Rheumatology Advance Access published April 22, 2013

Original article

doi:10.1093/rheumatology/ket140

Assessing tactile acuity in rheumatology and musculoskeletal medicine—how reliable are two-point discrimination tests at the neck, hand, back and foot?

Mark J. Catley[1], Abby Tabor[1], Benedict M. Wand[2] and G. Lorimer Moseley[1,3]

Abstract

Objective. Chronic pain from rheumatic and musculoskeletal conditions is associated with cortical changes and altered tactile acuity. Tactile acuity is considered a clinical signature of primary somatosensory representation. The two-point discrimination (TPD) threshold is increasingly used both clinically and in research. Remarkably, the reliability and precision of the measure at commonly used sites has not been determined. This study aimed to determine the utility, intra- and inter-rater reliability, bias and variability of TPD threshold assessment at the neck, back, hand and foot using mechanical callipers.

Methods. Intra- and inter-rater reliability of TPD was assessed at the back, neck, hand and foot of 28 healthy young adults by 28 clinicians. Each clinician received training in the assessment of TPD using mechanical callipers and following a standardized protocol. Intraclass correlation coefficients (ICCs) and Bland–Altman plots were used to assess reliability, bias and variability.

Results. Intra-rater assessments in all four regions and inter-rater assessments at the neck and foot were reliable (ICC range 0.79–0.86), but large variability was seen in all assessments. Inter-rater assessment of the back (ICC = 0.66) and hand (ICC = 0.62) was deemed unreliable. Negligible systematic bias suggested learning did not affect reliability.

Conclusion. Individual clinicians are able to reliably assess TPD threshold at the neck, back, hand and foot using mechanical callipers. Measures obtained by different clinicians were only reliable for the neck and foot. Large variability was observed in all assessments, which suggests clinicians should be cautious when interpreting changes in tactile acuity in individual patients.

Key words: tactile acuity, two-point discrimination, cortical reorganisation, reliability, repeatability.

Introduction

Two-point discrimination (TPD) has been used extensively in clinical practice to evaluate the severity of peripheral nerve injuries and monitor patient recovery and response to treatment [1, 2]. While dependent on innervation density and intact neural pathways, tactile acuity is also dependent on response profiles of primary sensory cortex (S1) cells and the sensory neuraxis [3].

Altered response profiles of S1 neurones are termed cortical reorganisation. The discovery of consistent and substantial cortical reorganization in chronic rheumatic and musculoskeletal pain has led to widespread use of tactile acuity as a clinical signature of cortical reorganization [4]. The extent of tactile acuity impairment and the extent of cortical reorganization have been shown to directly relate to pain intensity in phantom limb pain [5] and complex regional pain syndrome (CRPS) [6]. Moreover, when patients recover, tactile acuity and cortical reorganization normalize [7, 8]. Decreased tactile acuity [9, 10] and cortical reorganization [11] have also been noted in chronic back pain and painful OA [12, 13], where the

[1]Sansom Institute for Health Research, University of South Australia, Adelaide, [2]School of Health Sciences, University of Notre Dame Australia, Fremantle, Western Australia and [3]Neuroscience Research Australia, Sydney, Australia.

Submitted 14 October 2012; revised version accepted 4 March 2013.

Correspondence to: G. Lorimer Moseley, University of South Australia, GPO Box 2471, Adelaide 5001, Australia.
E-mail: lorimer.moseley@gmail.com

extent of the changes has been associated with the duration of pain. Tactile discrimination training can normalize tactile acuity and cortical reorganization and, more importantly, reduce pain in phantom limb pain patients [14] and CRPS patients [15, 16], and there is preliminary evidence to these effects in chronic low back pain too [17] (see [4] for review).

TPD assessment has been widely criticized for the unexplained variability observed within subjects, between subjects and between studies [18]. Despite this criticism and despite its growing popularity, TPD has not been thoroughly investigated. Proponents rely on reliability measures for the fingertip, which is not relevant for chronic pain states. Moreover, the data were obtained using the Disk-Criminator [19], an octagonal-shaped tool with standardized inter-stimulus spacings that range from 1 to 15 mm, or its crude alternative, the paperclip (see [20] for a review). The Disk-Criminator and the paper clip are inappropriate for assessment of the neck, back and feet, where TPD thresholds are thought to be notably larger [21–23].

