

Intrarater Reliability of Pain Intensity, Tissue Blood Flow, Thermal Pain Threshold, Pressure Pain Threshold and Lumbo-Pelvic Stability Tests in Subjects with Low Back Pain

Aatit Paungmali*¹, MPhty, PhD; Patraporn Silitertpisan², MSc, PhD; Khanittha Taneyhill¹, MSc, PhD; Ubon Pirunsan¹, MPhty, PhD; Sureeporn Uthaihp¹, MPhty, PhD

Authors' Affiliation:

1. Neuro-Musculoskeletal and Pain Research Unit, Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand
2. Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand

* Corresponding Authors;

Address: Neuro-Musculoskeletal and Pain Research Unit, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand 50200

E-mail: aatit@chiangmai.ac.th

Received: May 12, 2011

Accepted: Jun 12, 2011

Key Words: Reliability; Low back pain; Outcome measures; Lumbo-pelvic stability; Pain

Abstract

Purpose: This preliminary study aimed to determine the intrarater reliability of the quantitative tests for the study of non-specific low back pain.

Methods: Test-retest reliability of the measurements of ratio data was determined by an intraclass correlation coefficient (ICC), standard error of measurements (SEMs), coefficient of variation (CV), and one-way repeated measures ANOVA using the values collected from 13 young individuals (25.8 ± 6.2 years) with chronic non-specific low back pain on two occasions separated by 2 days. Percent agreement of the ordinal data was also determined by Cohen's Kappa statistics (kappa). The measures consisted of tissue blood flow (BF), average pain visual analog scales (VAS), pressure pain threshold (PPT), cold pain threshold (CPT), heat pain threshold (HPT) and lumbo-pelvic stability test (LPST). An acceptable reliability was determined as the ICC values of greater than 0.85, SEMs less than 5%, CV less than 15%, the kappa scores of greater than 80% and no evidence of systematic error (ANOVA, P > 0.05).

Results: ICC of all measures in the lumbo-sacral area were greater than 0.87. The kappa was also greater than 83%. Most measures demonstrated a minimal error of measurements and less potential of systemic error in nature. Only the SEMs and the CV of the CPT exceeded the acceptable level.

Conclusions: It is concluded that most of the quantitative measurements are reliable for the study of non-specific low back pain, however the CPT should be applied with care as it has a great variation among individuals and potential of measurement error.

Asian Journal of Sports Medicine, Volume 3 (Number 1), March 2012, Pages: 8-14

INTRODUCTION

Chronic low back pain is an increasing health problem among young athletes [1-5]. For professional athletes, such as weightlifters, gymnasts, golfers, rowers, wrestlers and tennis players; low back pain is one of the most common reasons for missed playing time and loss in competition [1-4].

Inclusions of reliable and quantifiable measurement tools are necessary in both clinical and research settings as part of a path to success in diagnosis and

management of low back pain among athletes. Along with the pain scales, mechanical pain (i.e. pressure pain threshold) and thermal pain (i.e. cold and heat pain threshold) have been used to evaluate the severity and characteristics of hyperalgesia in various musculoskeletal conditions such as tennis elbow, ankle sprain, neck and shoulder pain [6-9]. Tissue blood flow is one of the factors indicating quality of healthy tissue and its potential for healing [10]. Most of the recent clinical studies also include the tissue blood flow as one of the primary measures for evaluating the physiological

effects of therapeutic treatments^[11-13]. In addition, core stabilization has been extensively mentioned in back pain literature as it is related to the severity of low back pain and function^[14-15]. Most athletes pay attention to gaining their core stability for minimizing back pain and injury, as well as promote their physical performance^[16-17]. These additional evaluating tools are potentially valuable in management of low back pain among athletes.

In order to include these measures in a study, it is necessary to establish the reliability of the measurements. At present, little information is available for the reliability of the quantitative outcome measures for the study of chronic low back pain. Therefore, the purpose of this study was to investigate the test-retest reliability of pain intensity, tissue blood flow, thermal pain threshold, pressure pain threshold and lumbo-pelvic stability tests that could be used to evaluate pathology and assess effects of treatment interventions for low back pain.

