Neuroscience Letters 579 (2014) 157–162

Contents lists available at ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet

Test–retest reliability of thermal quantitative sensory testing on two sites within the L5 dermatome of the lumbar spine and lower extremity

I.A. Knutti[a,b], M.R. Suter[c], E. Opsommer[b,*]

[a] Bern University of Applied Sciences (BFH), Health, Murtenstrasse 10, 3008 Bern, Switzerland
[b] University of Health Sciences (HESAV), University of Applied Sciences and Arts Western Switzerland (HES-SO), Av. de Beaumont 21, 1011 Lausanne, Switzerland
[c] Pain Center, Department of Anesthesiology, University Hospital Center (CHUV) and University of Lausanne (UNIL), Rue du Bugnon 46, 1011 Lausanne, Switzerland

HIGHLIGHTS

- Quantitative sensory testing (QST) is used to assess sensory disturbances.
- These findings showed excellent intra-rater reliability of thermal QST on the feet.
- Intra-rater reliability of thermal QST on the lumbar spine was fair to excellent.

ARTICLE INFO

Article history:
Received 15 March 2014
Received in revised form 25 May 2014
Accepted 11 July 2014
Available online 23 July 2014

Keywords:
Quantitative sensory testing
Reproducibility of results
Pain threshold
Pain measurement

ABSTRACT

Introduction Quantitative sensory testing (QST) is widely used in human research to investigate the integrity of the sensory function in patients with pain of neuropathic origin, or other causes such as low back pain. Reliability of QST has been evaluated on both sides of the face, hands and feet as well as on the trunk (Th3-L3). In order to apply these tests on other body-parts such as the lower lumbar spine, it is important first to establish reliability on healthy individuals. The aim of this study was to investigate intra-rater reliability of thermal QST in healthy adults, on two sites within the L5 dermatome of the lumbar spine and lower extremity. *Methods* Test–retest reliability of thermal QST was determined at the L5-level of the lumbar spine and in the same dermatome on the lower extremity in 30 healthy persons under 40 years of age. Results were analyzed using descriptive statistics and intraclass correlation coefficient (ICC). Values were compared to normative data, using *Z*-transformation. *Results* Mean intraindividual differences were small for cold and warm detection thresholds but larger for pain thresholds. ICC values showed excellent reliability for warm detection and heat pain threshold, good-to-excellent reliability for cold pain threshold and fair-to-excellent reliability for cold detection threshold. ICC had large ranges of confidence interval (95%). *Conclusion* In healthy adults, thermal QST on the lumbar spine and lower extremity demonstrated fair-to-excellent test–retest reliability.

1. Introduction

Quantitative sensory testing (QST) is a method used to test all somatosensory submodalities (i.e. touch, vibration, temperature or pain) with different kinds of calibrated stimuli, to examine the presence of negative or positive sensory signs [1]. QST can be applied to investigate the integrity of the sensory function in order to define and classify pathologies, to analyze pathogenesis or to evaluate changes in diseases [2]. Over the past decades, there has been an increasing interest in QST in clinical and research settings, for example to determine the conduction velocity of peripheral nerve fibres [3] or to assess a treatment's effectiveness [4]. QST is a psychophysical measurement, relying on the subjective perception of a physical stimulus [5]. Therefore, reliability needs to be assessed with a very rigorous methodology. In their systematic review, Moloney and colleagues [6] reported large variability in

* Corresponding author. Tel.: +41 21 316 81 24; fax: +41 21 316 81 02.
E-mail addresses: iknutti@bluewin.ch, isabelle.knutti@hesav.ch (I.A. Knutti), marc.suter@chuv.ch (M.R. Suter), emmanuelle.opsommer@hesav.ch (E. Opsommer).

http://dx.doi.org/10.1016/j.neulet.2014.07.023
0304-3940/

methodological quality of published studies with poor-to-excellent reliability of thermal QST. The German Research Network on Neuropathic Pain (DFNS) published a standardized QST protocol and normative data for quantitative sensory evaluations [7]. Under this protocol, that QST battery was evaluated bilaterally on the face, hands and feet using the method of limits, with reliability then assessed on these different sites [8]. The availability of this standardized protocol tends to improve reliability [9]. Indeed, before using QST on other body parts, its reliability needs to be assessed as recently shown by Pfau and colleagues [10] for the upper (Thoracic(Th)2-Th8) and the lower back (Th10-Lumbar(L)3). Among the diverse sensory submodalities described in the DFNS standardized protocol, the thermal modalities include cold and warm detection threshold (CDT, WDT), paradoxical heat sensation (PHS) as well as cold and heat pain perception threshold (CPT, HPT) [8]. All thresholds are given in °C (continuous scale) and can be tested using a thermotesting device by trained investigators.

QST has mostly been used in patients with neuropathic pain [2] but also in different musculoskeletal pathologies such as low back pain (LBP) [1]. People with LBP may have sensory disturbances in the back with or without radicular symptoms in the lower limb. QST has been used in some studies to assess these disturbances [11–13]. Nevertheless, stimulation sites were not combining two sites within the same dermatome of the lumbar spine and the lower extremity in these studies and reliability of thermal QST has, to our knowledge, never been assessed on the lumbar paravertebral area, corresponding to the L5-dermatome.

