ORIGINAL

Stuart Cathcart
Donald Pritchard

Reliability of pain threshold measurement in young adults

S. Cathcart (✉)
Paraworklinks Inc.,
Suites 6&10 Sidney Chambers,
Elizabeth City Centre,
South Australia 5112, Australia
e-mail: scathcart@paraworklinks.com.au
Tel.: +61-8-8287-0578
Fax: +61-8-8287-0080

S. Cathcart • D. Pritchard
Department of Psychology,
University of Adelaide,
Adelaide, South Australia, Australia

Abstract The objective was to examine reliability of pressure and thermal (cold) pain threshold assessment in persons less than 25 years of age, using intra-class correlation (ICC) and coefficients of repeatability and variability. We measured thresholds to pain from pressure algometry and ice placed at the hand and head in 10 healthy volunteers aged 18–25. Intra-rater reliability was examined with ICC. Coefficients of repeatability (CR) and variability (CV) were estimated. Reliability of repeat assessments was high as assessed by ICC, although coefficients of repeatability and variation indicated considerable inter-individual variation in repeat measurements. Pressure algometry and strategically placed ice appear to be reliable techniques for assessing pain processing in young adults. Reliability studies employing ICC may benefit from complementary estimation of CR and CV.

Introduction

Previous reports have demonstrated high intra- and inter-rater reliability for assessment of pressure pain thresholds (PPT) in clinical [1–4] and healthy subjects [5–11]. However, many of the earlier reliability analyses used Pearson's *r* correlation coefficient [e.g.; 5, 9–11]. Pearson's *r* is a measure of association, but is not appropriate for assessing agreement between raters or ratings, as it is not sensitive to systematic bias between raters (or ratings) [12–14]. The intra-class correlation (ICC) is a more appropriate measure as it is sensitive to systematic bias. The ICC can provide a measure of absolute agreement between raters, which answers the question of whether raters or ratings are interchangeable [12–14]. The coefficient of repeatability (CR), specified as 2 standard deviations of the mean test-retest differences [14], provides an estimate of retest ranges expressed in the measurement units. Hence, 95% of repeat measurements for the sample will be in the range: mean difference±CR. The coefficient of variation (CV) is specified as half of the CR expressed as a percentage of the mean [14]. ICC, CR and CV therefore provide complementary information on reliability.

More recently, some studies of PPT reliability have used ICC to confirm reliability assessed by analogue algometers [15], electronic algometers [7, 16], dolorimeters [17] and novel computer-controlled devices [18].

Reliability has been confirmed in healthy [6–8] and chronic pain subjects such as fibromyalgia [2] and temporomandibular disorder sufferers [4]. However, most studies examining PPT reliability using ICC have assessed thresholds at either muscular [e.g.; 6, 7] or non-muscular locations [e.g.; 16]. Also, while reliability of repeat assessment over consecutive days has been established [2, 6, 19], few studies have assessed repeat within-session assessment [e.g.; 7]. We are not aware of any research using ICC, CV and CR to examine within-session repeatability of PPT at extra-cephalic and cephalic muscular and non-muscular locations in healthy subjects. Comparing ICC, CR and CV would be valuable, as ICC depends on CR and population standard deviation [13]. Hence, ICC results can be inflated by variability, reflected in CR and CV.

Assessment of thermal pain thresholds has been used in conjunction with mechanical thresholds to examine the relative importance of peripheral and central pain mechanisms in myofascial pain conditions such as tension-type headache [20, 21]. However, while there is considerable research on the reliability of PPTs, there is little on the reliability of thermal pain thresholds. That which has been conducted has predominantly focused on response to heat-induced pain and utilised Pearson's r as a reliability measure [22, 23]. Perkins et al. [23], however, found considerable within-subject variability in heat pain thresholds. Such a result indicates the potential benefit of using ICC and CV in future studies of thermal pain reliability. We are not aware of any studies using ICC, CR and CV to examine reliability of cold pain threshold assessment.

