THE RELIABILITY OF PRESSURE ALGOMETRY IN ASYMPTOMATIC INDIVIDUALS OVER CONSECUTIVE DAYS

Laura Frank*
Patrick McLaughlin[*,^]
Brett Vaughan[*,^]

*Osteopathy Unit
School of Biomedical & Health Sciences
Victoria University

^Institute of Sport, Exercise & Active Living
Victoria University

Corresponding Author

Brett Vaughan
Osteopathy Unit
School of Biomedical & Health Sciences
Victoria University
PO Box 14428
Melbourne VIC 8001

Phone 61 3 9919 1210
Fax 61 3 9919 1030
E-mail brett.vaughan@vu.edu.au

ABSTRACT

Background

To determine if electronic pressure algometry is a statistically stable measure of spinal pressure pain threshold (PPT) in asymptomatic individuals, in particular, to determine if repeated measurements at the same site changes the PPT, and to determine if repeatability differs in each of the spinal regions tested.

Design

Repeated measures design.

Setting

University teaching clinic.

Participants

Thirty-three asymptomatic participants.

Interventions

The PPT of three spinal segments (C6, T6 and L4) was measured three times in consecutive measures (10 seconds apart), then repeated one day and two days post-initial measurement. Measurements were taken using an electronic pressure algometer.

Main outcome measures

PPT, intra-class correlation coefficient and test of significant equality.

Results

Results demonstrated that the PPT measurement is statistically stable both between days ($p<0.001$) and within day ($p<0.001$). The intra-class correlation coefficient (ICC) values between the mean scores of daily trials demonstrated excellent concordance for each spinal segment (ICC = 0.860 - 0.953), with the exception of the correlation between day 1 and day 3 at T6, demonstrating good concordance (ICC = 0.676). All trial-to-trial correlations demonstrated excellent concordance both within trials of the same day (ICC = 0.833 – 0.988) and subsequent days (ICC = 0.823 – 0.940).

Conclusion

Electronic pressure algometry is a repeatable and statistically stable measure of the spinal PPT, both between days and within-day. The results provide evidence that the use of this device may be of value as an outcome measure for primary spinal complaints such as low back or thoracic spine pain.

Keywords

Pressure pain threshold; algometry; algometer; data correlation; reliability

THE REPEATABILITY OF PRESSURE ALGOMETRY IN ASYMPTOMATIC INDIVIDUALS OVER CONSECUTIVE DAYS

INTRODUCTION

Palpation provides relevant information about underlying tissue and the nociceptive system,[1] however, quantification of tenderness by palpation is subjective, and makes comparison over time periods and between patients difficult, in both clinical and research settings. The application of pressure by a device to a tissue has the advantage of triggering the same type of nociceptors as that of palpation.[2] A controlled application of pressure, as applied with an algometer, may therefore be a more appropriate technique for quantifying pain thresholds and pain tolerance levels of various musculoskeletal tissues. Algometry is also appropriate given the rate of application of pressure and the direction of the pressure being applied can be controlled.[3]

Pressure algometry has been described as a semi-objective method, or subjective measure,[4] for establishing the pressure-pain threshold (PPT) of various tissues. The PPT is defined as the minimum amount of pressure which induces pain or tenderness.[5, 6] The use of pressure algometry has been demonstrated in many studies to be a reliable and repeatable tool for quantifying local pain and tenderness in various tissues.[5, 7, 8] PPT values have been used in studying a variety of musculoskeletal conditions including fibromyalgia, headaches (such as cervicogenic and tension-type headaches), arthritis, spinal conditions, and Delayed Onset Muscle Soreness.[9,10]

Much of the previous literature investigating the effect of repeated PPT measurements on musculoskeletal tissues has focused on soft tissues, with most studies demonstrating reliability over repeated measures.[5, 6, 8, 11-13] Two studies[1, 9] have investigated the effect of repeated algometry in the spine however none have investigated the stability of PPT

measurements over consecutive days. Keating et al.[1] tested the repeatability of electronic algometry over a short period of time (30 minutes) on four spinal segments (C6, T4, T6, L4), using the mean of three trials on each segment. Sterling et al.[9] investigated repeatability over a longer period of time (one week between testing) on two segments of the cervical spine, using only one trial on each segment. Both of these studies demonstrated that the PPT was stable over both short and long durations.

