Reliability and validity of quantitative sensory testing in persons with spinal cord injury and neuropathic pain

Elizabeth R. Felix, PhD;1–2 Eva G. Widerström-Noga, DDS, PhD1–4*

1Research Service, Department of Veterans Affairs Medical Center, Miami, FL; 2Department of Rehabilitation Medicine, 3The Miami Project to Cure Paralysis, and 4Department of Neurological Surgery, University of Miami Leonard M. Miller School of Medicine, Miami, FL

Abstract—Quantitative sensory testing (QST) has been used to assess neurological function in various chronic pain patient populations. In the present study, we investigated the ability of QST to reliably characterize somatosensory dysfunction in subjects with spinal cord injury (SCI) and neuropathic pain by measuring mechanical, vibration, and thermal detection and pain thresholds. Test-retest reliability was determined based on data collected from 10 subjects with SCI and neuropathic pain who underwent QST on two occasions approximately 3 weeks apart. The intraclass correlation coefficients for mechanical, vibration, warm, and cool detection thresholds were in the “substantial” range, while thresholds for cold pain and hot pain demonstrated “fair” stability in this sample of patients. To determine the validity of QST in persons with SCI-related neuropathic pain, we evaluated the relationship between somatosensory thresholds and severity of neuropathic pain symptoms with multiple linear regression analysis. Thermal pain threshold was the only QST variable significantly related to the severity of neuropathic pain symptoms. The present study provides preliminary evidence that QST is a reliable and valid adjunct measurement strategy for quantifying the neurological dysfunction associated with neuropathic pain in persons with SCI.

Key words: intractable pain, neuropathic pain, pain, pain measurement, pain threshold, psychophysics, quantitative sensory testing, rehabilitation, sensory thresholds, spinal cord injuries.

INTRODUCTION

Approximately 70 percent of persons with spinal cord injury (SCI) develop chronic pain after their injury [1–5]. Although many types of pain may be present in the same individual with SCI [6–8], pains with neuropathic-like characteristics are particularly refractory to treatment [9–10], partly because the precise underlying mechanisms responsible for this type of pain in persons with SCI are still unknown.

One potentially promising method for assessing the mechanisms that contribute to the development and/or maintenance of neuropathic pain after SCI is the use of quantitative sensory testing (QST) to evaluate the hypo- and hypersensitivity of the somatosensory system [11]. QST has been used extensively to assess the functional integrity of the somatosensory system in a number of

Abbreviations: ASIA = American Spinal Injury Association, ATDT = average thermal detection threshold, ATPT = average thermal pain threshold, CDT = cool detection threshold, CPT = cold pain threshold, HPT = hot pain threshold, ICC = intraclass correlation coefficient, LOI = level of injury, MDT = mechanical detection threshold, NPSI = Neuropathic Pain Symptom Inventory, QST = quantitative sensory testing, SCI = spinal cord injury, SD = standard deviation, VA = Department of Veterans Affairs, VDT = vibration detection threshold, WDT = warm detection threshold.

*Address all correspondence to Eva G. Widerström-Noga, DDS, PhD; University of Miami, The Miami Project to Cure Paralysis, PO Box 016960 R-48, Miami, FL 33101; 305-243-7125; fax: 305-243-3921. Email: ewiderstrom-noga@med.miami.edu

DOI:10.1682/JRRD.2008.04.0058
patient populations, including persons with diabetic neuropathy [12–14], herpes zoster [15], complex regional pain syndrome [16–18], and SCI [19–30]. Assessment of tactile function, including measurement of thresholds for light pressure and vibration, can be used to evaluate large-fiber and dorsal column function, while measurement of thermal detection and pain thresholds can be used to assess the integrity of small-fiber and spinothalamic tract function [31]. Although QST has been promoted for the precise assessment of neurological dysfunction in some neuropathic pain populations [32–33], the use of QST protocols for diagnostic purposes or as outcome measures in persons with SCI and chronic pain has not been fully established [11,34].

An important feature of an outcome or diagnostic measure, for both research and clinical purposes, is its reliable reproduction under the same conditions. Several studies in nondisabled subjects have found sufficient test-retest reliability between sessions for mechanical detection thresholds (MDTs) [35], vibration detection thresholds (VDTs) [24,36–38], and thermal detection thresholds [24,35–36,38–39], but reliability of thermal pain threshold measurements has been lower [35,40]. In contrast to studies in healthy, nondisabled control subjects, research involving the reliability of QST in the SCI population has been sparse [22,24,29] and information regarding the pain status of study participants is rarely included [24,29]. Overall, these studies support adequate stability of QST in persons with SCI, with some exceptions noted for particular test sites [24] and for cold pain thresholds (CPTs) [29]. Furthermore, the validity of QST protocols as diagnostic and outcome measures in those who experience persistent neuropathic pain after injury relies heavily on the ability to link these measures of neurological dysfunction to symptoms of the neuropathic pain condition. Although research in other patient populations indicates that QST may be useful for differentiating neuropathic pain subtypes [12,33], previous studies in persons with SCI have not clearly linked measures of specific somatosensory dysfunctions to the presence or severity of neuropathic pain [20,22,26,28].

Currently, little conclusive evidence exists regarding the use of QST in persons who have SCI and neuropathic pain. Further research establishing the reliability and validity of QST in patients with SCI and pain is needed [11,34]. The present study (1) assessed the test-retest reliability of QST in persons with SCI and chronic neuropathic pain and (2) examined the validity of QST measurements as indicators of neuropathic pain in a sample of individuals with SCI.