The use of ice placed at the temple and wrist may be a simple and reliable way to measure cephalic and extra-cephalic sensitivity to thermal (cold) stimulation in healthy and clinical subjects. Only one report has examined the reliability of thresholds to pain induced by strategically placed ice cubes [24]. That report found pain detection and tolerance thresholds were reliable at the temple over assessments separated by five days in healthy subjects. However, Pearson's r correlation was used. The within-session repeatability of strategically place ice cubes has not been examined to our knowledge, nor has the reliability at extra-cephalic locations.

Pain thresholds in the general population vary by age [25, 26]. The reliability of pain threshold measurement using both PPT and thermal (cold) stimulus in subjects less than 25 years of age has not been reported to our knowledge. Determining reliability of both procedures within the same session would be useful for future studies wishing to compare mechanical and thermal processing. In the present study, we used ICC, CR and CV to estimate the retest reliability and estimates of repeatability for within-session PPT assessment at the finger and at cephalic muscular and non-muscular sites. We also calculated ICC, CR and CV for discomfort and pain detection thresholds to ice cubes placed at the wrist and temple. The CR in this case is, therefore, within session (CRw). The aim was to confirm reliability of PPT and threshold to ice discomfort and pain in healthy subjects less than 25 years of age.

Methods

Subjects

Jensen et al. [9] demonstrate intra-individual variation of PPT can be estimated in groups of 10 healthy subjects with 80% power at the 0.05 significance level. We are not aware of a more recent study addressing this issue. We therefore recruited 10 healthy volunteers aged 18–25 from advertisements at this institution. The mean age of participants was 21 years (SD=2.6 years). Forty percent were male. Although gender differences have been observed in PPTs [27], we are not aware of data indicating gender differences in PPT reliability, particularly for the present method, which examined intra-individual repeatability over a 10-min period. We therefore included males and females in our sample and analyses. Subjects were included in the study following examination to rule out any current or previous major medical or psychiatric condition, and no spontaneous or chronic pain complaints. Subjects were also assessed for tenderness in cephalic and pericranial sites to rule out the possibility of latent algogenic foci in the areas to be examined (parieto-temporal muscular and non-muscular locations, as well as neck and shoulder tenderness). Subjects were excluded if they exhibited tenderness in these areas.

Procedures

Potential subjects were explained the procedures when they contacted the researchers and were invited to the electrophysiology laboratory at the University of Adelaide Psychology Department. Written consent was obtained. The study was approved by the University's Human Ethics Research Committee. Thresholds were measured while subjects were seated, with both hands rested on a table in front of them at approximately elbow height. Thresholds were assessed twice by one rater (SC), each assessment conducted 10 min apart in the order below. The first assessment was taken after a 5-min baseline during which subjects sat quietly. Subjects browsed through the local newspaper between assessments. All assessments were conducted between 1 pm and 4 pm Monday to Friday. The laboratory was temperature controlled at 22°C.

Pressure pain measurements

An analogue pressure algometer constructed in-house was used to measure pressure pain detection thresholds. Briefly, current

from a multi-meter is attenuated by a linear resistor attached to a spring-loaded probe. The probe tip is circular, 0.39 cm^2, with a hard rubber tip. The output is in kilo-ohms on the multi-meter. Data can be converted to kg/cm^2 according to previous calibration analyses, which demonstrated a linear relationship of the form $y=0.89+0.56x$. PPTs were measured in the following order: (1) the dorsal surface of the middle segment of the 1st phalange; (2) the central fibres of the temporalis muscle, identified by palpation above the superior margin of the ear; and (3) an adjacent parietal location without overlying muscle. The latter point was identified by having subjects alternately raise their eyebrows and clench their teeth while the investigator felt the anteromedial border of the temporal and medial border of the frontal muscle. The PPT was taken from the middle of the non-muscular region bisecting an approximately 50° imaginary line from the temporal muscle PPT point. Single measurements were taken at each location bilaterally with the left-side measurements taken first. Pressure was increased at approximately 0.5 kohms/s, which is equal to approximately 1.2 kg/s. Subjects were asked to say 'pain' at the point the pressure first became painful. Pressure was released when either the pain detection threshold had been reached and the investigator noted the kohms readout, or when the maximum pressure of the algometer (equal to 9.98 kohms) had been reached. Before the 5-min baseline, subjects were familiarised with pressure in non-painful ranges (0–0.5 kohms) to relieve potential anxiety over the assessment.