Concurrent validity testing of the electronic pressure algometer used in this study was conducted by Vaughan et al.[14] Results obtained were similar for all rates of pressure application according to the standard correlation-based measure (ICC > 0.9). However, when tested for equality the two data sets were not significantly the same. These authors suggested that future studies investigate the validity and repeatability of the device, particularly with human tissues, as it was implied that repeated application of force may lead to an erroneous result.

The aim of the present study was to determine if electronic pressure algometry was a repeatable measure of spinal pressure-pain thresholds in asymptomatic individuals. In particular, the study investigated if repeated measurements on the same spinal site affected the stability of the PPT measures over consecutive days, and to determine if reliability differed in each of the three regions of the spine.

METHODS

Participants

Thirty-three (N=33) asymptomatic participants were recruited from the student population at Victoria University. Twenty female and thirteen male participants were recruited with a mean age of 19.36 ± 2.23 years. Participants were excluded if they had current complaints of cervical, thoracic, or lumbar spine pain, pain syndromes (e.g. fibromyalgia), were currently seeking manual therapy treatment or using analgesic/NSAID medication for a spine complaint. Participants were also excluded if they had a history of spinal pain in these areas that required ongoing manual therapy treatment or medication, a history of spinal trauma to the C6/T6/L4 levels, spinal congenital abnormality, spinal deformity, or the presence of a pathological condition such as osteoporosis, rheumatoid arthritis, osteoarthritis, or ankylosing spondylitis. Each participant provided written informed consent and completed a form to identify if any of the exclusion criteria applied for that participant. This study was approved by the Victoria University Human Research Ethics Committee.

Equipment

The PPT was measured using a hand held electronic pressure algometer (Somedic Algometer Type II, Sweden). This device consists of a gun shaped handle with a pressure-sensitive strain gauge at the tip and a 1cm diameter rubber plate fitted to the tip. Connected to the device is a hand-held button which, when activated by the

participant, automatically stops the pressure from increasing and subsequently displays a pressure reading in kilopascals (kPa) on the algometer's screen (Figure 1).

INSERT Figure 1 here

Training

Prior to testing, the examiner underwent four hours of training in order to become familiar with the operation of the device. Training consisted of at least twenty trials against a hard surface to practice the application of a constant pressure rate and practice trials were performed on three willing participants' spines during the training session. These participants were not experiencing pain in the cervical, thoracic or lumbar spine and did not participate in data collection phase of the study.

Procedure

Prior to commencement of measurement, a demonstration of pressure algometry was performed on the participant's ulnar styloid process in order to familiarise them with the procedure. Participants were then asked to lie prone on a treatment table with their back exposed. The spinous processes of C6, T6 and L4 were then located and marked with a skin pen by the examiner. The spinous processes of C6, T6 and L4 were selected for assessment as they have been used in a previous study,[1] and therefore discussion of the results of the present in relation to this previous study is possible. The PPT for each segment was measured three times in order to obtain an average PPT value, as studies

have shown that the mean of more than one trial is the most reliable estimate of the PPT,[11] in particular if the first trial is excluded,[6, 15] The algometer was positioned perpendicular to the body surface at the specified spinous processes and the pressure was increased at a rate of 40 kPa/sec.[9, 15] Previous research has indicated that a constant rate of pressure application increases reliability.[10, 16]

Participants were instructed to press a button held in their hand once the sensation on their spine changed from pressure to pain. Once the participant pressed the hand held button, the algometer measurements automatically stopped and the examiner removed the algometer from contact with the participant's spine and recorded the reading from the algometer. This process was performed three consecutive times at each spinal level with a period of 10 seconds between measurements. The examiner waited 20 seconds before commencing measurements on the next spinal level.