METHODS AND SUBJECTS

Design:

Test-retest intratester reliability was determined with a 48 hours interval between two occasions. This pattern of reliability study was utilized to replicate the study protocol of within-subject model for the study of low back pain.

Participants:

Thirteen young male and female (25.8 ± 6.2 years; 4 male, 9 female) with chronic non-specific low back pain volunteered to participate in this study. This amount of sample size was sufficient to establish the significant alpha level of 0.05 and power analysis of 0.80. They were recruited from the community and university areas during October 2010 to March 2011. The inclusion criteria were being 20-35 years old with mild to moderate back pain (VAS 2-7/10) of greater than 3 months in the area between the 12th rib to gluteal

folks. Their average (\pm standard deviation [SD]) height, body mass, pain intensity, and duration of onset were 165.2 ± 7.0 cm, 60.5 ± 10.2 kg, 3.9 ± 0.9 VAS, and 14.4 ± 13.2 mo, respectively. The subjects had no referred pain or neurological involvement in lower limbs, had no experience of surgery, and had no history of injury in the last 3 months before attending this study. The subjects were also requested not to take stimulants, medications, alcohol or participate in heavy physical activities at least 8 hours prior to the test. The study was approved by the institutional ethics committee and a written consent was obtained from each individual.

Procedure:

The measurements were taken over the most sensitive local spot over L1-S5, and the remote areas over the deltoid insertion and a proximal part of the tibialis anterior (5 cm distal to the Girby's tubercle) on both dominant and non-dominant sides. The measures consisted of tissue blood flow (BF), average pain intensity over the 10 centimeter pain visual analog scales (VAS), and pain thresholds including thermal pain threshold [cold pain threshold (CPT) and heat pain threshold (HPT)], and pressure pain threshold (PPT). The lumbo-pelvic stability test (LPST) was also included as the primary outcome measure for the study of low back pain. The order of measurements was standardized as follows; VAS, BF, CPT, HPT, LPST, and PPT, to consider possible carry-over effects from other measures. The interval between different measures was at least 5 minutes, and the rest period between trials in the same measure was 30 to 60 s as indicated in each test protocol shown below. One day prior to the study, all participants underwent a complete series of familiarization trials. The reliability assessments were based on the measures between two occasions at the same time of the day with a 48-hour interval. The same investigator performed all measurements and was blinded from the previous scores. All tests were conducted in a controlled environment laboratory room (24.5 ± 0.5 degrees Celsius [$^{\circ}$ C]).

Calibration and resolution: All of the instruments were calibrated before the measures according to the respective recommended procedures.

Outcome measures:

Pain intensity: The visual analogue scale (VAS) was used to rate the average intensity of pain over the lumbo-sacral area. The VAS consisted of a 10 cm line anchored with “no pain” on the left end and “extreme pain” on the right end. Subjects were asked to rate their perceived level of pain at rest.

Tissue blood flow: Blood flow of the tissue in unit of flux/min was monitored using a laser Doppler blood flow meter (Moor instruments DRT4, UK). An electrode of the laser Doppler blood flow meter was recommended to put over a center of the target area being investigated^[11-13]. In this study, the tissue at the most tender spot over the lumbo-sacral area (L1-S5) of each individual subject was evaluated. Each subject lay in prone position with arm by side, and the electrode was applied on the marked area. The tissue blood flow was recorded every minute for a period of 5 minutes. The mean value of tissue blood flow was used for further analysis.

Thermal pain threshold: Temperature or thermal pain threshold was the level of temperature that induces initial pain, and was assessed using a Thermal Sensory Analyzer (Medoc Ltd., Neuro Sensory Analyzer Model TSA-II, Israel) for cold pain threshold (CPT) and heat pain threshold (HPT). Each subject lay down on the bed (i.e. prone or supine) with arm by the side, and the thermode (5 cm²) was applied on the marked areas (i.e. lumbar, deltoid insertion, tibialis anterior) with a Velcro strap. The initial temperature of the thermode was set at 32 °C, and then it was modulated at a controlled rate (2 °C·s⁻¹ for cold pain and 1 °C·s⁻¹ for heat pain). The subject held a control switch, and was instructed to press the button when they felt the sensation changing from cold or heat to pain. The pain threshold in the unit of °C was assessed three times with a 30-s interval between trials. The mean value of the 3 trials was used for further analysis.