Ice pain measurements

Ice cubes in sealed plastic satchels were held against the wrist then temple on the left then right sides. The ice cubes had a 3.5x2.5 cm flat surface area. A small square of wax paper was placed between the ice and the subject to avoid startle effects from the initial ice application. Marlowe [24] demonstrated that the pressure of ice application ('gentle' and 'firm', subjectively defined) does not affect ice pain thresholds at the temple. We attempted to maintain a constant gentle pressure. Subjects were asked to say 'discomfort' and 'pain' at the respective times after ice application. A stopwatch with lap function was used to record time until the 2 thresholds were reached after the ice was first applied. The ice was removed after 3 min if pain was not reported.

Results

Analyses were conducted using the Statistical Package for the Social Sciences v8.0 (Vesta Services Inc., USA). Table 1 presents the means and standard deviations for the threshold measurements. Matched samples t-tests indicated no difference at the 0.05 significance level between time 1 and time 2 assessments for any measures. Table 1 also presents

Table 1 Reliability of pressure and cold threshold measurements

Measure	Time 1 Mean (SD)	Time 2 Mean (SD)	T1–T2 diff Mean (CRw)[a]	t[b]	ICC[c]	CV (%)[d]
Pressure pain threshold*						
Finger left	4.33 (2.1)	3.98 (1.8)	0.35 (2.92)	0.76	0.73	33.7
Finger right	4.86 (1.8)	4.28 (1.7)	0.58 (2.42)	1.50	0.75	24.8
Parietal left	3.23 (1.5)	3.24 (1.4)	0.01 (1.48)	0.04	0.88	22.9
Parietal right	3.03 (1.3)	3.34 (1.3)	0.32 (1.22)	1.63	0.87	20.1
Temporal left	2.58 (1.4)	2.16 (1.4)	0.42 (2.14)	1.24	0.69	41.4
Temporal right	2.26 (1.7)	2.04 (1.1)	0.22 (2.16)	0.63	0.72	47.8
Ice discomfort threshold#						
Temple left	21.3 (7.0)	19.8 (4.5)	1.47 (16.4)	1.12	0.75	38.0
Temple right	18.8 (5.8)	17.3 (4.6)	1.57 (14.4)	1.39	0.75	38.3
Wrist left	21.8 (8.7)	20.9 (9.6)	0.83 (15.2)	0.68	0.92	34.8
Wrist right	21.7 (8.6)	22.4 (7.4)	0.66 (14.0)	0.59	0.91	32.3
Ice pain threshold#						
Temple left	41.1 (19.4)	42.5 (20.7)	1.45 (16.4)	0.56	0.92	20.0
Temple right	39.4 (16.0)	38.1 (16.4)	1.33 (17.6)	0.48	0.86	22.3
Wrist left	45.2 (15.6)	38.4 (18.1)	6.75 (19.2)	2.22	0.79	21.2
Wrist right	41.9 (14.4)	39.8 (15.8)	2.14 (10.6)	1.28	0.94	12.6

* k-ohms/0.4 cm^2/s

Time in seconds

[a] Coefficient of repeatability; 2(SD mean T1/T2diff)

[b] Matched samples t-test, 2-tailed, df=9; all $p>0.10$, except left wrist pain threshold for which $p>0.05$

[c] Agreement model intra-class correlation

[d] Coefficient of variation; ((CR/2)/mean threshold)x100

ICC results and CRw estimates. Intra-class coefficients of 0.75 or above are generally considered as excellent [12]. The ICC coefficients for all measures were 0.75 or above, except PPT measurements at temporal locations bilaterally (r=0.69, 0.72) and at the left-side finger (r=0.73).