METHODS

Spinal Cord Injury Participants

Individuals with SCI were recruited through advertisements posted at the Miami Department of Veterans Affairs (VA) Medical Center and the University of Miami medical campus, including The Miami Project to Cure Paralysis, and by word of mouth. The study was approved by the institutional review boards of the Miami VA Medical Center and the University of Miami.

Potential subjects were screened over the telephone or in person to confirm eligibility. Participants had to be over the age of 18, be fluent in English, have experienced a traumatic SCI at least 1 year before participation in the study, have an injury level above the first lumbar, and have neuropathic-like pain that had been present for at least the past 3 months and was rated as at least a 4 on a 0–10 numerical rating scale for average pain intensity.

Control Subjects

Nondisabled participants were recruited in a similar manner as the participants with SCI. They were screened to confirm that they had no current or recent pain and/or health problems, had no history that may have put them at risk for peripheral or central neuropathies, and were not regularly taking any prescription or over-the-counter medications other than on an as-needed basis.

General Protocol

Subjects with SCI who met the inclusion criteria were scheduled for their first study visit. After informed consent was obtained, a neurological examination was conducted and a second visit was scheduled. During the second visit, a battery of questionnaires was administered in an interview format and QST was conducted. Of the 22 participants with SCI and neuropathic pain, 12 were part of a clinical trial, and therefore, only their baseline values before starting treatment were used. The remaining 10 subjects completed an identical test session approximately 1 to 4 weeks later to provide data for the test-retest analysis portion of the present article.

Nondisabled control subjects completed two visits. During the first visit, eligibility was confirmed, informed consent was obtained, demographic and health history
questionnaires were completed, and QST was performed. Nondisabled subjects were tested with the same questionnaires and QST protocol during a second session to examine data for test-retest reliability. All participants were paid $50 for the completion of each session.

**Demographic and Injury Characteristics**

Each participant’s age, sex, recent medical history, living situation, and racial/ethnic background were recorded as part of a structured interview. For subjects with SCI, additional questions regarding the cause of injury and time since injury were included. For each participant with SCI, a physician with extensive SCI experience conducted a physical examination, including the American Spinal Injury Association (ASIA) standard examination [41–43], to assess neurological status and determine the severity (complete or incomplete) of injury. To determine the location of pain areas with respect to level of injury (LOI), if the LOI was different for the left and right sides and/or for the motor and sensory examinations, the overall LOI was taken as the most rostral level determined by the ASIA Impairment scale.

**Pain History Interview**

Participants with SCI were interviewed in a quiet, private room with a questionnaire to obtain details regarding pain in SCI [8,44–46]. Participants were first asked to indicate where they were currently experiencing chronic pain by shading in the areas on a drawing of the dorsal and frontal views of the human body; if physically unable to complete the drawing, participants described the locations to the interviewer, who shaded in the areas described. If more than one location of pain existed, the participant was asked whether the pains located in different areas were distinguishable from one another; if so, the participant was asked to answer each of the remaining questions separately for each different pain.

A number of other questions were included in the pain history interview in order to identify the most probable etiology of each subject’s pains (i.e., neuropathic or nociceptive): subjects were asked to choose from a list the words that described the quality of their pain, indicate the temporal constancy of their pain (e.g., intermittent or constant), and choose the degree of aggravation or relief a number of factors or situations had on their pain. Neuropathic pain was defined as being located at and/or below the LOI; being described as “sharp,” “shooting,” “burning,” “stabbing,” and/or “electric” [47–48]; and not having characteristics primarily associated with nociceptive pains (e.g., exacerbation due to movement, responsiveness to nonsteroidal anti-inflammatory drugs).

**Neuropathic Pain Symptom Inventory**

In the present study, we used the Neuropathic Pain Symptom Inventory (NPSI) [48] as a comparison variable to examine the validity of using somatosensory thresholds as an assessment of neuropathic pain characteristics. The NPSI is a self-report questionnaire specifically designed to measure the quality and severity of neuropathic pain. Validity and reliability of the NPSI have been established in patients with neuropathic pain of both peripheral and central origin [48–49]. The NPSI total intensity score is calculated from answers to 10 questions regarding the severity of common neuropathic pain qualities (e.g., burning, pressure, squeezing, electric shocks, stabbing, tingling, and pins and needles) and of pain evoked by brushing, pressure, and cold. We used the NPSI total intensity score as a measure of neuropathic pain severity.

**Quantitative Sensory Testing**

QST was used to examine the functional integrity of somatosensory pathways in all participants. MDTs and VDTs were measured to assess dorsal column function, and thermal detection thresholds (cool and warm) and thermal pain thresholds (cold pain and hot pain) were measured to assess spinothalamic tract function [31].

**Protocol**

All QST was performed in a quiet room with an approximate temperature between 21 °C and 23 °C. Control subjects were seated in a comfortable chair with armrests and a semireclining back. Subjects with SCI and chronic pain were tested in their own wheelchair. Test sites were identified based on anatomical landmarks to ensure that the same site could be accurately located for the repeat session. Room temperature, time of testing, and momentary rating of pain intensity (on a 0–10 numerical rating scale, with 0 = “no pain” and 10 = “most intense pain imaginable”) were recorded just before testing.