Pressure pain threshold: Pressure pain threshold (PPT) was measured by a pressure algometer (Somedic Production, Algometer type II, Sweden) with a probe of 1.0 cm². It was recalibrated in the laboratory with a 100-kPa calibrating weight before experimentation. The PPT was assessed in a similar manner as the thermal pain threshold. The pressure was increased at a

rate of 40 kPa·s⁻¹ until the subject felt the sensation changing from the pressure to pain, which was indicated by the subject pressing a button. PPT in the unit of kilo Paskal (kPa) was assessed 3 times for each site with 30-s rest between trials, and the mean of the 3 trials was used for further analysis.

Lumbo-pelvic stability test: There were 7 levels of the lumbopelvic stability control as recommended by Hagins and colleagues^[18]. Lumbo-pelvic stability test (LPST) was tested in supine position with knee flexion of 70 degrees. The pressure biofeedback unit (PBU) was placed under the lumbar spine (L2-L4) to monitor the stability of lumbo-pelvic position and the pressure transducer was pumped to 40 mmHg. The subjects were maintaining the stability of trunk in each level. Subjects received a pass category for each tested stability level, if the pressure gauge reading was within 40±4 millimeters of mercury (mmHg). In contrast, if the pressure gauge reading was out-off the target range, the subject received a fail category^[18,19].

Reliability analysis:

The test-retest reliability was primarily determined by intraclass correlation coefficients [ICC (3,1) for VAS], [ICC(3,5) for BF], [ICC(3,3) for CPT, HPT, PPT], and percent agreement by Cohen’s kappa for LPST. Coefficient of variation (CV) and standard error of measurements (SEMs) were included for determining variability of measurements. The presence of systematic bias between trials was also analyzed using one-way repeated measures ANOVA. The statistical significance was set at the alpha level of 0.05. The results of ICCs and one-way repeated measures ANOVA were obtained from the SPSS statistical package. In addition, CV and SEMs values were calculated from the following formula:

$$CV = (SD/\bar{X})100$$

$$SEMs = SD\sqrt{1-ICC}$$

where \bar{X} is the mean of the data, SD is the standard deviation of observed test scores, and ICC is the reliability coefficient for that measurement.

For SEMs interpretation, the percent value of the actual SEMs was determined by the proportion in percentage of SEMs value to the mean of data^[20,21].

RESULTS

Table 1 and 2 show the intraclass correlation coefficients (ICC), coefficient of variation (CV), standard error of measurements (SEMs) and analysis of systematic error (ANOVA) for all measures. The VAS, BF and PPT were considered to be reliable (i.e. ICC > 0.85, CV < 10%, SEMs < 3.5%) and no potential of systematic errors.

In addition, the LPST also showed an acceptable percent agreement of the test-retest (kappa = 83.1%). Thermal pain threshold, especially the CPT, was greater in CV (>20%) and potentially larger in measurement errors (>4.4%) when compared to that of the other measurements for both local (Table 1) and remote (Table 2) sites.

DISCUSSION

In this study, the test-retest reliability of the valuably quantitative measures (i.e. VAS, BF, CPT, HPT, PPT, and LPST) was determined for the study of low back pain using a series of ICC, percent agreement, CV, SEMs, and one-way repeated measures ANOVA. These measurement outcomes were evaluated in an attempt to utilize these measures for exploring characteristics and examining effectiveness of an intervention in both clinical and research settings among athletes or individuals with low back pain. The results of this current study showed that the data from