The CRw estimates for PPT ranged from 1.22 kohms for the right parietal region to 2.92 kohms for the left finger. The CVs for PPT ranged from 20.1% for the right parietal to 47.8% for the right temporal thresholds. The CRs for ice discomfort threshold measurements ranged from 14 s at the right wrist to 16.4 s at the left temple. The CVs for ice discomfort ranged from 32.3% at the right wrist to 38.3% at the right temple. Ice pain detection CRs ranged from 10.6 s at the right wrist to 19.2 s at the left wrist. Ice pain CVs ranged from 12.6% at the right wrist to 22.3% at the right temple.

Left and right side differences in thresholds were examined with matched samples t-tests (Table 2). Results indicated significant differences for the temple discomfort thresholds, with the right side being lower at both time 1 (p<0.05) and time 2 (p<0.05). No other laterality differences were significant. Left and right side measurements at

Table 2 Laterality of pressure and cold thresholds

Paired variables (L/R)[#]	L/R diff Mean (SD)	t[a]	r[b]
Time 1			
PPT finger	0.53 (1.66)	1.01	0.65*
PPT temporal	0.32 (1.12)	0.89	0.74*
PPT parietal	0.20 (0.54)	1.17	0.93+
IDT wrist	0.01 (5.19)	0.02	0.82+
IPT wrist	3.27 (7.19)	1.44	0.89+
IDT temple	2.45 (2.39)	3.25*	0.95+
IPT temple	1.65 (6.69)	0.78	0.95+
Time 2			
PPT finger	0.30 (0.81)	1.19	0.90+
PPT temporal	0.12 (0.61)	0.60	0.91+
PPT parietal	0.10 (0.75)	0.44	0.85+
IDT wrist	1.45 (7.17)	0.64	0.67*
IPT wrist	1.34 (11.19)	0.38	0.79+
IDT temple	2.55 (2.93)	2.75*	0.79+
IPT temple	4.43 (9.56)	1.47	0.89+

[#] Left *vs.* right side thresholds (means and SDs in Table 1)

PPT, pressure pain detection threshold (k-ohms/0.39 cm^2/s); *IDT*, ice discomfort detection threshold (s); *IPT*, ice pain detection threshold (s)

[a] Matched samples t-test, 2-tailed, df=9

[b] Pearson's r correlation

* p<0.05

+ p<0.01

each location were significantly correlated as assessed by Pearson's r coefficient. Pearson's r is preferable to ICC for such analysis as left and right side measurements represent different variables, hence, the appropriate question is one of association rather than 'agreement'.

Discussion

The ICC results for PPT are of similar magnitude to previous findings in healthy subjects [5, 7, 9, 11] and indicate excellent intra-rater reliability for all measures except temporal location and left finger, for which reliability as assessed by ICC was moderate to high. The ICC analyses also indicate excellent reliability for ice pain threshold measurements at both cephalic and extra-cephalic sites. The use of strategically placed ice cubes allows assessment of thermal pain processing at different body locations, and may therefore be useful for examining pathophysiology of disorders such as chronic tension-type headache, hypothesised to involve trigeminal level dysfunction particularly [20, 21]. For example, cephalic and extra-cephalic sites could be compared on cold pain sensitivity, as has been done previously for heat and pressure pain sensitivity in headache sufferers [20, 21].