A standard set of instructions with an overview of the testing procedures was then read to the subject. For each different modality, specific instructions were read just before beginning the test. Measurement of a particular type of threshold was first demonstrated, and at least two practice trials were conducted on the subject’s left cheek.
After practice trials were completed, data collection began for each test modality. Measurements of MDT were recorded at each test site first, followed by measurements of VDT and thermal thresholds (cool detection threshold [CDT], warm detection threshold [WDT], CPT, and hot pain threshold [HPT]). Skin surface temperature was measured at each site with a Raytek MiniTemp non-contact thermometer (Raytek Corporation; Santa Cruz, California) just before thermal threshold measurements commenced. Vibrotactile and thermal threshold measurements were obtained with the TSA-II Neurosensory Analyzer and accompanying software (Medoc Ltd; Ramat Yishai, Israel).

For each stimulus type (monofilaments, vibration, thermal), the first measurements were taken at the subject’s right cheek and testing progressed to more caudal sites. By proceeding in this manner, we could ask subjects with SCI to compare the quality of the sensation evoked by each test stimulus in areas at and below the LOI to the quality evoked at the cheek, an area above the LOI where sensation was expected to be within normal limits. For measures obtained with the TSA-II (vibration, thermal detection, and thermal pain thresholds), when subjects verbally responded that threshold was reached, the experimenter immediately pressed a button to record the threshold and stop the stimulus trial. Although this method of response causes an increase in reaction time and, thereby, slightly higher threshold measurements, it was necessary to accommodate participants with upper-limb dysfunction and was used across both subject groups (SCI and nondisabled) to ensure consistency.

**Test Sites**

Two standard sites in every subject with SCI and neuropathic pain were tested: (1) a site located above-level for all subjects (right cheek) and (2) a site located below-level for all subjects (right medial calf). Other test sites were selected based on each individual’s LOI and pain distribution so as to include sites where chronic neuropathic pain was present and sites where neuropathic pain symptoms were absent both at and below the LOI. “At” the LOI was defined as a band of dermatomes including the dermatome of the neurological LOI defined by the ASIA examination and three dermatomes below this level, and “below” the LOI was defined as areas at least four dermatomes below the neurological LOI [50]. Because SCI subjects in the present study had lesion levels from cervical 4 to thoracic 10 and because the distribution of neuropathic pain areas in these subjects varied widely, standardization of the anatomical location of most test sites across all subjects was impossible.

Eight standard body sites were chosen for testing in our sample of healthy, nondisabled controls subjects (Figure 1). As far as possible, these standard sites were selected to match the majority of test sites in subjects with SCI and to be distributed no more than five dermatomes from another standard test site.

To examine the validity of QST data from SCI subjects with neuropathic pain, we attempted to match each test site in an individual with SCI to one of the standard test sites in control subjects that was located in the same or closest dermatome no more than two dermatomes away. Because previous studies have shown that somatosensory thresholds obtained bilaterally do not differ with regards...
to body side in healthy participants [33,37], we chose to place all test sites in control subjects on the right side of the body.

**Mechanical Detection Threshold**

A standard set of Semmes-Weinstein monofilaments (Touch Test™ Sensory Evaluator, North Coast Medical, Inc; Morgan Hill, California) was used to measure MDTs. The set consists of 20 graded monofilaments, with a range of target forces between 0.008 and 300 g. Each subject was instructed to close his or her eyes during this portion of the testing and to respond with a “yes” if he or she could feel the test stimulus when it was delivered or with a “no” if he or she could not feel the stimulus. For each trial, the monofilament was applied perpendicular to the skin surface and, once the filament was fully bent, was held in place for approximately 1 s before being lifted off the skin. Movement of hair follicles during stimulus presentation was avoided if possible.

Four stimulus series were performed at each site according to the method of limits. For each of the two descending series, an average was calculated using the force for the last monofilament that was detected and the force for the first monofilament that was not detected [51]. For each of the two ascending series, an average was calculated using the force for the last monofilament that was not detected and the force for the first monofilament that was detected [51]. “Catch trials” were performed periodically at each site by making the motion of applying a monofilament but without actually touching the skin in order to document the participant’s bias. Threshold values obtained at test sites in which a subject reported a sensation during a catch trial were eliminated from analysis. On series during which the lowest force was detected (0.008 g), this value was taken as threshold for that series; on series in which the highest force (300 g) was not detected, this ceiling value was recorded as threshold. Threshold values for MDT at each test site were defined as the arithmetic mean of the values obtained during the four stimulus series.

**Vibration Detection Threshold**

The handheld VSA-3000 component of the Medoc system was used to measure VDTs for a 100 Hz stimulus frequency. The circular contact tip (1.22 cm²) was held in place by the experimenter during testing so that there was a slight and maintained indentation of approximately 1 to 2 mm. Stimulus presentation was programmed with the software accompanying the vibratory equipment to control the rise rate of stimulus amplitude, the number of trials, and the time between each trial. Three trials, separated by approximately 10 s each, were conducted using the ascending method of limits: vibratory amplitude began at 0 μm and increased until the subject indicated that the stimulus was felt or until the maximum amplitude of 130 μm was reached. During most VDT trials, the vibratory amplitude increased at a rate of 0.5 μm/s. For test sites on which there was very little or no vibratory sensation, the rate of increase was changed to 5.0 μm/s to avoid unusually long test trials (>2 min). Subjects were asked to indicate the “first moment” that they felt the vibration at the test site. The mean value across the three trials was recorded as the VDT for that site.