Participants were then required to return at the same time of day over the next two days where the same procedure was performed, thus providing PPT readings over the three consecutive days (Table 1). The participants were also requested to avoid receiving spinal manual therapy treatment over the three days of the study. The order of examination was determined by random; the examiner drew 3 cards from a container with that contained 3 cards marked with C6, T6 and L4. Once each participants measurements were completed, the cards were returned to the container.

INSERT Table 1 here

Data analysis

A two-way average measure Intraclass Correlation Coefficient (ICC 2, 1) statistic was used, which directly compares one value in a trial to the same value in another trial. The ICC technique is considered to be the most appropriate technique for assessing the repeatability and reliability of data. The ICC was used to compare trial-to-trial reliability and day-to-day reliability. Trial reliability was estimated by correlating trial 1 and 2 measures, trial 1 and 3 measures, and trial 2 and 3 measures each day, as well as the correlation among all three trials on each day. Day-to-day reliability was estimated by correlating the mean PPT score of one day with the mean PPT score of subsequent days, as well as computing the correlation among the mean scores of all three days. Correlations were also calculated between single trials of like numbers in the sequence of daily trials (e.g. trial 1 of the 1st day, with trial 1 of 2nd and 3rd day).

Interpretation of the ICC was based on criteria published by Fleiss:[17] ICC <0.4 represents poor concordance, 0.4-0.75 represents fair to good concordance, and >0.75 represents excellent concordance. The mean and standard deviation, ICC calculations and calculation of the 95% confidence interval for the ICC, was performed using SPSS 15.0 for Windows (SPSS, Chicago, USA).

Further analysis of the data was conducted using the test for significant equality,[18] a far more stringent test of "sameness" than the correlational techniques. This test was employed to determine whether the mean algometer values day-to-day were statistically the same.

The test of significant equality requires the selection of a minimum acceptable difference between the data sets (eg. 5% of the mean) and the calculation of the F-score, degrees of freedom (df) and a noncentrality parameter (f). Readers should refer to the original publication for further detail on the calculations involved in this technique.[18]

For the purposes of this analysis, the minimum acceptable difference (PD) between each day was set at 1% with calculations also made at 5%, 10% and 15% of the mean. A one-way ANOVA comparing the mean data sets for each day was used to calculate the F score and degrees of freedom in SPSS 15 for Windows (SPSS, Chicago, USA). The noncentrality parameter was calculated in MS Excel. The R Foundation for Statistical Computing (Version 2.1.1) was used to calculate the equality statistic using the non-central F test.

RESULTS

A total of 297 measurements for each spinal segment were collected providing a total of 891 PPT values over the 3 days of the study. Table 2 reports the PPT mean and standard deviation for each segment for each individual day of testing, as well as the mean PPT value of each segment obtained from all measurements over the three days. The standard deviation values of Table 2 demonstrate a wide variability of the mean PPT values for each segment.

INSERT Table 2 here

Table 3 reports the day-to-day repeatability of PPT measurements taken from the mean of multiple daily measurements performed on the specified spinal segments. The ICC values and confidence intervals of each correlation are presented. The data presented in Table 3 demonstrates excellent concordance for each segment of the spine (ICC 0.86 – 0.946) with the exception of the correlation between day 1 and day 3 at the T6 segment (ICC = 0.676), which demonstrates fair to good concordance.

INSERT Table 3 here

A scatterplot of the mean PPT scores of T6 for day 1 and day 3 is presented in Figure 2. Figure 3 demonstrates a number of outliers along the line of best fit, particularly within the mid section of the graph. Figure 3 represents a scatterplot of the mean PPT scores of

L4 for day 2 and day 3, and provides a comparison for Figure 2. The mean PPT values of Figure 4 are more closely distributed along the line of best fit.

INSERT Figure 2 here

INSERT Figure 3 here

Tables 4 and 5 report trial-to-trial repeatability of single trials performed on each segment correlated with single trials of the same day, or subsequent days respectively. The ICC values reported in Tables 4 and 5 indicate excellent concordance across the trials.