Although differences in pressure and cold thresholds between time 1 and time 2 were not significant as assessed by t-test, the CRws indicate considerable inter-individual variation in T1–T2 changes for the temporal PPTs and for the temple ice discomfort thresholds. Thus, while the mean difference in temporal PPT between time 1 and time 2 was 0.22 kohms, 95% of repeat observations will, in fact, be within ±2.16 kohms. Such variability may have inflated the ICC results, which depend on the CR and the population standard deviation [1]. Similarly, the PPT CVs are relatively large for all measures when compared to other CV findings (e.g.; 18% after 15 min, 14% after 45 min and 29% after 5 weeks [9, 15]), but are of similar magnitude to more recent findings of PPT intra-individual variation [16]. The CVs indicate generally larger percent variation in repeat assessment for ice discomfort thresholds than for PPT, while ice pain thresholds had less percent variation on repeat assessment. The results are consistent with the ICCs in indicating ice pain detection as the most reliable of the measurements on repeat assessment. The greater variation in ice discomfort compared to pain thresholds was expected, thus, while threshold to pain is assumed to represent activation of nociceptors, we are not aware of a postulated discrete physiology for threshold to discomfort.

Consistent with previous findings, the present results indicate little laterality in pain detection thresholds to

pressure [11, 26] or ice [24] in healthy subjects. The finding that the left side temple discomfort threshold was consistently lower than that on the right side is a marked exception. That the pain detection threshold was not lower may suggest the difference is not due to nociceptive sensitivity. Pain detection thresholds to ice at the temple were lower on left compared to right sides in Marlowe's [24] data (20.0 s *vs*. 22.2 s), although the statistical significance of the difference was not reported. Our cold pain threshold findings were considerably higher than those found by Marlowe [24], which were similar to our discomfort thresholds. We intentionally modified Marlowe's method in order to provide greater sensitivity of the measure through increasing its scale; thus, whereas Marlowe used frozen satchels of water, we used ice placed in room temperature satchels, which were further separated from the skin by wax paper.

When converted to kilograms per square centimetre, the present mean kilo-ohms results are in the range of 2.04 kg at the right temporal location to 3.62 kg at the right finger. The ranges are similar to those in some previous studies [9, 10] but lower than others [1, 15]. The differences may be due to the size of the algometer applicator tip, which is 0.39 cm^2 in our study compared to the more commonly used 0.50–1 cm^2 tip in previous studies [5, 6, 16]. Intra-individual comparisons indicate that smaller algometer tips produce lower PPTs [9]. Application rate, which positively affects PPT [9], was similar in our study to that commonly used (e.g.; 1 kg/s). The possibility of a systematic over-estimation by the investigator of the actual application rate applied during testing cannot be ruled out at this stage.

Conclusions

The present results indicate that pressure algometry and strategically placed ice cubes are reliable measures for assessing pain sensitivity in young adults. Specifically, our results indicate acceptable reliability of the measures over within-session repeat assessment at extra-cephalic and cephalic muscular and non-muscular locations, in healthy subjects less than 25 years of age. While PPT and ice have been previously shown to be reliable when used individually, our results indicate that PPT and ice are also reliable when used in conjunction. However, our results also indicate large intra-individual variability in repeated pain threshold assessments, which may have inflated the ICCs. Future studies examining reliability and temporal characteristics of pain thresholds may therefore benefit from estimating CR and CV, particularly if ICC is used.