**Cool and Warm Detection Thresholds**

The method of limits was used to obtain measures of CDTs and WDTs. A 1.6 × 1.6 cm thermode connected to the TSA-II Neurosensory Analyzer was used to deliver thermal stimuli. The experimenter held the thermode firmly against the skin with light pressure during all thermal testing procedures.

Each trial began with the thermode temperature set at 32 °C. Once the trial began, the temperature increased (for CDT) or decreased (for WDT) at a rate of 1 °C/s until the subject perceived the stimulus or until the stimulus reached the cutoff value (0 °C for CDT and 50 °C for WDT). If no change was detected, the cutoff value for that stimulus modality was recorded as the threshold. Each trial was separated by approximately 10 s. After four CDT trials were completed, four WDT trials were conducted in the same way. The average of the change in temperature needed to evoke the appropriate sensation (cool or warm) across the four trials at each test site was recorded as threshold for that modality.

**Cold and Hot Pain Thresholds**

CPTs and HPTs were obtained using the same equipment and in a similar manner as the CDTs and WDTs. Subjects were read a standard set of instructions that informed them to indicate as soon as the sensation changed from “just being cold to being painfully cold” or from “just being hot to being painfully hot.” Each trial began at 32 °C and was either decreased (CPT) or increased (HPT) at a rate of 1.5 °C/s until pain threshold was reached or the cutoff value was reached (0 °C for CPT and 50 °C for HPT). Each trial was separated from
the next by at least 20 s. The arithmetic mean across three trials at each test site was calculated as the threshold.

**Quantitative Sensory Testing Data Processing**

To examine the function of the dorsal column and the spinothalamic tracts, we included in the data analysis only those threshold measurements that accurately reflected the appropriate stimulus modality. Therefore, a measurement taken during any trial on which the subject reported feeling a sensation that was qualitatively different from the specific modality being tested (e.g., reporting an “electrical” pain sensation during VDT or reporting only a “muscle contraction” sensation during thermal testing) was eliminated from the present analyses. If pain was the first sensation reported during a CDT or WDT trial and no cold or warm sensation preceded the pain, the threshold measurement was interpreted as pain threshold (i.e., CPT or HPT). To include data for analyses at sites where no sensation was evoked during testing, we recorded the maximum amplitude of that stimulus modality.

Data from the right cheek, the above-level standard test site in all SCI subjects, were analyzed to determine whether transformation of threshold data was needed to approximate the normal distribution. The skewness, kurtosis, and Kolmogorov-Smirnov *d* statistics were calculated for raw data and log-transformed data for thresholds obtained for each stimulus modality. Using Rolke et al.’s method [52], we determined the geometric mean of skewness and kurtosis and multiplied this value by the Kolmogorov-Smirnov *d* for each distribution as a measure of goodness of fit to the normal distribution. If the ratio for the raw data to the log-transformed data exceeded 3, then the log-transformed data were considered to be a better approximation of a normal distribution [52] and were used in all further analyses for that stimulus modality.

Test-retest reliability, a measure of the stability of a test when it is administered across time without changes in other variables, was evaluated separately for SCI subjects and nondisabled subjects for each QST modality by using intraclass correlation coefficients (ICCs) (one-way random effects model) [53]. The assessment of the level of reliability was based on Shrout’s recommendations [54]: an ICC between 0.41 and 0.60 is considered “fair,” an ICC between 0.61 and 0.80 is considered “moderate,” and an ICC between 0.81 and 1.00 is considered “substantial.”

For analysis of the relationship between QST measures and NPSI scores, only the results from SCI subjects were used. However, to compare across SCI subjects and test sites, we used the mean and standard deviation (SD) values of the thresholds obtained in the healthy, nondisabled control subjects to calculate *z*-scores for the threshold measures obtained in the SCI neuropathic pain group. Thus, we were able to limit the effect of variance in threshold measures due to body location [39–40,55–56] and to examine relationships between NPSI scores and QST measures obtained across all SCI participants and test sites. The *z*-scores were calculated by obtaining the difference between the SCI patient’s threshold at a particular test site and the mean of the nondisabled control subjects’ thresholds at a comparable test site and then dividing by the SD of the control subjects’ data at that site.

Although we attempted to match all SCI test sites with control sites, not every test site in all SCI subjects had a suitable control comparison site because site selection in persons with SCI depended on both LOI and location of clinical pain symptoms. If a test site in an SCI participant was located more than two dermatomes away from any comparison site or the SCI test site was in a noncomparable area for the closest dermatomal match to a nondisabled control site (i.e., hairy vs glabrous skin sites, dorsal vs ventral body sites), the data for that site were not included in statistical analyses that used *z*-score calculations.

Independent samples *t*-tests were conducted to examine differences between *z*-scores for thresholds obtained in pain areas and *z*-scores for thresholds obtained in nonpain areas for each stimulus modality.

Before conducting analyses regarding the validity of our QST measures, we calculated Pearson correlation coefficients among thresholds obtained for each stimulus modality to assess potential collinearity between these measures. For any pair of thresholds that had a correlation coefficient exceeding 0.70, the threshold measures for the two modalities were combined by calculating the average of the *z*-scores of the threshold values obtained for these two modalities. These averaged values were used in all further statistical analyses.
Because our sample size was relatively small and therefore a large number of predictor variables in the multiple regression analysis would increase the likelihood of a type I error, we wanted to include only those sensory modalities whose thresholds were significantly related to NPSI score based on bivariate correlations. Since differences have been found between QST measures obtained in areas where chronic pain is manifest and areas where it is absent in other patient populations [16,57], we performed two sets of correlation analyses between each type of threshold and NPSI scores: the first set included data from test sites that were in neuropathic pain areas and the second set included data collected in test sites where chronic pain was not present.