INSERT Table 4 here

INSERT Table 5 here

When the first trial of each segment on each day was excluded, and the mean of the other two trials correlated across days, all segmental recordings demonstrated excellent concordance. This is presented in Table 6 where the smallest ICC value was calculated between day 1 and day 3 for the T6 segment (ICC = 0.713 [0.420-0.858]).

INSERT Table 6 here

The test of equality statistic demonstrated that very few of the data sets were statistically the same ($p>0.05$) (Table 7). The exception was the T6 segment between days 2 & 3, which demonstrated that the data between these two days was statistically the same at a 1% difference between the means ($p=0.048$).

INSERT Table 7 here

DISCUSSION

The data obtained from this study suggests that electronic pressure algometry is a repeatable and statistically stable method for evaluating the PPT of the lumbar and cervical spine over consecutive days. Repeated PPT measurements of the specified segments of the spine were shown to be of a high concordance over the three consecutive days with the exception of the correlation between the mean trials of day 1 and day 3 for the T6 segment. The lower ICC value of this correlation (ICC = 0.676) is still within the fair to good concordance range. Trial-to-trial measures of all segments also demonstrated excellent concordance both between single trials of the same day (ICC = 0.833-0.988) and subsequent days (ICC = 0.823-0.940).

The mean of the individual PPT scores for the thoracic segment were higher than that of the cervical segment, though lower than that of the lumbar segment over all three days of testing. This finding is the same as that reported by Keating et al.[1] These authors tested four segments of the spine (C6, T4, T6 and L4) and found that PPT values increased in a caudal direction; values were lowest in the cervical region and highest in the lumbar region. Interestingly, these authors suggested that thoracic tenderness was therefore not a common finding among asymptomatic individuals.

There are few comparative studies that have assessed the effect of repeated algometry in the spine, and to the authors' knowledge, none have determined the effect over consecutive days. Sterling et al.[9] tested the reliability of repeated PPT measurements on two segments of the cervical spine, and concluded no significant change was found in

the PPT. A problem with this study was that only one trial was performed on each cervical segment, and then repeated a week later. Keating et al.[1] tested repeatability on four segments of the spine (C4, T4, T6 & L4) by obtaining the mean PPT of three trials performed on each segment (20 seconds apart), then repeated 30 minutes later. The ICC results showed that the reproducibility of PPTs in the cervical and thoracic spine was excellent (>0.9), and good in the lumbar spine (>0.75), therefore indicating that pressure algometry is reliable over a shorter period of time. The results of the present study concur with the results obtained by Keating et al.[1] and also indicate that PPT measures of the spine may be repeatable over a period of days.

Previous studies have found that reliable estimates of the PPT are greatest when the mean of more than one trial is used,[11] and when the first trial of each measurement is excluded.[6] The exclusion of the first trial when calculating the mean PPT is believed to eliminate the measurement that has the greatest variability in measurements when using heat, cold and chemical stimulation of a tissue.[17] However, the results of the present study refute previous findings that results are more reliable when the first trial is excluded. Overall, concordance between the mean PPT scores of each segment excluding the first trial were marginally lower than concordance containing all three trials, though still of excellent concordance (ICC = >0.847) with the exception of day 1 versus day 3 for T6 which maintained good concordance (ICC = 0.713).

The high concordances of the results of this study are also consistent with other literature that has determined the effect of repeated PPT measurements in other musculoskeletal tissues. Nussbaum & Downes[6] tested three trials (10 seconds apart) on

the biceps brachii muscle over three consecutive days and found no significant change in the PPT between trials of the same day or day-to-day. Potter et al.[12] tested two trials (5 seconds apart) on the belly of four spinal muscles on three occasions separated by one week. No significant change was found in the mean PPT of either within-session or between-session measurements for each muscle, therefore indicating that repeated algometry is also reliable among soft tissues.