the local site of symptoms (i.e. back pain area) was more precise and relatively consistent than the non-symptomatic remote sites (i.e. deltoid, tibialis anterior). These findings were supported by the previous studies in which the local primary area of injury was sensitive to both mechanical (e.g., PPT) and thermal stimuli (i.e. CPT, HPT), but the remote site was mainly sensitive to mechanical stimulus [6,22]. A previous study of the test-retest reliability for pressure pain threshold measurements of the upper limb and scapular region also reported a high reliability of test-retest with a 2-day interval between sessions (i.e. ICC ranged from 0.90-0.98) [23]. This range of ICC value was similar to our current study (i.e. ICC at deltoid site ranged from 0.92-0.95) which is considered as an acceptable level of reliability. Interestingly, the reliability of the PPT at the remote site on tibialis anterior (TA) seemed to show a relatively lesser ICC value (i.e. ICC of 0.88-0.91) and larger variability than the remote site on the deltoid (i.e. ICC of 0.92-0.95) as well as the local area of lumbar region (i.e. ICC of 0.99). Some explanations were that the symptomatic lumbar area might be more sensitive to mechanical stimulus than the asymptomatic deltoid and leg areas. In addition, the leg might to some degree relate to an impairment of lumbo-sacral nerve roots which are distributed along the dermatome of lower limb. Another interesting point from this present study was that the pain intensity as described by the VAS scale remained relatively stable within a 2-day interval of an evaluation for the individuals with chronic low back pain. This finding was supported by the recent study which found that pain and symptoms were relatively unchanged with the chronic conditions [6,24].

Table 1: The test-retest reliability analysis of pain visual analog scale, tissue blood flow, lumbo-pelvic stability test as well as cold pain threshold, heat pain threshold and pressure pain threshold at the local lumbar area

Measurements	ICC	%CV	SEMs	P. values*
Pain visual analog scale	0.90	7.29%	0.09 (2.4%)	0.19
Tissue blood flow	0.89	9.43%	0.27 (3.2%)	0.07
Cold pain threshold	0.89	40.74%	0.13 (13.7%)	0.84
Heat pain threshold	0.87	1.33%	0.22 (0.5%)	0.43
Pressure pain threshold	0.99	3.31%	1.19 (0.3%)	0.56
Lumbo-pelvic stability test	Kappa = 83.1%	2.32%	0.02 (1.0%)	0.34

* P. value of one-way analysis of variance

ICC: Intraclass correlation coefficients; CV: Coefficient of variation; SEMs: Standard error of measurements

Table 2: The test-retest reliability analysis for cold pain threshold, heat pain threshold, and pressure pain threshold at the remote sites (i.e. Deltoid and Tibialis Anterior)

Measurements	ICC	%CV	SEMs	P. values*
Cold pain threshold [D]	0.87-0.96	21.95-41.86%	0.08-0.26 (4.5-15.2%)	0.38-0.56
Heat pain threshold [D]	0.68-0.82	1.67-2.03%	0.33-0.53 (0.7-1.2%)	0.06-0.10
Pressure pain threshold [D]	0.92-0.95	7.35-7.85%	4.29-5.33 (1.7-2.0%)	0.59-0.64
Cold pain threshold [TA]	0.76-0.85	37.63-45.17%	0.28-0.44 (14.5-22.0%)	0.07-0.69
Heat pain threshold [TA]	0.62-0.78	2.09-2.90%	0.46-0.82 (1.0-1.8%)	0.11-0.64
Pressure pain threshold [TA]	0.88-0.91	9.71-9.88%	10.80-11.43 (2.9-3.4%)	0.14-0.34

* P. value of one-way analysis of variance

ICC: Intraclass correlation coefficients; CV: Coefficient of variation; SEMs: Standard error of measurements

D: Deltoid; TA: Tibialis Anterior

The tissue blood flow (BF) showed that its reliability was suitable, this result was similar to the laser Doppler flowmetry study by Roeykens et al [25] which found that blood flow of the pulpal tissue was reliable (i.e. agreement of blood flux ranged from 0.85-0.88) with an interval of 1 week. However, a minimal diurnal variation of tissue blood flow might also be evident as it was influenced by the sympathetic tone and psychological stage [26]. For lumbo-pelvic stability test, its intratester reliability in this study was considerably acceptable with percent agreement (kappa's score) of 83.1%. Harris and Lahey [27] suggested that agreement scores of greater than 80% were adequate and considered as a conventional level for most scientific studies. However, the percent agreement of the lumbo-pelvic stability test in our current study seemed to be less than the study previously reported by Phrompaet et al [19], which reported kappa's score of 95%. One factor which might contribute to the different result was that the subjects in Phrompaet's study were healthy volunteers, therefore they might perform lumbo-pelvic stability test better than the individuals with low back pain in our current study.