The utility of any clinical measure depends on the cost of equipment and required clinician training. Hardware-style mechanical callipers provide a wide-range alternative to custom tools such as the Disk-Criminator for about 20% of the cost, but they have not been tested. The aim of this study was to determine in a large cohort of clinicians with variable experience and minimal training, in a clinically pragmatic fashion, the utility, intra- and inter-rater reliability, bias and variability of TPD threshold assessment at the neck, back, hand and foot using hardware-style mechanical callipers.

Materials and methods

Subjects and clinicians

In this study, the clinicians and the healthy volunteers were considered participants. For the sake of clarity, however, we refer to the healthy volunteer participants as subjects and the clinician participants as clinicians. A convenience sample of healthy subjects was sought via flyers and social media. Previous studies suggest tactile acuity is reduced in patients with neurological disease [19], peripheral nerve injury [1] and chronic pain [6, 11] and, in healthy persons, diminishes over the age of 35 [24]. For this reason, subjects were excluded if they had current pain, neurological disease (or overt neurological signs), were unable to detect light touch or were over 35 years of age. For clinicians, physiotherapists were recruited from several practices within the metropolitan area as part of a separate project. Each physiotherapist had completed a 4-year bachelor programme in an Australian university and several clinicians had completed an additional masters programme in either sports physiotherapy or musculoskeletal physiotherapy. Previous experience in the assessment of TPD was not required. Clinicians completed a short questionnaire related to years and nature of their clinical experience.

Study design

Both intra-rater and inter-rater reliability were assessed at four locations on each subject. Clinicians and subjects were randomly assigned into seven groups of four (two clinicians and two subjects in each group). To establish intra-rater reliability, each clinician assessed one of the subjects twice, with a 30-min interlude between each assessment. To establish inter-rater reliability, each clinician assessed the other subject within the group prior to the interlude. That is, each clinician assessed the TPD thresholds of two subjects. Clinicians were instructed not to discuss the results with the other clinician or the subjects. Each measurement was recorded by the clinicians and collected by the authors immediately following each assessment. This study was conducted in compliance with the Helsinki Declaration and ethics approval was provided by the University of South Australia Human Research Ethics Committee.

Test areas and positioning

Subjects nominated their dominant side according to their response to the question—would you judge yourself to be more right-handed or more left-handed? TPD was then assessed at the neck, back, hand and foot on the side they nominated as dominant. Subjects lay prone, with their dominant arm pronated to enable assessment of the palmar aspect of the hand. Neck assessment was conducted with the neck in neutral position and the subject looking downward through the face hole of the plinth. The subject's feet were supported on a pillow placed under the ankles.

For the neck and back, clinicians were instructed to locate the spinous process of C7 and L3, respectively, and assess TPD horizontally out from the midline towards the subject's dominant side. For the hand, the clinicians were instructed to locate the pisiform on the palmar aspect of the hand and assess TPD distally along the hypothenar eminence. For the foot, clinicians were instructed to locate the base of the fifth metatarsal on the lateropalmar aspect of the foot and assess TPD distally along the lateral margin of the sole.

TPD assessment

All clinicians underwent a brief training session on the evaluation of TPD according to the method described by Moberg [25]. The sites of assessment were not marked because locating the site of testing was deemed to be a crucial part of TPD assessment. Mechanical sliding callipers (Duratech TA-2081) purchased from a hardware store were used (supplementary Fig. S1, available at *Rheumatology* Online). The callipers have a precision of 1 mm and were applied with pressure sufficient to first blanch the skin. Assessment commenced with 0 mm between the two points and gradually increased until the subject discerned two points [9]. A series of five ascending and descending assessments, centred around the subject's TPD threshold, was conducted and the average of these assessments was analysed. Each clinician was given the opportunity to practise the technique at each of

the four test areas on multiple subjects prior to commencing the trial.

Each subject was instructed to report after each application, 1 if they felt one point or 2 if they felt two points [25]. If they were unsure, they were instructed to report one point. The only feedback they were to give to the assessor was if they discerned two points because of a temporal delay between each point. When this occurred, that report was rejected.

Statistical analysis

Intra- and inter-rater reliability was assessed via intraclass correlation coefficients (ICCs) and Bland and Altman's 95% limits of agreement [26]. ICCs, used to determine the absolute agreement between assessments, were calculated using PASW Statistics 18 (v18.0.0, IBM Corporation, New York). A two-way random model was chosen because both the subjects and the clinicians were considered random effects [27]. ICC values were categorized according to Portney and Watkins [28]: $\geqslant 0.75$ was interpreted as good reliability and <0.75 was interrupted as poor to moderate reliability.