A possible reason for the lower concordance between day 1 and 3 measurements of the T6 segment may be attributed to the fact that during measurement, a small number of participants reported tenderness over the segment, particularly on the third day of testing. These participants reported a feeling of bruising over the spinous process as the pressure was applied with the algometer. No visible bruising was seen in any of the participants over the three days and this sensation was not reported over the C6 & L4 segments. This sensation of bruising may indicate an increased sensitivity of the tissues overlying the T6 segment and therefore produce a lower PPT value. This finding is not associated with lower PPT values of T6 in comparison to that of the cervical and lumbar segments, nor is it associated with a decline in the mean PPT value of T6 over the three days of measurement. The mean PPT value of T6 for day 3 was lower than that of day 1, though it was higher than that of day 2. The correlation between the mean trials of day 1 and day 3 for this segment did, however, marginally improve when the first trial of each day was excluded (ICC = 0.713).

In addition, the thoracic segment was reported by the examiner to be the most difficult spinous process to measure the PPT. In some participants, particularly those with

prominent thoracic segments, the increased prominence made it more difficult to remain centred on the spinous process whilst increasing the pressure on the segment with the algometer. On a few occasions during measurement the algometer slipped off the spinous process which meant a repeat measurement was taken in order to obtain a more reliable value. The subsequent recording may have therefore had an effect on the reliability of the results for the thoracic segment and could be regarded as a limitation of the present study. This finding adds merit to a statement made by Nussbaum & Downes[6] that the PPT may be more reliable when the bony measurement site is flat and broad.

Trial-to-trial repeatability of PPT measurements demonstrated high concordance, both within trials of the same day and of subsequent days. For the purpose of this study, the examiner underwent a training session in the days prior to testing which consisted of multiple practice trials on several willing participants' spines. The high correlation values may be a reflection of this training, however the training procedure was minimal and may allow the device to be used as a single pre-test and post-test measure[15] and therefore appropriate for routine use in clinical practice. The use of the device could be as outcome measure for spinal complaints to quantify the subjective changes in pain reported by patient. The training procedure used in the present study could easily be replicated in clinical practice.

These data were also assessed for equality between days using a more stringent approach – the non-central F test.[18] Generally, the data were not equal across days and the explanation provided above has relevance to these findings also. It is also reasonable

to assume that individual variation between days, both within the patient and the examiner, makes it highly unlikely that data of this nature will ever be equal within a tolerance of 10%. Variability in PPT does not seem to fall within such a small criterion range when measured over consecutive days. The non-central F test seems more suited to assessing data collected simultaneously from different devices, rather than the data provided by participants over consecutive days.

CONCLUSION

Spinal PPT assessment by electronic algometry is repeatable, both between and within-sessions. This study has demonstrated that with a trained examiner, assessment of the PPT using electronic pressure algometry is reliable at different spinal segments over consecutive days. The method utilised in the present study may be useful for symptomatic individuals with primary spinal complaints however further investigation is required.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Jane Mulcahy for her review of the manuscript.

REFERENCES

1. Keating L, Lubke C, Powell V, Young T, Souvlis T, Jull G. Mid-thoracic tenderness: a comparison of pressure pain threshold between spinal regions, in asymptomatic subjects. *Man Ther* 2001;**6**(1):34-9.

2. Vatine J-J, Shapira S, Magora F, Adler D, Magora A. Electronic pressure algometry of deep pain in healthy volunteers. *Arch Phys Med Rehab* 1993;**74**:526-30.

3. Farella M, Michelotti A, Steenks M, Romeo R, Cimino R, Bosman F. The diagnostic value of pressure algometry in myofascial pain of the jaw muscles. *J Oral Rehab* 2000;**27**:9-14.

4. Ylinen J. Pressure algometry. *Aust J Physio* 2007;**53**:207.

5. Fisher A. Pressure algometry over normal muscles. Standard values, validity and reproducibility of pressure threshold. *Pain* 1987;**30**:115-26.

6. Nussbaum E, Downes L. Reliability of clinical pressure-pain algometry measurements obtained on consecutive days. *Phys Ther* 1998;**78**:160-9.