Multiple linear regression analysis was used to examine the relationship between z-score transformed thresholds and severity of neuropathic pain (NPSI total intensity score) and the potential moderators of this relationship (location of test site relative to injury, complete vs incomplete injury). Threshold values determined by bivariate correlations to be related to NPSI scores and interaction terms needed to evaluate possible confounding or moderating effects (threshold × location of site relative to injury, threshold × “completeness” of injury) were entered into a stepwise regression.

All statistical tests were performed with SPSS 14.0 for Windows (SPSS, Inc; Chicago, Illinois), and statistical significance was set at $\alpha = 0.05$.

RESULTS

Participants

Twenty-two individuals with SCI and neuropathic pain and ten nondisabled control subjects participated in the study. Demographic information for these two groups of participants is listed in Table 1, along with characteristics of injury for the subjects with SCI. Data collected during two identical test sessions for 10 of the SCI subjects and all 10 of the nondisabled subjects were used in test-retest reliability analyses.

Distributions of Quantitative Sensory Testing Data

The distribution properties of quantitative threshold measures for each stimulus modality were assessed to determine whether transformation was needed to better approximate a normal distribution using the methods described by Rolke et al. [52]. Table 2 presents the skewness, kurtosis, and Kolmogorov-Smirnov $d$ statistics for the raw thresholds and for the log-transformed threshold values. On the basis of a comparison of these parameters, as described in the “Methods” section, we used the log-transformed data for the MDT, VDT, CDT, and WDT measures and the raw data for the CPT and HPT measures when performing all further statistical tests.

Test-Retest Reliability

Subjects in the reliability portion of the study completed two identical test sessions with approximately 1 to 4 weeks between each session (mean SCI between-session interval = 22.2 days, mean nondisabled between-session interval = 17.2 days). The ICCs for each test modality are presented in Table 3. For subjects with SCI, threshold measures for MDT, VDT, CDT, and WDT showed substantial reliability, with ICCs ranging from 0.84 to 0.95, while threshold measures for CPT and HPT were somewhat less reliable (ICCs = 0.50). For the nondisabled control sample, only the ICC for VDT fell into the substantial category (0.86), while MDT, CDT, WDT, and HPT all demonstrated moderate reliability (0.63–0.70), and CPT measurements had fair reliability (0.49).

Validity

Multicollinearity was detected between $z$-scores for CDT and WDT ($r = 0.73$) and between CPT and HPT ($r = 0.71$) in persons with SCI and neuropathic pain. Therefore, average values for these pairs were calculated to create an average thermal detection threshold (ATDT) $z$-score and an average thermal pain threshold (ATPT) $z$-score to be used in further analyses. Independent samples t-tests performed to compare $z$-scores for thresholds obtained in pain sites and thresholds obtained in nonpain sites did not reveal any significant group differences for any of the stimulus modalities tested (MDT, VDT, ATDT, ATPT).

Relationships between $z$-scores for somatosensory thresholds and severity of neuropathic pain (NPSI total intensity score) are displayed in Figure 2. Threshold data for test sites located in neuropathic pain (open symbols in Figure 2) and for test sites located in areas exhibiting neuropathic-like pain symptoms (filled symbols in Figure 2) were assessed separately with regards to their relationship with NPSI total intensity scores. NPSI scores and ATPT values obtained within painful test sites were significantly correlated ($r = 0.58$, $p < 0.02$). No other
measure of somatosensory function was significantly related to NPSI scores.

To further investigate the relationship between ATPT and NPSI scores within sites where neuropathic pain was present, we performed a stepwise multiple regression analysis \((n = 17)\) with NPSI total intensity score as the dependent variable and ATPT, completeness of injury, location of test site relative to LOI, and appropriate interaction terms \((\text{ATPT} \times \text{completeness of injury}, \text{ATPT} \times \text{location of test site relative to LOI})\) as independent variables. Adding these interaction terms into the regression allowed for the assessment of the potential moderating effects of completeness of injury and location of test site relative to LOI on the relationship between thermal pain thresholds and neuropathic pain severity. Only ATPT \(z\)-scores were significantly related to NPSI score \((R^2 = 0.331, p = 0.016)\). None of the other factors, or interactions of those factors with ATPTs, significantly added to the model. Regression results are shown in Table 4.

### DISCUSSION

The present study determined the use of QST measures as potential diagnostic and outcome variables for SCI clinical pain trials. Specifically, the primary aims of the present study were to determine the test-retest reliability of somatosensory threshold measurements and to examine the validity of QST by evaluating the strength of the relationship between sensory thresholds and the severity of neuropathic pain. Although the sample size was small, this study provides preliminary support for the reliability and validity of this methodology in persons with SCI and neuropathic pain.