Given the increased use of QST to assess sensitivity in the lumbar spine in patients with LBP [13], reliability of QST-measurement in this region needs to be assessed. Furthermore, testing a second site on the same dermatome will allow further studies to be performed on patients. Therefore, the aim of this study was to determine test–retest reliability of thermal QST on two sites within the same dermatome (L5) of the lumbar spine and the lower extremity in healthy adults younger than 40 years of age, according to the protocol of the DFNS [8]. After examining reliability, data obtained on the lower extremity were compared to the normative values of the DFNS for the same site [7] and then to the data obtained in our study on the lumbar spine within the same dermatome.

2. Methods

Healthy volunteers were recruited from the staff and student population of the University of Health Sciences, Physical Therapy Department and the University Hospital Centre (CHUV), Lausanne, through an e-mail campaign. Inclusion criteria were: (1) good health status with no lower-back pain or lower-limb pain, (2) age less than 40 years, and (3) ability to read and speak French. Participants had to fill out the French-translated "Delphi definitions of Lower-Back Pain Prevalence" (DOLBaPP) – questionnaire [14] to ensure that they were not suffering from lower-back or lower quadrant pain. Other exclusion criteria were the following: diabetes, endocrine dysfunction, cognitive disorders, spinal pain, neurological or rheumatologic disorders and known pregnancy. The first 30 volunteers meeting all inclusion criteria were included in the study (non-probabilistic voluntary sample). Subjects all confirmed not being on any pain medication. All volunteers signed a written informed consent form. The study was approved by the local Ethics Committee, and consistent with the Declaration of Helsinki.

2.1. Experimental procedure

Tests were conducted in a quiet room at the CHUV, with an ambient temperature between 20 and 22 °C. Subjects were assessed on two occasions at the same time of day, within a one-week interval. Tests were performed bilaterally on the dorsum of both feet and on the lumbar paravertebral area (L5 dermatome). The stimulation sites were randomized (Microsoft Excel 2008, version 12.3.6) for each participant and each measurement session, first to define the site (lumbar or foot) and then the side (left or right). Prior to both sessions, a demonstration of the procedure was performed on the left hand. Tests were done using a Neuro Sensory Analyzer TSA-II (Medoc, Israel). A Peltier thermode (16 × 16 mm) was attached directly on the skin of the tested area. CDT and WDT were measured first. The number of PHS was then determined during the thermal sensory limen (TSL) procedure of alternating cold and warm stimuli. CPT and HPT were finally recorded. The baseline temperature was set at 32 °C and increased or decreased at a rate of 1 °C/s. To prevent thermal injury, cut-off levels were set at 50 °C and 0 °C. Standardized instructions were read out to each subject before testing [8]. Tests were conducted under the same conditions as in clinical practice.

All measurements were performed in the same protocolled way in both sessions, by the same trained observer (physiotherapist), using the same equipment. Prior to the beginning of the study, the investigator used exactly the same protocol as that used for participants on 10 persons who were not included in the study. Volunteers were blinded to their own prior results. The Quality Appraisal for Reliability Studies (QAREL)-Checklist [15] was used to guarantee optimal methodological rigour. Guidelines for reporting reliability and agreement studies (GRRAS) in the medical field have been followed [16].

2.2. Statistical analysis

All statistical calculations were performed using SPSS v.21 (IBM, SPSS, Inc., Chicago, IL, USA), except *t*-test statistics that were done using the internet-based statistical software Simple Interactive Statistical Analysis (SISA) URL: http://www.quantitativeskills.com/sisa/ (accessed 2014, May 10), as proposed by Magerl and colleagues [17] and Bland–Altman plots that were realized with MedCalc Software v.13.2 (Ostend, Belgium). Reliability of measures was assessed with the intraclass correlation coefficient for average measurement ($ICC_{agreement}$) with 95% confidence interval (CI) for each modality. Sample size was estimated as follows: with a sample size of 30 and an expected ICC of 0.85, the lower boundary of the 95% CI would still be above 0.7 for at least 80% of all ICC calculated. This is accepted as a sufficiently high level of reliability [18]. Kolmogorov–Smirnov test was used to assess normality of data distribution.

2.2.1. Test–retest reliability

Means and standard deviations (SD) of three consecutive measurements were calculated for CDT, WDT, CPT and HPT for each tested area in both sessions. In addition, mean intra-individual differences (MID) were calculated for each pair of data sets. PHS was made to follow the DFNS protocol [8] but was not further developed in this study. Calculations of relative reliability were done using a two-way random effects analysis of variance (ANOVA) model, with absolute agreement. Obtained ICC can range from 0 (no correlation) to 1 (perfect correlation). Strength of the correlation was interpreted as follows: ICC < 0.40 is considered as poor, 0.40–0.59 as fair, 0.60–0.75 as good and >0.75 as excellent [19]. Bland–Altman plots were created to determine absolute reliability.

2.2.2. Comparison with normative data

Data obtained from the lower extremity were compared to corresponding normative data for the feet (L5 dermatome) from the DFNS [7] using a *Z*-transformation for each parameter.