7. Reeves J, Jaeger B, Gradd-Radford S. Reliability of the pressure algometer as a measure of myofascial trigger point sensitivity. *Pain* 1986;**24**:313-21.

8. Vanderweeen L, Oostendorp R, Vaes P, Duquet W. Pressure algometry in manual therapy. *Man Ther* 1996;**1**(5):258-65.

9. Sterling M, Jull G, Carlsson Y, Crommert L. Are cervical physical outcome measures influenced by the presence of symptomatology? *Physio Res Inter* 1996;**7**:113-21.

10. Jensen K, Anderson H, Olesen J, Lindblom U. Pressure-pain threshold in human temporal region: evaluation of a new pressure algometer. *Pain* 1986;**25**:313-23.

11. Ohrbach R, Gale E. Pressure-pain thresholds in normal muscles: reliability, measurement effects, and topographic differences. *Pain* 1989;**37**:257-63.

12. Potter L, McCarthy C, Oldham J. Algometer reliability in measuring pain pressure threshold over normal spinal muscles to muscles to allow quantification of anti-nociceptive treatment effects. *Int J Ost Med* 2006;**9**:113-9.

13. van Wilgen P, van der Noord R, Zwerver J. Feasibility and reliability of pressure pain threshold measurements in patellar tendinopathy. *J Sci Med Sport* 2011 doi:10.1016/j.jsams.2011.05.004.

14. Vaughan B, McLaughlin P, Gosling C. Validity of a pressure algometer. *Int J Ost Med* 2007;**10**:24-8.

15. Persson A, Brogardh C, Sjolund B. Tender or not tender: retest repeatability of pressure pain thresholds in the trapezius and deltoid muscles of healthy women. *J Rehabil Med* 2004;**36**:17-27.

16. Kinser A, Sands W, Stone M. Relaibility and validity of a pressure algometer. *Journal of Strength and Conditioning Research* 2009;**23**(1):312-4.

17. Fleiss J. *The design and analysis of clinical experiments*. New York: John Wiley and Sons; 1986.

18. Londoree B, Speckman P, Clapp D. Testing for hypothesized equality. *Res Q Exerc Sport* 1990;**61**:275-9.

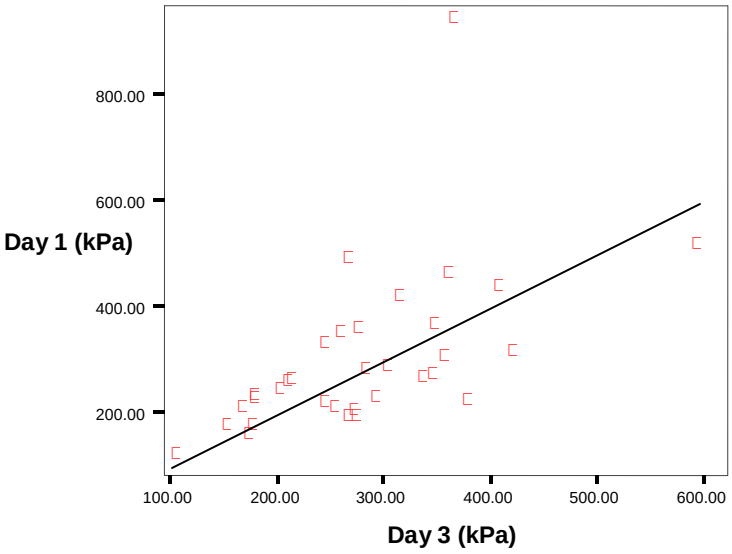

Figure 2. The mean PPT scores for the T6 segment with day 1 means plotted against the mean scores of day 3.