Bland–Altman plots were constructed using MedCalc (v12.2.0.0) to determine bias, variability and agreement. This method plots the difference between two measurements against the mean of the two measurements. The difference between any two measurements is expected to lay within the limits of agreement, i.e. 1.96 S.D. above or below the mean of the difference [29]. The actual width of these limits will have implications for the tool's clinical use. Prior to analysis, histograms displaying the difference between intra- and inter-rater measurements were inspected for normality, which is an assumption of the limits of agreement method [26]. Agreement, variability, bias and heteroscedasticity were assessed initially visually by examining the spread of difference scores in relation to the mean difference and 95% limits of agreement. Heteroscedasticity (whether the differences are dependent on the magnitude of the mean) was formally assessed by calculating the correlation coefficient between the differences and the means. Mean differences substantially different from zero are indicative of learning or fatigue effects [29].

The effect of clinician experience on test–retest reliability was investigated using an independent *t*-test. Because this was an exploratory analysis and we wanted to minimize the likelihood of a false-negative result, significance was set at $\alpha = 0.1$. An arbitrary cut-off of greater than 5 years clinical practice was chosen a priori to categorize clinicians as inexperienced or experienced.

Results

Demographic data

Tactile acuity was assessed on 28 (11 male) healthy subjects by 28 (19 male) clinicians. Demographic data for both subjects and clinicians are shown in Table 1. The mechanical callipers used in this study were purchased for AU$15.00. The total training duration was approximately 30 min and each individual assessment of TPD took approximately 3 min.

Normative values and reliability

Normative values and summary statistics are shown in Table 2 and Figure 1. Intra-rater reliability at each of the four sites was good (ICC 0.84–0.96). Inter-rater reliability at the neck and foot was also good (ICC 0.78–0.82), and inter-rater reliability at the back and hand was moderate (ICC 0.62–0.66). The CIs were greater for inter-rater reliability than for intra-rater reliability, but were large for both. Interestingly, the lower bounds were all <0.75.

Bland–Altman plots for the intra- and inter-rater performances are shown for each assessment site in Fig. 2 and the mean differences are shown in Table 2. Differences were all judged to be normally distributed, which meant that the limits of agreement method could be used [26]. The mean differences in all plots were close to zero, suggesting that there was no systematic learning or fatigue effects. The plotted differences showed large variability, indicative of error, suggesting that TPD assessment is reliable but not precise. However, the plotted differences were spread evenly and randomly above and below the mean, suggesting homoscedasticity. Homoscedasticity was further confirmed by a lack of correlation between the absolute differences and the individual means for each of the comparisons. The experience of the clinician had no effect on TPD measures. That is, there were no effects of the independent *t*-test, even with a liberal α of 0.1 ($P > 0.15$ for all tested regions).

Discussion

The aim of this study was to determine, in a large cohort of clinicians with variable experience and minimal training, and in a clinically pragmatic fashion, the utility, intra- and inter-rater reliability, bias and variability of TPD threshold assessment at the neck, back, hand and foot using hardware-style mechanical callipers. The results suggest that if the same clinician performs the test each time, TPD is a reliable measure across all four sites. This is important because tactile acuity is recognized as a clinical signature of cortical reorganization and TPD threshold is considered the key clinical assessment of cortical reorganization in rheumatology and musculoskeletal medicine [4]. The current results appear to endorse the use of TPD threshold as a method for a single clinician to evaluate tactile acuity and response to treatment. However, if different clinicians perform the test, the reliability becomes unacceptable, particularly for the back or hand. This means that pooling or comparison of TPD thresholds between patients becomes problematic if measures are taken by more than one clinician, which is often the case when patients see multiple practitioners.

Particularly relevant to clinical practice is the finding that clinical experience had no effect on TPD measures. Clinicians underwent basic training in the technique and had a chance to practise it, but this whole process took

less than 30 min. Some of the clinician cohort had extensive experience in assessing TPD as part of their usual practice, but they appeared to be no more or less reliable than the complete novices. This is encouraging insofar as the utility of any measure depends on how much the equipment costs and how time consuming and difficult it is to perform. The callipers used here cost about AU$15 each, the assessment took about 3 min and proficiency was obtained with minimal training. On the other hand, the lack of effect of experience is discouraging insofar as we can expect no improvement in reliability or precision with practice. On these grounds, TPD threshold assessment appears to be a useful clinical tool, but an imprecise absolute measure.