References

1. Sand T, Zwart JA, Helde G, Bovim G (1997) The reproducibility of cephalic pain pressure thresholds in control subjects and headache patients. Cephalalgia 17:748–755
2. Maquet D, Croisier JL, Demoulin C, Crielaard JM (2004) Pressure pain thresholds of tender point sites in patients with fibromyalgia and in healthy Controls. Eur J Pain 8:111–117
3. Smidt N, van der Windt DA, Assendelft WJ, Mourits AJ, DeVille WL, de Winter AF, Bouter LM (2002) Interobserver reproducibility of the assessment of severity of complaints, grip strength, and pressure pain threshold in patients with lateral epicondylitis. Arch Phys Med Rehab 83:1145–1150
4. Brown FF, Robinson ME, Riley JL, Gremillion HA, McSolay J, Meyers G (2000) Better palpation of pain: reliability of a new pain protocol in TMD. Cranio 18:58–65
5. Reeves JL, Jaeger B, Graff-Radford SB (1986) Reliability of the pressure algometer as a measure of myofascial trigger point sensitivity. Pain 24:313–321
6. Farella M, Michelotti A, Cimino R, Martina R (1999) An investigation of central and peripheral factors affecting pressure pain thresholds of the human jaw muscles. J Muscle Pain 1–2:253–259
7. Persson AL, Brogardh C, Sjolund BH (2004) Tender or not tender: test–retest repeatability of pressure pain thresholds in the trapezius and deltoid muscles of healthy women. J Rehab Med 36:17–27
8. Isselee H, De Laat A, Lesaffre E, Lysens R (1997) Short-term reproducibility of pressure pain thresholds in masseter and temporalis muscles of symptom free subjects. Eur J Oral Sci 105:583–587
9. Jensen K, Anderson HO, Olesen J, Lindblom U (1986) Pressure-pain threshold in human temporal region: evaluation of a new pressure algometer. Pain 25:313–323
10. Buchanan HM, Midgley JA (1987) Evaluation of pain threshold using a simple algometer. Clin Rheumatol 6:510–517
11. Orbach R, Gale EN (1989) Pressure pain thresholds in normal muscles: reliability, measurement effects, and topographic differences. Pain 37:257–263

12. Streiner DL (1995) Learning how to differ: agreement and reliability statistic in psychiatry. Can J Psychiatry 40:60–66
13. Shrout PE, Fleiss JL (1976) Intraclass correlations: uses in assessing rater reliability. Psychol Bull 86:420–428
14. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 8:307–310
15. Antonaci F, Sand T, Lucas GA (1998) Pressure algometry in healthy subjects: inter-examiner variability. Scand J Rehab Med 30:3–8
16. Brennum J, Kjeldsen M, Jensen K, Jensen TS (1989) Measurement of human pressure pain thresholds on fingers and toes. Pain 38:211–217
17. MacGregor AJ, Griffiths GO, Baker J, Specto TD (1997) Determinants of pressure pain threshold in adult twins: evidence that shared environmental influences predominate. Pain 73:253–257

18. Edwards T, Baker S, Eston R (1996) A method of detecting the muscle pain threshold using an objective software-mediated technique. Percept Mot Skills 82:955–960
19. Nussbaum EL, Downes L (1998) Reliability of clinical pressure-pain algometric measurements obtained on consecutive days. Phys Ther 78:160–169
20. Jensen R, Olesen J (1996) Initiating mechanisms of experimentally induced tension-type headache. Cephalalgia 16:175–182
21. Jensen R (1995) Mechanisms of spontaneous tension-type headache: an analysis of tenderness, pain thresholds and EMG. Pain 64:251–256
22. Langemark M, Jensen K, Jensen T, Olesen J (1989) Pressure pain thresholds and thermal nociceptive thresholds in chronic tension-type headache. Pain 38:203–210

23. Perkins KA, Grobe JE, Jennings JR, Epstein LH (1992) A technique for rapid, reliable assessment of thermal pain threshold in humans. Behav Res Methods Instrum Comput 24:6066
24. Marlowe N (1991) Pain sensitivity and headache: an examination of the central theory. J Psychosom Res 36:17–24
25. Lautenbacher S, Kunz M, Strate P, Nielsen J, Arendt-Nielsen L (2005) Age effects on pain thresholds, temporal summation and spatial summation of heat and pressure pain. Pain 115:410–418
26. Jensen R, Rasmussen BK, Pederson B, Lous I, Olesen J (1992) Cephalic muscle tenderness and pressure pain threshold in a general population. Pain 48:197–203
27. Chesterton LS, Barlas PF, Nadine E, Baxter GD, Wright CC (2003) Gender differences in pressure pain threshold in healthy humans. Pain 101:259–266