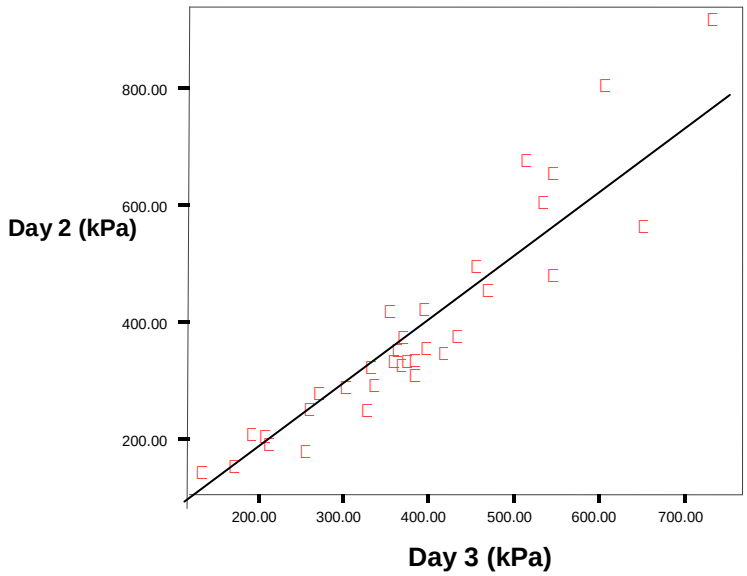

Figure 3. The mean PPT scores for the L4 segment with day 2 means plotted against the mean scores for day 3.

C6, T6 & L4		
Day 1	**Day 2**	**Day 3**
Trial 1 Trial 2 Trial 3	Trial 1 Trial 2 Trial 3	Trial 1 Trial 2 Trial 3
= Mean PPT score for Day 1	= Mean PPT score for Day 2	= Mean PPT score for Day 3

Table 1. Interpretation of the method

	Mean PPT (SD)		
	C6 segment	**T6 segment**	**L4 segment**
Day 1	271.5 (148.0)	292.8 (153.0)	389.8 (165.5)
Day 2	250.5 (112.6)	286.7 (133.0)	376.7 (186.0)
Day 3	267.7 (107.6)	288.6 (100.9)	394.4 (143.3)
Day 1, 2 & 3	263.2 (124.0)	289.4 (130.3)	387.0 (165.5)

Table 2. Mean PPT (kPa) of each spinal segment for each individual day and for all three days combined.

	ICC value [95% Confidence Interval]		
	C6 segment	**T6 segment**	**L4 segment**
Day 1, 2 & 3	0.918 [0.854 – 0.957]*	0.872 [0.772 - 0.933]*	0.953 [0.916 – 0.975]*
Day 1 & 2	0.862 [0.720 – 0.932]*	0.882 [0.760 - 0.942]*	0.923 [0.844 – 0.962]*
Day 1 & 3	0.860 [0.716 – 0.931]*	0.676 [0.345 - 0.840]#	0.922 [0.842 – 0.962]*
Day 2 & 3	0.940 [0.878 – 0.970]*	0.884 [0.766 - 0.943]*	0.946 [0.890 – 0.973]*

This table shows the ICC and 95% confidence interval of the mean PPT score of one day correlated with the mean PPT score of subsequent days (* $p<0.0001$, # $p<0.001$).

Table 3. Day-to-day reliability of mean PPT measurements of the spine.

	ICC [95% confidence Interval]		
	C6 segment		
	Day 1	Day 2	Day 3
Trial 1x2x3	0.908 [0.964-0.989] *	0.949 [0.908-0.973] *	0.954 [0.908-0.973] *
Trial 1x2	0.975 [0.949-0.988] *	0.893 [0.783-0.947] *	0.926 [0.851-0.964] *
Trial 1x3	0.954 [0.908-0.977] *	0.938 [0.874-0.969] *	0.928 [0.853-0.962] *
Trial 2x3	0.979 [0.958-0.990] *	0.947 [0.893-0.974] *	0.944 [0.887-0.972] *
	T6 segment		
	Day 1	Day 2	Day 3
Trial 1x2x3	0.988 [0.973-0.993] *	0.902 [0.826-0.949] *	0.945 [0.902-0.971] *
Trial 1x2	0.977 [0.953-0.988] *	0.888 [0.774-0.945] *	0.913 [0.824-0.957] *
Trial 1x3	0.987 [0.973-0.993] *	0.833 [0.662-0.918] *	0.877 [0.751-0.939] *
Trial 2x3	0.980 [0.960-0.990] *	0.865 [0.726-0.933] *	0.961 [0.921-0.981] *
	L4 segment		
	Day 1	Day 2	Day 3
Trial 1x2x3	0.964 [0.936-0.981] *	0.977 [0.959-0.988] *	0.963 [0.934-0.981] *
Trial 1x2	0.933 [0.865-0.967] *	0.964 [0.927-0.982] *	0.959 [0.917-0.980] *
Trial 1x3	0.948 [0.894-0.974] *	0.953 [0.904-0.977] *	0.925 [0.847-0.963] *
Trial 2x3	0.962 [0.922-0.981] *	0.979 [0.958-0.990] *	0.952 [0.904-0.976] *