The large variability seen in assessment of all four regions was highlighted by the Bland–Altman plots. Random error due to within-subject factors and clinician factors in TPD assessment has been reported extensively in the literature, with a focus on neurological assessment [1, 18]. Within-subject factors known to influence threshold measurement include age [24, 30], skin temperature [31, 32], epidermis thickness [33], body hair [21], patient co-operation [1] and fatigue [34]. Additionally, clinician factors such as the protocol [1], tool [35], skill and technique [21], concentration [36] and application pressure [25, 37] contribute to variability. It is likely many of these factors, with the exception of the protocol, which was standardized here, contributed to the variability seen in this study, and would also contribute to variability in the clinic. As such, caution is advised because large changes in TPD threshold would need to be evident before they can be confidently attributed to a true change rather than variability of the measure (i.e. chance). For example, TPD threshold of the lower back in this study was about 55 mm, which corroborates the only other report we could find of normative values [23]. Discounting the larger variability seen in patient populations [38], these values suggest that a change in TPD threshold would need to be in excess of 15 mm to be attributed to anything but error. According to our results, this minimal detectable change would increase to 24 mm for the neck. Future studies will clearly need to consider these limits when calculating sample sizes [29].

A person's tactile acuity itself improves with practice. That is, twice-daily assessment of TPD resulted in a substantially reduced threshold over a 1-month period (Dressler 1894 cited in [18]). Milerian and Tkachenko (1963, cited in [18]) repeated Dressler's study in eight healthy subjects, noting similar improvement accompanied by consequent improvement in 15 remote sites distributed over both arms and back. Together these studies imply that repeated assessment over time constitutes a training effect and this improvement in tactile acuity could thus influence reliability. In patient populations with reduced tactile acuity, as little as 30 min of tactile training was sufficient to induce a significant improvement in TPD threshold, which was maintained for up to 48 h [16]. Godde *et al.* [39] found that 2 h of tactile training induced significant improvements in TPD, but 30 min of training was insufficient, and Finnell *et al.* [35] found no learning effect with repeated TPD assessment in the fingertip. Thus

Table 1 Clinician and subject demographic characteristics

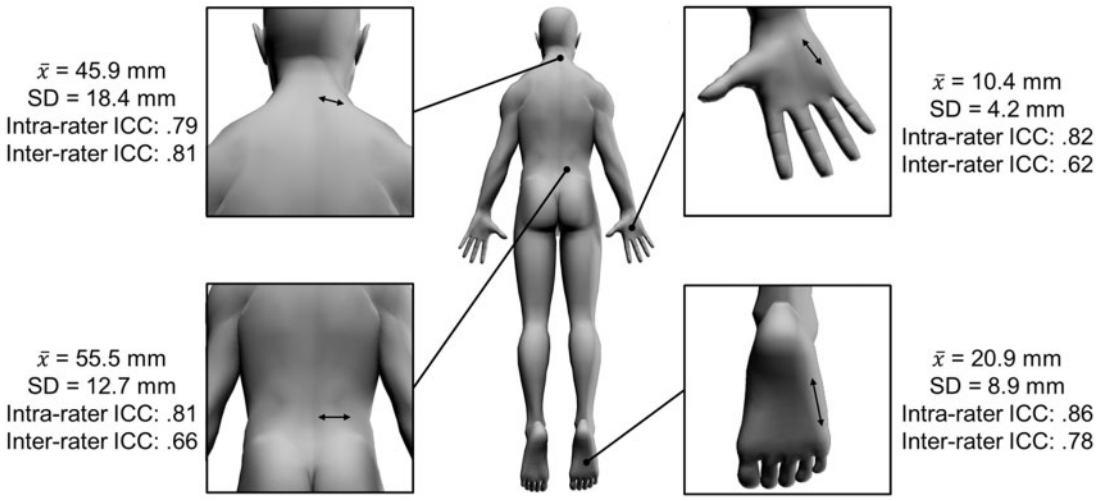
	Clinicians ($n = 28$)	Subjects ($n = 28$)
Gender, *n* (%)		
Males	11 (39.3)	19 (67.9)
Females	17 (60.7)	9 (32.1)
Age, mean (S.D.), years	32.8 (13.1)	24.1 (4.7)
Qualification, *n* (%)		
Bachelor	19 (67.9)	—
Masters	9 (32.1)	—
Experience, mean (S.D.), years	9.8 (13.1)	—
Handedness, *n* (%)		
Right	—	24 (85.7)
Left	—	4 (14.3)
TPD threshold, mean (S.D.), mm		
Back	—	55.5 (12.7)
Neck	—	45.9 (18.4)
Hand	—	10.4 (4.2)
Foot	—	20.9 (8.9)

