Adams, 1998). As a result, the characteristics of PA movements that clinicians consider during motion palpation remain elusive (Petty, Maher, Latimer, & Lee, 2002).

Findings are emerging suggesting how changes in intersegmental stiffness may be related to pathology and how such changes might impact on PA movements. Gay, Ilharreborde, Zhao, Zhao, & An (2006) found differences in segmental stiffness occurring with lumbar disc degeneration were more pronounced near the neutral position rather than near the end of range. We performed computer-based modelling of PA movements where alterations of the ‘neutral zone’ suggested that effects on PA movements would be complex with the greatest differences likely to occur in the early to middle portion of the PA movement (Tuttle, Laakso, & Barrett, 2006).

The single values of displacement or stiffness used previously may not be the clinically relevant characteristics of PA movements. Other, as yet unknown parameters of the FD relationship may be necessary to adequately characterise PA movements. The repeatability of measurements of the FD relationship of PA movements must therefore be established if measurements of PA movements are to be used to detect the clinically relevant characteristics of PA movements. In order to ensure appropriate experimental design for future studies, it is also important to know whether repeated measures can be reliably performed by the same or different practitioners and on the same or different days.

The purpose of this study was to determine the inter-rater, intra-rater and inter-day repeatability of the Posteroanterior Movement Assessment Device (PMAD) which was developed to assess PA movements.
of the cervical spine. It was hypothesised that the repeatability would be best for the same operator on the same day and that the variation from the tests being performed on different days would be greater than the variation resulting from different operators.

Methods

Subjects and experimental design

Subjects were recruited from university staff and students. Inclusion criteria were asymptomatic subjects defined as having no neck symptoms within the past six months that required treatment and no contraindications or precautions to manual therapy assessment or treatment (H. Lee, Nicholson, & Adams, 2004). Ten participants (six females and four males; mean age 37.2, range 21 to 50; mean weight 72.7 kg, range 52 to 92 kg; mean height 169.9, range 155 to 179 cm) were recruited for the study. The experimental protocol was approved by the Griffith University Human Research Ethics Committee and all individuals provided written confirmation of their informed consent prior to participation.

The procedures were explained and the subjects familiarised with the equipment prior to the first trial. Subjects were assessed by two trials on each of two consecutive days with all trials performed by the same operator except the second trial on day two which was performed by a second operator. Both operators were qualified musculoskeletal physiotherapists with over 10 years experience. Each trial consisted of PA movements to a total of six locations: both sides at each of three levels separated by 12 mm along the long axis of the treatment bed. The intention of this study was to assess the repeatability of the PMAD not to assess the ability to locate anatomical locations. It was therefore not necessary to undertake the imaging that would be required to locate positions anatomically. Rather the positions of the PMAD selected in the first trial were repeated for subsequent trials as it was necessary only to ensure that the indentor and patient were located in the same position for repeated measurements. After the first trial on each day, the subject stood and walked a few steps before adopting the same position for a second trial.

Instrumentation and data collection protocol

‘Unilateral’ PA movements were assessed with the force applied over the articular pillar. In order to maximise the relevance of our instrumented assessment to the clinical setting, our intention was to develop a methodology capable of assessment of PA movements of the cervical spine in a manner as similar as possible to that which occurs during manual palpation of PA movements. In principle, the device was similar to methods reported previously for the lumbar and thoracic regions, but was adapted in two ways. Firstly, the indentor was constructed to be similar in size to the human thumb and secondly, the force was applied manually rather than mechanically. The force being applied manually enables the device to be used in future studies to establish the physical factors corresponding to clinicians’ perceptions. The principal requirements of the PMAD (Figure 5.1) developed for this purpose were
identified as being able to:

1. Apply a PA force to repeatable locations.
2. Produce repeatable movements as similar as possible to manual assessment of PA movements.
3. Be perceived by the subject and operator to be as similar as possible to those occurring with manual assessment.
4. Accurately measure the force and displacement characteristics of the PA movements.

The indentor was intended to resemble the human thumb as used in manual assessment of PA movements to ensure maximum subject comfort without the indentor being larger than the contact used by a practitioner or wider than the height of a cervical vertebra. An indentor was constructed of a 25 mm length of 12 mm square aluminum section with edges rounded to a radius of approximately 1 mm.

Figure 1. Posteroanterior Movement Assessment Device (PMAD). The device measures force and displacement during movement produced by a force applied manually to the thumb-hold.