### Table 1.

Demographic and injury characteristics of participants with spinal cord injury (SCI) and nondisabled control subjects.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>SCI ((n = 22))</th>
<th>Control ((n = 10))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr), Mean ± SD</td>
<td>41.7 ± 15.5</td>
<td>30.4 ± 4.3</td>
</tr>
<tr>
<td>Sex, (n) (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>3 (13.6)</td>
<td>4 (40.0)</td>
</tr>
<tr>
<td>Male</td>
<td>19 (86.4)</td>
<td>6 (60.0)</td>
</tr>
<tr>
<td>Race/Ethnicity, (n) (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White Non-Hispanic</td>
<td>10 (45.5)</td>
<td>4 (40.0)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>10 (45.5)</td>
<td>3 (30.0)</td>
</tr>
<tr>
<td>African American</td>
<td>2 (9.1)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Other</td>
<td>0 (0.0)</td>
<td>3 (30.0)</td>
</tr>
<tr>
<td>Time Since Injury (yr), Mean ± SD</td>
<td>6.6 ± 5.7</td>
<td>—</td>
</tr>
<tr>
<td>Level of Injury, (n) (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cervical</td>
<td>12 (54.5)</td>
<td>—</td>
</tr>
<tr>
<td>Below Cervical</td>
<td>10 (45.4)</td>
<td>—</td>
</tr>
<tr>
<td>Completeness of Injury, (n) (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incomplete</td>
<td>10 (45.4)</td>
<td>—</td>
</tr>
<tr>
<td>Complete</td>
<td>12 (54.5)</td>
<td>—</td>
</tr>
</tbody>
</table>

SD = standard deviation.

### Table 2.

Properties of threshold distributions for right cheek site in participants with spinal cord injury and neuropathic pain.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>K-S (d)</th>
<th>Weighted Ratio (raw/log)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw MDT</td>
<td>2.96</td>
<td>9.07</td>
<td>0.41</td>
<td>3.32</td>
</tr>
<tr>
<td>Log MDT</td>
<td>1.82</td>
<td>4.03</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Raw VDT</td>
<td>0.60</td>
<td>0.45</td>
<td>0.14</td>
<td>3.55</td>
</tr>
<tr>
<td>Log VDT</td>
<td>–0.723</td>
<td>0.011</td>
<td>0.225</td>
<td></td>
</tr>
<tr>
<td>Raw CDT</td>
<td>2.19</td>
<td>4.45</td>
<td>0.30</td>
<td>3.46</td>
</tr>
<tr>
<td>Log CDT</td>
<td>1.28</td>
<td>1.36</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Raw WDT</td>
<td>1.77</td>
<td>4.75</td>
<td>0.159</td>
<td>5.36</td>
</tr>
<tr>
<td>Log WDT</td>
<td>0.74</td>
<td>1.05</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>Raw CPT</td>
<td>–0.15</td>
<td>–1.53</td>
<td>0.20</td>
<td>0.51</td>
</tr>
<tr>
<td>Log CPT</td>
<td>–0.64</td>
<td>–0.92</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Raw HPT</td>
<td>0.25</td>
<td>–1.02</td>
<td>0.10</td>
<td>0.93</td>
</tr>
<tr>
<td>Log HPT</td>
<td>–0.245</td>
<td>–1.05</td>
<td>0.10</td>
<td></td>
</tr>
</tbody>
</table>

CDT = cool detection threshold, CPT = cold pain threshold, HPT = hot pain threshold, K-S = Kolmogorov-Smirnov, MDT = mechanical detection threshold, VDT = vibration detection threshold, WDT = warm detection threshold.
QST has been used successfully in other patient populations to aid in diagnosis [12–18] and also has a long history of use in healthy, nondisabled control subjects to assess mechanisms of somatosensory processing, including the processing of nociceptive information. Threshold measures for MDT, VDT, CDT, and WDT in persons with SCI and neuropathic pain were normally distributed only after logarithmic transformation, while thermal pain thresholds (CPT and HPT) needed no transformation. Rolke et al. found similar distribution properties in healthy, nondisabled subjects, with the exception of VDT, which was measured using a different technique than in the present study [52].

Reliability

In our sample of nondisabled control subjects, we found moderate to substantial levels of reliability for MDT, VDT, CDT, WDT, and HPT (0.63–0.86) but fair reliability for CPT (0.49). These results are consistent with a number of studies in healthy, nondisabled subjects and in various patient populations other than SCI that showed similar reliability results [35–38,58], although some other studies have reported unacceptable variation from session to session [39–40,59]. Similar to the findings of the present study, reports have suggested that thresholds for thermal pain, and for cold pain in particular, are more variable across sessions than other QST measures [29,35].

In our sample of individuals with SCI and neuropathic pain, the test-retest reliability of threshold measures for mechanical, vibratory, cool, and warm detection showed substantial reliability (0.84–0.95) and measures of cold pain and hot pain showed fair reliability (0.50) across a number of test sites. These results suggest that the degree of reliability of QST in persons with SCI and chronic pain is similar to that seen in healthy, nondisabled subjects and in other patient populations. A study by Defrin et al. examined thresholds for warmth, cold, and heat pain across three test sessions in a sample of SCI patients with chronic pain and a sample of SCI patients without chronic pain [22]. Threshold measures were reported to be consistent across the sessions: no significant differences between the ranges of values across sessions were found using the Student t-test. Although interpreting test-retest reliability is difficult using t-test analysis techniques, Defrin et al.’s study results [22] agree with our results, suggesting reasonable reliability between sessions in persons with SCI and chronic pain.