Table 2 Intra- and inter-rater comparison of TPD assessment in young healthy adults

			Intra-rater reliability			Inter-rater reliability		
				95% CI limits			95% CI limits	
Site	Intra-rater difference, mean (S.D.), mm	Inter-rater difference, mean (S.D.), mm	$ICC_{(2,1)}$	Lower	Upper	$ICC_{(2,1)}$	Lower	Upper
Neck	−0.2 (12.4)	0.6 (12.4)	0.79[a]	0.59	0.90	0.81[a]	0.63	0.91
Back	0.6 (7.5)	1.9 (10. 7)	0.81[a]	0.63	0.91	0.66	0.38	0.82
Hand	0.4 (2.4)	−0.1 (4.2)	0.82[a]	0.65	0.91	0.62	0.33	0.80
Foot	1.5 (4.27)	−0.7 (6.4)	0.86[a]	0.72	0.96	0.78[a]	0.57	0.89

[a]Indicates good reliability (i.e. $\geqslant 0.75$) [28].

Fig. 1 TPD was assessed at the neck, back, hand and foot.



Mean TDP thresholds measured in 28 healthy subjects are reported for each of the four locations assessed. The reliability of TPD assessment, measured with mechanical callipers, was assessed by 28 trained physiotherapists. ICCs for each location are reported for intra- and inter-rater reliability.

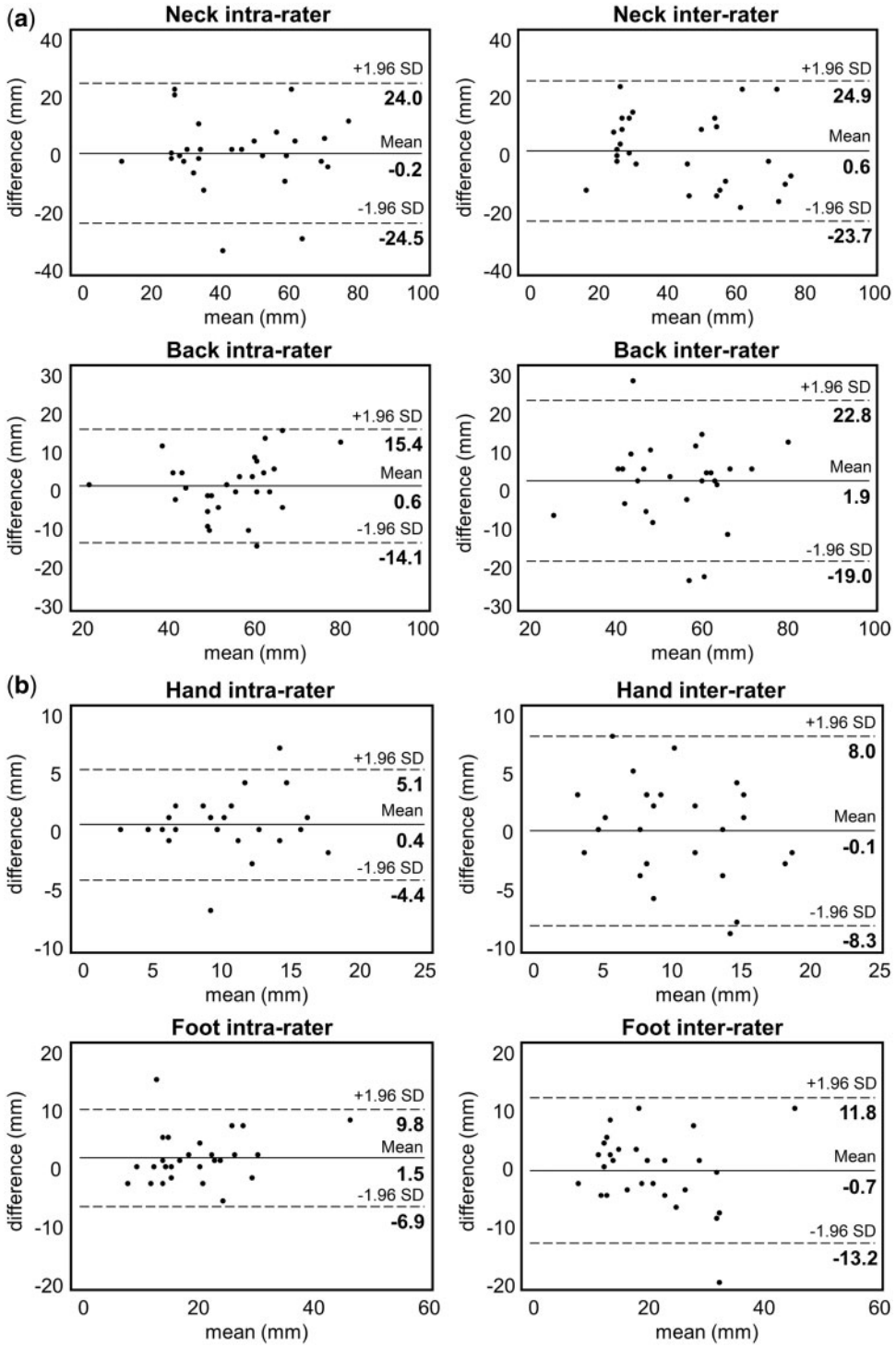
it seems that the learning threshold for tactile acuity lies somewhere between a repeated assessment and about 30 min of training.