The assessment of each location consisted of five gradual applications of force up to 25 N performed at an intended frequency of approximately 1 Hz (actual frequency 0.40 Hz, SD 0.082) which is within the range used in manual assessment (Snodgrass, Rivett, & Robertson, 2006). The force and displacement data were simultaneously recorded while the operator gradually applied force to the thumb-hold until hearing an audible sound produced by the data acquisition program when the force reached 25 N. The
five applications of the force were angled medially by 10 degrees and positioned such that the shaft was 15 to 25 mm from the midline when the indentor contacted the subject. The shaft was secured to a linear bearing to ensure smooth repeatable movement and the entire assembly could be repositioned to a corresponding position on the contralateral side. Both the operator and subjects found this combination of indentor size, position and direction most closely approximated the sensations felt during manual assessment of unilateral PA movements.

The applied force was measured with a load cell (Transducer Technologies MLP-25) between the thumb-hold and the indentor. The corresponding displacements were measured with a linear potentiometer (Hollywell LTS04N04KB5C) attached to the shaft of the PMAD. The non-repeatability of the load cell is reported by the manufacturer to be 0.05% and the linearity of the potentiometer to be ± 0.1%. A two-point linear calibration was performed for both the potentiometer and load cell prior to each day’s testing. Three readings were taken at each of two displacements (separated by 69 mm) and two loads (deadweights of mass 27 and 2362 g). The sensors were connected to a PC through a USB DAQ card (USB-6008, National Instruments) and sampled at 100 Hz. Data collection and storage was performed with custom software written in Labview Version 7i.

A repeatable subject position was achieved by replacing the head section of a standard height-adjustable treatment bed with a specifically designed head-cradle lined with 2 mm of high density EVA foam. To determine how much movement was likely to occur as a result of compression of the head-cradle and treatment bed an anatomical skull model was placed in the head cradle to represent the subject’s head and a flat plate 20 cm × 20 cm weighing 5 kg was placed on the treatment table as a conservative representation of the subject’s chest. Forces of 12.5 N were applied to each and displacements measured using the PMAD. The measured displacement of the skull model of 0.2 mm and of the plate of 0.4 mm suggested that less than 0.4 mm of the PA movement on subjects would result from compression of the supporting head-cradle and treatment table. The reproducibility of positioning of the subject was tested by measuring the angle of a pointer held between the subject’s teeth and the position of the vertex. The angle and position were then re-measured after the subject stood up and repositioned themselves on the treatment bed. The ICC (3,1) were 0.94 and 0.99 for head angle and position respectively and the corresponding 95% limits of agreement (Bland & Altman, 1999) were −1.9 to 1.9 degrees and −1.7 to 1.7 mm, respectively. A sliding frame able to be fixed at 12 mm intervals along the long axis of the bed was fixed to the head cradle to enable repeatable positioning of the device for the three levels to be tested.

**Data analysis**

Following data collection, the data was processed with custom software using Matlab Version 7.04. Data from the second, third and fourth applications from 0.5 N to 25 N of force were filtered using a second order low-pass Butterworth filter with a cut-off frequency of 2.5 Hz. The displacement at 0.5 N was
assigned a value of zero to create a common starting point for further comparisons and the three resulting curves were averaged. In preliminary trials soft tissue occasionally appeared to move under the indentor and the force in the resulting measure did not increase continuously. Therefore, when processing the data, if the force data did not continuously increase through the range being assessed for any of the middle three applications, the fifth and, when necessary, the first applications were used in their place for further analysis. Although data from the first application were used on only three occasions, it is important to note that our preliminary testing did not suggest the need for ‘preconditioning’ the movement as we did not detect any differences between the first and subsequent force applications. Preconditioning may not have been required due to aligning all curves at a force level of 0.5 N rather than at a common location in space. Ninety-nine displacement values corresponding to forces from 0.5 N to 25 N at 0.25 N intervals were then determined using a cubic spline interpolation.

**Statistical analysis**

The coefficient of multiple correlation (CMC) was advocated by Kadaba, Ramakrishnan, Wootten, Gainey, Gorton, & Cochran (1989) and has since been used extensively to assess the repeatability of curve data related to gait. The adjusted coefficient of multiple determination (CMD) is defined as the square of the CMC and was used in the current study as it indicates the proportion of the variance accounted for within the data. CMDs have been used to assess the repeatability of measures of both gait (Kavanagh, Morrison, James, & Barrett, 2005) and active spinal movement (Lee, Laprade, & Fung, 2003), but to our knowledge they have not been used for the assessment of the repeatability of passive movements. CMD is defined as:

\[
\text{CMD} = 1 - \frac{\sigma^2_e}{\sigma^2_g}
\]

Equation 1

Where:

- \(\sigma^2_e\) represents the variance from the ensemble average curve, and
- \(\sigma^2_g\) represents the variance from the grand mean of the FD curves.