Two other studies have remarked on the stability of threshold measures obtained in persons with SCI [24,29]. However, the chronic pain status of the participants in these studies was not reported. Krassioukov et al. found that ICCs in SCI patients for CPT and VDT were in the acceptable range at all test sites (0.65–0.90), ICCs for CDT and WDT were acceptable in some test sites (0.55–0.81), but ICCs for CDT and WDT in other sites (one of six sites for CDT, and four of six sites for WDT) were low (0.25–0.46) [24]. A later article using similar QST methods, although not reporting specific statistical values for test-retest reliability, noted that “mean values obtained from the QST testing patients on the 2 days of testing showed considerable variation from day to day . . . . Most unreliable were the measures of cold pain perception” [29, p. 1614]. Although several other studies have used QST to compare thresholds in subjects with SCI and chronic pain and subjects with SCI and no pain, none of these has examined the stability of QST across test sessions.

### Table 3.

Test-retest reliability for participants with spinal cord injury (SCI) and neuropathic pain and for nondisabled control subjects.

<table>
<thead>
<tr>
<th>Modality</th>
<th>SCI ICC (95% CI)</th>
<th>SCI No. Test Sites</th>
<th>Nondisabled ICC (95% CI)</th>
<th>Nondisabled No. Test Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log MDT</td>
<td>0.84 (0.75–0.90)</td>
<td>56</td>
<td>0.63 (0.45–0.76)</td>
<td>58</td>
</tr>
<tr>
<td>Log VDT</td>
<td>0.90 (0.84–0.94)</td>
<td>57</td>
<td>0.86 (0.79–0.91)</td>
<td>80</td>
</tr>
<tr>
<td>Log CDT</td>
<td>0.90 (0.83–0.94)</td>
<td>55</td>
<td>0.68 (0.54–0.78)</td>
<td>80</td>
</tr>
<tr>
<td>Log WDT</td>
<td>0.95 (0.91–0.97)</td>
<td>50</td>
<td>0.70 (0.57–0.80)</td>
<td>79</td>
</tr>
<tr>
<td>CPT</td>
<td>0.50 (0.28–0.67)</td>
<td>56</td>
<td>0.49 (0.31–0.64)</td>
<td>80</td>
</tr>
<tr>
<td>HPT</td>
<td>0.50 (0.28–0.66)</td>
<td>62</td>
<td>0.68 (0.55–0.79)</td>
<td>79</td>
</tr>
</tbody>
</table>

CDT = cool detection threshold, CI = confidence interval, CPT = cold pain threshold, HPT = hot pain threshold, ICC = intraclass correlation coefficient, MDT = mechanical detection threshold, VDT = vibration detection threshold, WDT = warm detection threshold.
Validity

In addition to an examination of the reliability of quantitative measures of sensory perception in persons with SCI and neuropathic pain, a preliminary analysis of the validity of QST procedures as diagnostic and outcome measures in SCI pain trials was also examined. Understanding the relationship between QST measures and clinical pain components is imperative if these measures are to be used as tools to diagnose specific pain mechanisms or as clinical trial outcome measures.

The present study used the distribution of threshold values from a number of test sites in healthy, nondisabled control subjects to calculate $z$-scores for thresholds obtained in SCI patients with chronic pain of neuropathic origin, as suggested by the German Research Network on Neuropathic Pain [33,52]. In this way, we were able to combine data across test sites, SCI subjects, and test modalities in order to determine the strength of relationship between each of the QST measures and the severity of neuropathic pain symptoms (NPSI total intensity score) in this subject population. We found that the $z$-scores obtained for the two measures of thermal detection (CDT and WDT) and the two measures of thermal pain (CPT and HPT) were significantly correlated with one another (0.73 and 0.71, respectively). Hayes et al. also reported high correlations (0.57–0.84) between WDT and CDT obtained in sites below the LOI in subjects with SCI [29]. The magnitude and significance of the correlations between these measures suggests that the same quantitative measures of sensory function used to determine the integrity of the spinothalamic tract system in healthy, nondisabled subjects and in persons in other patient populations may also be used in persons with SCI to reflect similar mechanisms.

Results from the multiple linear regression analysis in the present study suggest that thermal pain thresholds may be particularly useful as adjunct measures to self-reported symptoms of neuropathic-like pain. Specifically, we found that lower average $z$-scores for thermal pain thresholds (CPTs and HPTs) in areas where clinical pain was present were significantly related to higher neuropathic pain severity (NPSI total intensity score), regardless of the location of the site relative to injury (at vs below) or the severity of injury (complete vs incomplete). In contrast, no significant correlations were found between NPSI scores and any of the threshold measures obtained in pain-free test sites. The results from this analysis should be viewed with caution, however, as the data available for this analysis (thermal pain threshold measures in painful areas at or below the LOI) were small ($n = 17$). Despite the small number of available data points, a significant association was found between average thermal pain

![Figure 2](image-url)

Neuropathic Pain Symptom Inventory (NPSI) scores and threshold values (reference: matched control site) for neuropathic pain (●) and nonpain (○) test sites in participants with spinal cord injury and neuropathic pain. (a) Mechanical detection threshold. (b) Vibration detection threshold. (c) Average thermal detection threshold. (d) Average thermal pain threshold. $r = $ Pearson correlation coefficient for average thermal pain threshold in neuropathic pain sites. *$p < 0.05$. 