Normative values have been reported for TPD on the neck and back [23], but to our knowledge, the current data represent the first normative data for the volar aspect of the foot and the palmar aspect of the hand in an adult population. The intra- and inter-rater reliability of TPD assessment at the fingertip has been debated in the literature. Drawing comparisons between studies is difficult because they vary in their choice of tools, their protocol, their statistical analysis and the sample used for subjects and clinicians. The Disk-Criminator is the most widely reported tool in the reliability literature [20]. Finnell *et al.* [35] reported high inter-rater reliability ($\kappa = 0.79–1.00$) in healthy subjects, but used a protocol whereby TPD was assessed at four intervals (4, 5, 6 and 8 mm) and was only powered to detect a 25% difference between assessors. Novak *et al.* [40] also reported similar high inter-rater reliability (ICC 0.989) in a mixed sample of healthy, blind and nerve-injured persons, and Dellon *et al.* [41] reported good inter-rater reliability ($r = 0.917–0.961$) in a sample of nerve-injured patients but used the interclass coefficient, thus assessing association rather than agreement [42]. In contrast, Rozental *et al.* [43] reported only slight to fair inter-rater ($\kappa = 0.12–0.22$) and intra-rater ($\kappa = 0.09–0.24$) reliability, suggesting that the discriminator was an unreliable assessment in healthy subjects. The paperclip was only reported in one study, where the authors suggested it as a cheap, readily available alternative to the Disk-Criminator and noted that, when calibrated, it provides similar results [35]. Despite the lack of consensus as to whether it is reliable, the Disk-Criminator is only suitable for assessment of regions with dense innervation. Accurate measurement (1 mm increments) is limited to between 0 and 15 mm, hence it is not suited to assessment of any of the regions assessed in this study. Tools such as mechanical callipers or compass-style tools such as that used by Nolan [21] appear to be more appropriate for assessment of the limbs and body.

In the present study, TPD assessment was conducted under conditions similar to those encountered in clinical practice. The test protocol was standardized but not monitored. Each subject was assessed by each clinician on only two occasions and we did not attempt to control for environmental conditions such as the ambient temperature of the room. Follow-up assessment of TPD threshold was not practical given the number of clinicians and subjects involved and hence the findings are only generalizable to assessments taken on the same day. It is plausible that reassessment over time could show larger variability due to fluctuation of within-subject factors or systematic learning, but future research will need to address this issue. The sample was also limited to young healthy subjects, as we controlled for age by excluding persons older than 35 years. As both age [24] and pain are associated with larger error margins [38], it is also plausible that older persons and persons in pain will have even greater within-subject variability, despite clinician factors remaining stable. We chose specific regions that are commonly assessed in clinical practice and research. Although we would predict that the values for the back would be replicated at nearby segmental levels, or the values for the foot would be replicated on the dorsal or medial surface, we cannot be sure.