This table shows the ICC and 95% confidence interval of single trials of the PPT correlated with single trials of the same day (*p<0.000).

Table 4. Trial-to trial reliability of individual PPT measurements of the spine.

	ICC value [95% Confidence Interval]		
	C6 segment	**T6 segment**	**L4 segment**
Trial 1x1x1	0.904 [0.829 – 0.949]*	0.823 [0.685 – 0.907]*	0.940 [0.893 – 0.968]*
Trial 2x2x2	0.858 [0.746 – 0.925]*	0.872 [0.772 – 0.933]*	0.940 [0.893 – 0.968]*
Trial 3x3x3	0.918 [0.855 – 0.957]*	0.831 [0.699 – 0.911]*	0.930 [0.874 – 0.963]*

This table shows the ICC and 95% confidence interval of single trials of the PPT correlated with single like-numbered trials of subsequent days (* $p<0.0001$).

Table 5. Day-to-day reliability of individual PPT measurements of the spine.

	ICC value [95% Confidence Interval]		
	C6 segment	**T6 segment**	**L4 segment**
Day 1, 2 & 3	0.903 [0.826 – 0.949]*	0.873 [0.773 - 0.933]*	0.946 [0.904 – 0.972]*
Day 1 & 2	0.862 [0.663 – 0.918]*	0.856 [0.709 - 0.929]*	0.918 [0.835 – 0.960]*
Day 1 & 3	0.862 [0.720 – 0.932]*	0.713 [0.420 - 0.858]*	0.902 [0.801 – 0.951]*
Day 2 & 3	0.847 [0.791 – 0.949]*	0.881 [0.759 - 0.941]*	0.940 [0.879 – 0.971]*

This table shows the ICC and 95% confidence interval of the mean PPT score of one day (excluding the first trial of each segment) correlated with the mean PPT score of subsequent days (*p<0.000)).

Table 6. Day-to-day reliability of mean PPT measurements of the spine (excluding the first trial of each segment).

Minimum acceptable difference	p-value								
	Day1/2	**Day1/3**	**Day2/3**	**Day1/2**	**Day1/3**	**Day2/3**	**Day1/2**	**Day1/3**	**Day2/3**
1%	0.479	0.094	0.473	0.225	0.579	**0.048**	0.235	0.093	0.329
5%	0.433	0.083	0.387	0.2	0.522	**0.041**	0.208	0.08	0.289
10%	0.381	0.071	0.472	0.172	0.459	**0.034**	0.178	0.066	0.246
15%	0.336	0.061	0.474	0.148	0.403	**0.028**	0.126	0.055	0.21

Table 7. p-values for the test of significant equality. Day 2/3 data was statistically significant ($p<0.05$).

AUTHOR CONTRIBUTION STATEMENT

PMcL and BV developed the idea for the study and developed the method. LF undertook the data collection. LF, PMcL and BV undertook the data analysis and developed the manuscript. All authors approved the final manuscript.

STATEMENT OF COMPETING INTERESTS

Brett Vaughan is an Editor of the Int J Osteopath Med but was not involved in review or editorial decisions regarding this manuscript.