The more similar the curves being compared, the more \(\sigma^2_e\) approaches 0 and the CMD approaches 1. Conversely, the more dissimilar the curves, the more \(\sigma^2_e\) approaches \(\sigma^2_g\) and the CMD approaches 0.

The offset (systematic bias) in pairs of FD curves was assessed by calculating the difference (in mm) between the grand means of the two curves. In order to facilitate comparison of the overall shape of the curves, adjusted CMDs were obtained by recalculating the CMDs for each pair of FD curves with the
offset removed. In other words the repeatability of the original curves considers the curves with a common starting point, the offsets indicate any systematic bias as might occur with altered tissue compliance or muscular contraction while the adjusted curves compare the overall shape of the curves.

The repeatability of the FD curves was assessed by calculating the CMDs for each of the three repeated measures that were assessed: intra-day, inter-rater (Inter-rater); intra-day, intra-rater (Intra-rater); and inter-day, intra-rater (Inter-day). The CMDs were also calculated to compare the repeatability of the shape of the curves with the offsets removed. ANOVAs and post hoc Scheffe tests were used to assess main effects of repeated measure interval, side, and level on CMDs, offsets and adjusted CMDs. Statistical significance was set at $p < 0.05$.

**Results**

Representative data for both sides of one subject showing the raw data and averaged curves are shown in Figures 5A.1 and 5A.2 in the appendix. Representative data of two pairs of repeated measures with the corresponding CMDs, offsets and adjusted CMDs are shown in Figure 5.2. The graph illustrates how the overall displacement and the shape of the FD curves can differ and gives an indication of the extent of agreement between pairs of curves that corresponds with CMD, offset and adjusted CMD values.

![Graph showing displacement and force with CMD values](image)

**Figure 2.** Representative data from two pairs of repeated measures with corresponding CMDs. The similarity of the Intra-day curves is confirmed by the high CMD. The adjusted CMD indicates the extent of agreement between the curves once the offset is removed. The high adjusted CMDs of both repeated measures indicate close agreement in the shape of the curves.
The CMDs for each repeated measure interval ranged from 0.72 to 0.90 with the Intra-rater mean CMD being 0.90 (Figure 5.3a). There were significant differences between CMDs for the three repeated measure intervals ($F = 6.57$, $dF = 2$), with Intra-rater repeatability being greater than Inter-rater by 0.16 (CI 0.03 to 0.28) and greater than Inter-day by 0.18 (CI –0.04 to 0.31). There were no differences in CMDs between sides or levels.

The offsets for the repeated measure comparisons are shown in Figure 5.3b. The mean offsets for the three repeated measure intervals ranged from –0.1 mm to 2.0 mm. There were significant differences in the magnitude of the offsets ($F = 8.47$, $dF = 2$) with the Inter-day offsets being larger than Inter-rater by 0.16 mm (CI –0.34 to 2.77 mm) and larger than Intra-rater by 2.22 mm (CI –0.80 to 3.63 mm). There were no significant differences in the offsets between sides or levels.

As shown in Figure 5.3c, the adjusted CMDs for the three retest intervals were higher than the standard CMDs and ranged from 0.96 to 0.99. Again there were significant differences between the three intervals ($F = 2.46$, $dF = 2$) with the Intra-rater repeatability being greater than Inter-rater by 0.03 (CI 0.01 to 0.05) and Inter-day by 0.02 (CI –0.01 to 0.04). There was a small, but statistically significant difference in the adjusted CMDs between sides ($F = 4.11$, $dF = 1$) with the left side being greater than the right by 0.01 (CI 0.00 to 0.03), but no differences between levels.
Figure 3. Repeatability for each level and side for each repeatability interval. CMDs are shown in a). Offsets (the difference between the ensemble mean values of each FD curve) are shown in b) and the Adjusted CMDs (comparing the shape of the FD curves with the offsets removed) are shown in c).
Discussion

The purpose of this paper was to report on the repeatability of assessment of PA movements of the cervical spine using the PMAD. To our knowledge this is the first report of the repeatability of assessment of PA movements of the cervical spine and the first study to assess the repeatability of entire FD curves rather than single values for any region of the spine.