In conclusion, clinicians with variable experience and minimal training are able to quickly and reliably assess TPD threshold at the neck, back, hand and foot

Fig. 2 Bland–Altman plots for TPD assessment of the neck, back, hand and foot.



Intra- and inter-rater reliability of TPD assessment is shown for (**a**) the neck and back and (**b**) the hand and foot. The difference in TPD measurements (*y*-axis) is plotted against the mean TPD measurement (*x*-axis) with the mean difference (bias) (continuous line) and 95% CI of the mean difference (limits of agreement) (dashed lines).

using hardware-style, inexpensive mechanical callipers. Measures obtained by different clinicians were only reliable for the neck and foot. Large variability was observed in all assessments, which suggests we should be cautious when interpreting changes in tactile acuity in individual patients. Researchers also need to account for this variability when calculating suitable sample sizes. Limited reassessments in healthy subjects do not appear to induce a training effect, but further research is needed to assess whether repeated assessments over time will influence reliability.

Rheumatology key messages

- Cortical reorganization and altered tactile acuity are both associated with chronic pain conditions including osteoarthritis.
- Individual clinicians can reliably assess TPD thresholds at the neck, back, hand and foot using callipers.
- Inter-rater comparisons of TPD are only reliable for the neck and foot.

Acknowledgements

M.J.C. is supported by an Australian post-graduate award. A.T. is supported by the University of South Australia President's Scholarship. G.L.M. is supported by the National Health and Medical Research Council of Australia (ID 571090).

Funding: This project was supported by NHMRC Project Grant (ID 1008017).

Disclosure statement: The authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at *Rheumatology* Online.

References

1 Lundborg G, Rosen B. The two-point discrimination test: time for a re-appraisal? J Hand Surg Am 2004;29:418–22.

2 Jerosch-Herold C. Assessment of sensibility after nerve injury and repair: a systematic review of evidence for validity, reliability and responsiveness of tests. J Hand Surg Am 2005;30:252–64.

3 Lotze M, Moseley GL. Role of distorted body image in pain. Curr Rheumatol Rep 2007;9:488–96.

4 Moseley GL, Flor H. Targeting cortical representations in the treatment of chronic pain: a review. Neurorehabil Neural Repair 2012;26:646–52.

5 Flor H, Elbert T, Knecht S *et al*. Phantom-limb pain as a perceptual correlate of cortical reorganization following arm amputation. Nature 1995;375:482–4.

6 Pleger B, Ragert P, Schwenkreis P *et al*. Patterns of cortical reorganization parallel impaired tactile discrimination and pain intensity in complex regional pain syndrome. Neuroimage 2006;32:503–10.

7 Pleger B, Tegenthoff M, Ragert P *et al*. Sensorimotor returning in complex regional pain syndrome parallels pain reduction. Ann Neurol 2005;57:425–9.

8 Maihöfner C, Handwerker HO, Neundörfer B *et al*. Cortical reorganization during recovery from complex regional pain syndrome. Neurology 2004;63:693–701.

9 Moseley GL. I can't find it! Distorted body image and tactile dysfunction in patients with chronic back pain. Pain 2008;140:239–43.

10 Luomajoki H, Moseley GL. Tactile acuity and lumbopelvic motor control in patients with back pain and healthy controls. Br J Sports Med 2011;45:437–40.

11 Flor H, Elbert T, Braun C *et al*. Extensive reorganization of primary somatosensory cortex in chronic back pain patients. Neurosci Lett 1997;224:5–8.

12 Gwilym SE, Filippini N, Douaud G *et al*. Thalamic atrophy associated with painful osteoarthritis of the hip is reversible after arthroplasty: a longitudinal voxel-based morphometric study. Arthritis Rheum 2010;62:2930–40.

13 Stanton TR, Lin CWC, Bray H *et al*. Tactile acuity is disrupted in osteoarthritis but is unrelated to disruptions in motor imagery performance. Rheumatology 2013, doi: 10.1093/rheumatology/ket139.

14 Flor H, Denke C, Schaefer M *et al*. Effect of sensory discrimination training on cortical reorganisation and phantom limb pain. Lancet 2001;357:1763–4.

15 Moseley GL, Zalucki NM, Wiech K. Tactile discrimination, but not tactile stimulation alone, reduces chronic limb pain. Pain 2008;137:600–8.