It should be mentioned that the CV of CPT was relatively large (22% - 45 %) for all testing sites. Park and colleagues [28] studied the reliability of the sensory testing on the volar aspect of forearm in 19 healthy subjects. They found that the variability of the CPT was approximately 25.5% and the 95% confidence interval (95% CI) for the CPT was 18-24 times higher than hot pain threshold. Large variability of the CPT was also evidenced in subjects with spinal cord injury and neuropathic pain [29]. Khamwong et al [30] also

found the similar result of greater in CV (i.e. 27 %) with the CPT measurement. Although, many studies suggested that CPT is more sensitive in detecting changes than the HPT [31], the CPT should be applied with caution because it had a large variation among individuals. This might be due the fact that cold sensation is activated in a wide range (e.g., from less than 15 °C) and signals are transmitted via the complicated pathways including C- and A-delta myelinated nerve fibers [32].

It should be considered that this preliminary study might have some limitations such that the subject of this study was a non-specific low back pain. Further studies are warranted to evaluate the reliability of outcome measurements in a specific pathologic condition, as well as in a specific group of athletes with low back pain. In addition, further studies should include the other quantitative measures for the study of low back pain such as the measurements of transabdominal muscles using the modern techniques of real-time ultrasonic imaging or magnetic resonance imaging (MRI).

CONCLUSION

In conclusion, the present study assessed the test-retest reliability of various quantitative measures that could be utilized in a study to investigate characteristics and evaluate effects of intervention for the study of low back pain. It suggests that most of the measures are

reliable for the study of non-specific low back pain, however the CPT should be applied with care as it has a great variation among individuals and potential of measurement error. Therefore, a robust measurement procedure and a familiarization protocol should be considered for obtaining an acceptable reliability, minimizing measurement and systematic errors, and eliminating the learning effect.

ACKNOWLEDGMENTS

This study was approved by the institutional human ethics committee. We would like to express our gratitude to all participants in the study and the Thailand Research Fund (TRF) and the Higher Education Commission for supporting the grant for this study. Special thanks also go to Prof. Dr. Bill Vicenzino and Prof. Dr. Watchara Kasinrerak for their valuable guidance.

Conflict of interests: None

REFERENCES

1. O'Dowd J. Back pain in athletes. *Curr Orthopaed* 2001;15:110-26.
2. Jackson DM, Verscheure SK. Back pain in whitewater rafting guides. *Wilderness Environ Med* 2006;17:162-70.
3. Evans K, Refshauge KM, Adams R, Aliprandi L. Predictors of low back pain in young elite golfers: A preliminary study. *Phys Ther Sport* 2005;6:122-30.
4. Raske A, Norlin R. Injury incidence and prevalence among elite weight and power lifters. *Am J Sports Med* 2002;30:248-55.
5. Shehab DK, Al-Jarallah KF. Non specific low-back pain in Kuwaiti children and adolescents: associated factors. *J Adolescent Health* 2005;36:32-5.
6. Wright A, Thurnwald P, Smith J. An evaluation of mechanical and thermal hyperalgesia in patients with lateral epicondylalgia. *Pain Clinic* 1992;5:221-7.
7. Collins N, Teys P, Vicenzino B. The initial effects of a Mulligan's mobilization with movement technique on dorsiflexion and pain in subacute ankle sprains. *Man Ther* 2004;9:77-82.
8. Teys P, Bisset L, Vicenzino B. The initial effects of a Mulligan's mobilization with movement technique on range of movement and pressure pain threshold in pain-limited shoulders. *Man Ther* 2008;13:37-42.
9. Sterling M, Pedler A, Chan C, et al. Cervical lateral glide increases nociceptive flexion reflex threshold but not pressure or thermal pain thresholds in chronic whiplash associated disorders: A pilot randomized controlled trial. *Man Ther* 2010;15:149-53.
10. Mani R, Cooper C, Kidd B, et al. Use of laser Doppler flowmetry and transcutaneous oxygen tension electrodes to assess local autonomic dysfunction in patients with frozen shoulder. *J Royal Soc Med* 1989;82:536-8.
11. Hinds T, McEwan I, Perkes J, et al. Effects of massage on limb and skin blood flow after quadriceps exercise. *Med Sci Sports Exerc* 2004;36:1308-13.
12. Okada K, Yamaguchi T, Minowa K, Inoue N. The influence of hot pack therapy on the blood flow in masseter muscle muscles. *J Oral Rehabil* 2005;32:480-6.
13. Poensin D, Carpentier PH, Fechoz C, Gasparini S. Effects of mud pack treatment on skin microcirculation. *Joint Bone Spine* 2003;70:367-70.
14. Hodges PW, Richardson CA. Inefficient muscular stabilization of the lumbar spine associated with low back pain. A motor control evaluation of transversus abdominis. *Spine* 1996;21:2640-50.
15. Richardson CA, Snijders CJ, Hides JA, et al. The relation between the transversus abdominis muscles, sacroiliac joint mechanics, and low back pain. *Spine* 2002;27:399-405.
16. Mulhearn S, George K. Abdominal muscle endurance and its association with posture and low back pain: An initial investigation in male and female elite gymnasts. *Physiother* 1999;85:210-60.
17. Shand D. Pilates to pit. *Lancet* 2004;363:1340.
18. Hagins M, Adler K, Cash M, et al. Effects of practice on the ability to perform lumbar stabilization exercise. *J Orthopaed Sports Phys Ther* 1999;29:549-55.
19. Phrompaet S, Paungmali A, Pirunsan U, Sitalertpisan P. Effects of Pilates training on lumbo-pelvic stability and flexibility. *Asian J Sports Med* 2011;2:16-22.

20. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 1998;26:217-38.
21. Portney LG, Watkins MP. *Foundations of Clinical Research Applications to Practice*. 2nd ed. New Jersey: Prentice Hall Health. 2000.
22. Sterling M. Testing for sensory hypersensitivity or central hyperexcitability associated with cervical spine pain. *J Manipulative Physiol Ther* 2008;31:534-9.
23. Jones DH, Kilgour RD, Comtois AS. Test-retest reliability of pressure pain threshold measurements of the upper limb and torsal in young healthy women. *J Pain* 2007;8:650-6.
24. Chanda ML, Alvin MD, Schnitzer TJ, Apkarian AV. Pain characteristic differences between subacute and chronic back pain. *J Pain* 2011;12:792-800.
25. Roeykens H, Van Maele G, De Moor R, Martens L. Reliability of laser Doppler flowmetry in a 2-probe assessment of pulpal blood flow. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1999;87:742-8.
26. Sindrup JH, Kastrop J, Kristensen JK. Diurnal variations in lower leg subcutaneous blood flow rate in patients with chronic venous leg ulcers. *Br J Dermatol* 1991;125:436-42.
27. Harris FC, Lahey BB. A method for combining occurrence and nonoccurrence interobservers agreement scores. *J Appl Behav Anal* 1978;11:523-7.
28. Park R, Wallace MS, Schulteis G. Relative sensitivity to alfentanil and reliability of current perception threshold vs von Frey tactile stimulation and thermal sensory testing. *J Peripheral Neuro Syst* 2001;6:232-40.
29. Felix ER, Widerstrom-Noga EG. Reliability and variability of quantitative sensory testing in persons with spinal cord injury and neuropathic pain. *J Rehab Res Develop* 2009;46:69-84.
30. Khamwong P, Nosaka K, Pirunsan U, Paungmali A. Reliability of muscle function and sensory perception measurements of the wrist extensor. *Physiother Theory Pract* 2010;26:408-15.
31. Scott D, Jull G, Sterling M. Wide spread sensory hypersensitivity is a feature of chronic whiplash-associated disorder but not chronic idiopathic neck pain. *Clin J Pain* 2005;21:175-81.
32. Guyton AC. Somatic sensations II: pain, headache, and thermal sensation. In: Guyton AC, Hall JE (eds). *Textbook of Medical Physiology*. 9th ed. Philadelphia: W.B. Saunders, 1996; Pp:619-20.