Repeatability of PA movements

As expected the best repeatability for all measures occurred for Intra-rater comparisons, where the repeated measures were taken by the same operator on the same day. The expected difference between Inter-rater and Inter-day repeatability was only found in relation to the adjusted CMDs and was most likely too small to be of practical significance. Therefore differences in PA movements would be expected to be easiest to detect if repeated measurements are taken by the same operator on the same day. The instrumentation can be considered to perform equally well regardless of side or level as indicated by the repeatability being comparable across sides and levels.

The adjusted CMDs were consistently larger than the non-adjusted CMDs demonstrating greater consistency in the shape of the curve than when the curves had a common starting point. If only the shape of the curve is being considered, the repeatability of the current instrumentation is excellent with CMDs over 0.96 for all intervals of repeated measures. The reason for or significance of the offsets being larger for intra-day than for the other comparisons is not clear.

Repeatability of FD curves from assessments of PA movements has not been reported previously. Previous studies on in vivo assessment of PA movements have been performed on lumbar, thoracic or animal spines. Lee and Evans (1992) assessed PA movements of the lumbar spine and reported ICCs for inter-day testing of 0.99 and 0.95 for displacement with a maximum error of less than 1 mm. Lee and Svensson (1990) also found good repeatability of the stiffness of the linear portion of the FD curve with ICCs of 0.88. A later device from the same group reported ICCs of 0.89 to 0.96, for intra-day, intra-rater repeated measures performed without repositioning the subject (Latimer et al., 1996). Kawchuk et al. (2001) added an ultrasound transducer to the tip of the indentor of a similar instrument to reduce possible errors in positioning as well as providing a more direct indication of the depth and movement of the target vertebra. The resulting accuracy for the linear region of the FD curve of porcine vertebrae was 1% to 2% for displacement and 6% for stiffness while for the non-linear region of the curve the corresponding accuracies were 3% to 4% and 14% respectively. It is not possible to make direct comparisons between the current study and previous studies as different measures of repeatability were employed in each study. It could be expected that assessment of the cervical spine may be less repeatable than other regions of the spine as the subject's position when assessing the cervical spine is less stable and the bony landmarks being assessed more difficult to locate accurately.
Relevance to clinical practice

The PMAD described in this paper was intended to assess PA mobility in a way comparable to that used in manual palpation. Even though PA movements involve the entire cervical spine, their intention is to assess symptomatic structures and it appears that clinicians are able to extract clinically useful information. It is not understood what parameters of PA movement inform the practitioner’s interpretations, but it is clear that single values of displacement or stiffness do not adequately characterize the movements being assessed.

At this time we can only speculate on ways symptoms or pathology might affect FD curves measured using the PMAD. For example, increased muscular contraction might produce a linear change in the entire curve, a difference in the size of an intervertebral 'neutral zone' might produce a greater difference in the early part of the PA movement, while alterations in viscosity as suggested by Nicholson, Maher, Adams, & Phan-Thien (2001) might result in rate-dependent differences that are uniformly distributed throughout the PA movement. The potential complexity of aspects of PA movements are further illustrated by Maher et al. (1998) who documented over 40 terms used by practitioners to describe their perceptions of mobility on motion palpation. Ramsay (1996) suggests that the significant features of continuous data are often more apparent when one considers derivatives of the raw data. For example the pattern of stiffness (the first derivative of the FD curve) or the change in stiffness (the second derivative of the FD curve) may be factors considered by practitioners. As the meaning of differences in offsets in the current study is unclear, it is interesting to note that the offsets are eliminated when derivatives of the data are considered. Changes in stiffness for example may be related to descriptors such as $R_1$ (the point where the first ‘resistance’ to the movement is felt) and end-feel (the change in stiffness at the end of the movement).

Some authors suggest that the difficulty in assessment of PA movements results from practitioners being informed by non-repeatable or idiosyncratic factors (Maher & Adams, 1995). Others have suggested using clinicians who are ‘gold standard palpators, against which others can be calibrated’ (Hansen, Simonsen, & Leboeuf-Yde, 2006). It is hoped that methodologies such as described in this paper will help reduce the subjectivity of motion palpation and enable more objective characterisation of the clinically relevant aspects of PA movements. A clearer understanding of the relevant characteristics of PA movements not only has the potential to improve the effectiveness of manual therapy diagnosis and treatment, but to improve the efficiency of teaching manual therapy skills to students.

It is expected that the repeatability of the PMAD found in the current study would be sufficient to enable studies using this methodology to detect differences relevant to clinical practice regardless of whether the differences are constant, rate dependent or unevenly distributed throughout the PA movements.
References