16 Moseley GL, Wiech K. The effect of tactile discrimination training is enhanced when patients watch the reflected image of their unaffected limb during training. Pain 2009;144:314–9.

17 Wand BM, O'Connell NE, Di Pietro F *et al*. Managing chronic nonspecific low back pain with a sensorimotor retraining approach: exploratory multi-baseline study of 3 participants. Phys Ther 2011;91:535–46.

18 Johnson KO, Van Boven RW, Hsiao SS. The perception of two points is not the spatial resolution threshold. In: Boivie J, Hansson P, Lindblom U, eds. Touch, temperature, and pain in health and disease: mechanisms and assessments. Seattle, WA: IASP Press, 1994:389–404.

19 Mackinnon SE, Dellon AL. Two-point discrimination tester. J Hand Surg Am 1985;10:906–7.

20 Auld ML, Boyd RN, Moseley GL *et al*. Tactile assessment in children with cerebral palsy: a clinimetric review. Phys Occup Ther Pediatr 2011;31:413–39.

21 Nolan MF. Two-point discrimination assessment in the upper limb in young adult men and women. Phys Ther 1982;62:965.

22 Nolan MF. Limits of two-point discrimination ability in the lower limb in young adult men and women. Phys Ther 1983;63:1424.

23 Nolan MF. Quantitative measure of cutaneous sensation. Two-point discrimination values for the face and trunk. Phys Ther 1985;65:181.

24 Woodward KL. The relationship between skin compliance, age, gender, and tactile discrimination thresholds in humans. Somatosens Mot Res 1993;10:63–7.

25 Moberg E. Two-point discrimination test: a valuable part of hand surgical rehabilitation, e.g. in tetraplegia. Scand J Rehabil Med 1990;22:127–34.

26 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.

27 McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996; 1:30–46.

28 Portney LG, Watkins MP. Foundations of clinical research: applications to practice. Upper Saddle River, NJ: Pearson/ Prentice Hall, 2009.

29 Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998;26:217–38.

30 Sohn SA, Simons BP. Functional anatomy of the hand. In: Wolford EG, ed. Acute hand injuries: a multidisciplinary approach. Boston: Little, Brown, 1980:32–42.

31 Mackworth NH. Finger numbness in very cold winds. J Appl Physiol 1953;5:533–43.

32 Sinclair D. Psychophysiology of cutaneous sensation. In: Jarret A, ed. Physiology and pathophysiology of the skin. London: Academic Press, 1973:429–73.

33 Omer GE. Sensibility testing. In: Omer GE, Spinner D, eds. Management of peripheral nerve problems. Philadelphia: Saunders, 1980:3–15.

34 Porter RW. New test for finger-tip sensation. Br Med J 1966;2:927–8.

35 Finnell JT, Knopp R, Johnson P *et al*. A calibrated paper clip is a reliable measure of two-point discrimination. Acad Emerg Med 2004;11:710–4.

36 Moberg E. Objective methods for determining the functional value of sensibility testing in the hand. J Bone Joint Surg Br 1958;40:454–76.

37 Kaneko A, Asai N, Kanda T. The influence of age on pressure perception of static and moving two-point discrimination in normal subjects. J Hand Ther 2005;18: 421–5.

38 Wand B, Di PF, George P *et al*. Tactile thresholds are preserved yet complex sensory function is impaired, over the lumbar spine of chronic non-specific low back pain patients: a preliminary investigation. Physiotherapy 2010; 96:317–23.

39 Godde B, Stauffenberg B, Spendler F *et al*. Tactile coactivation-induced changes in spatial discrimination performance. J Neurosci 2000;20:1597–604.

40 Novak CB, Mackinnon SE, Williams JI *et al*. Establishment of reliability in the evaluation of hand sensibility. Plast Reconstr Surg 1993;92:311–22.

41 Dellon AL, Mackinnon SE, Crosby PM. Reliability of two-point discrimination measurements. J Hand Surg Am 1987;12:693–6.

42 Field AP. Intraclass correlation. In: Everitt BS, Howell DC, eds. Encyclyopedia of statistics in behavioral science. Chichester: John Wiley & Sons Ltd, 2005.

43 Rozental TD, Beredjiklian PK, Guyette TM *et al*. Intra- and interobserver reliability of sensibility testing in asymptomatic individuals. Ann Plast Surg 2000;44:605–9.