Figure 2.
thresholds and the severity of neuropathic pain symptoms. Although most other variables included in the regression analysis (completeness of injury, test site relative to the LOI, and the interaction terms) displayed non-significant relationships with the dependent variable, the lack of a mediating effect of these variables is inconclusive as a result of the low power.

To our knowledge, no other study in SCI patients with neuropathic pain has reported a significant relationship between the severity of neuropathic pain symptoms and thermal pain thresholds at and below the LOI. However, Song et al. did find a significant inverse relationship between ratings of clinical pain severity and tactile two-point discrimination thresholds in SCI patients with dysesthetic pain [60]. The results from the present study found a relationship between functioning of the spinothalamic tract system and clinical pain severity, while those of Song et al. suggest that neuropathic pain severity may be related to the integrity of the dorsal column system (as measured by two-point thresholds) [60]. Despite these differences, both the current study and Song et al.'s study [60] found significant relationships between increasing levels of clinical pain and decreased somatosensory thresholds. In addition, Attal et al.'s recent study in a diverse group of individuals with peripheral or central neuropathic pain (including some individuals with SCI) reported a similar relationship between chronic pain symptom severity and pain thresholds: the severity of pressure-evoked pain symptoms, measured as part of the NPSI, were negatively correlated with mechanical pain thresholds [49].

Most previous studies in persons with SCI and neuropathic pain have focused on detecting threshold differences between groups (i.e., SCI participants with neuropathic pain vs SCI participants without pain, pain areas vs nonpain areas) instead of examining the correlations between continuous variables. These studies have typically failed to find group differences [20,22,26,28], which is consistent with the absence of significant differences from t-tests comparing thresholds in pain areas versus nonpain areas in the present study. Use of dichotomous data (e.g., pain vs no pain) can substantially reduce statistical power compared with use of continuous data (e.g., ratings of the severity of pain) [61]. Therefore, the differences in the relationship between thermal pain thresholds and neuropathic pain severity in pain sites compared with their relationship in nonpain sites, as indicated by our regression analysis, suggest that analyses aiming to detect significant group differences between pain areas and nonpain areas may not be as informative as analyses examining correlations within these groups.

Based on the results of the present study, areas affected by severe neuropathic pain may have more intact functioning within the nociceptive system than areas with less severe pain. This finding is in seeming contrast to a report in the literature showing greater thermal impairments in pain areas than in nonpain areas [22]. However, the level of clinical pain severity in the patient sample in that study was not reported. Figure 2 indicates that subjects with low NPSI scores (i.e., with less severe neuropathic pain) may have more impaired function in painful areas than in nonpainful areas (as Defrin et al. reported [22]), but this difference between pain and nonpain areas changes as clinical pain severity increases. Maybe the subjects recruited in the Defrin et al. [22] study reflected a subgroup of patients with relatively low intensities of spontaneous pain, and in this subgroup, painful areas may demonstrate more spinothalamic tract dysfunction than do nonpainful areas.

Table 4.
Multiple regression analyses predicting Neuropathic Pain Symptom Inventory total intensity score for participants with spinal cord injury and neuropathic pain.

<table>
<thead>
<tr>
<th>Variables in Model</th>
<th>β</th>
<th>t-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>-5.191</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>ATPT z-Score</td>
<td>-0.575</td>
<td>-2.723</td>
<td>0.016</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables Not in Model</th>
<th>β</th>
<th>t-Value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete vs Incomplete</td>
<td>0.038</td>
<td>0.157</td>
<td>0.88</td>
</tr>
<tr>
<td>Test Site re: LOI</td>
<td>-0.154</td>
<td>-0.633</td>
<td>0.54</td>
</tr>
<tr>
<td>Interaction: (ATPT) × (Complete vs Incomplete)</td>
<td>0.017</td>
<td>0.060</td>
<td>0.95</td>
</tr>
<tr>
<td>Interaction: (ATPT) × (Test Site re: LOI)</td>
<td>-0.299</td>
<td>-0.829</td>
<td>0.42</td>
</tr>
</tbody>
</table>

ATPT = average thermal pain threshold, LOI = level of injury.
Recommendations

Results from our analyses present a number of suggestions for future studies. First, these results need to be replicated in a larger study to further detail the reliability of each QST measure at different test sites in persons with SCI and neuropathic-like pain. Subgroup analysis of reliability results obtained in persons with complete versus incomplete spinal lesions and in areas at the LOI versus those below the LOI would be of particular interest within a larger sample of patients. On the basis of the current literature, clinical trials with repeated assessments across several days during each test period (i.e., baseline, drug administration, and washout phases of a study) may be recommended to increase the reliability and accuracy of pain threshold measures. Reliability of thresholds for the innocuous stimulus modalities tested in this study (MDT, VDT, CDT, WDT) are acceptable, and therefore, measuring these thresholds during only one test session for each phase of a study may be adequate.

Second, collection of age- and sex-matched reference data at a number of test sites in healthy, nondisabled subjects specifically for comparisons in the SCI population is advisable. In the present study, we used data from 10 healthy, nondisabled control subjects to calculate z-scores for threshold measures obtained in our sample of patients with SCI and neuropathic pain. Differences in injury level and locations of spontaneous pain among subjects with SCI necessitate the comparison of a number of different test sites across subjects. Since many patients with SCI have bilateral locations of pain, contralateral control sites cannot be used in this group of subjects even though this method is preferred for accurate assessment of the severity of sensory dysfunction in other patient populations [16,33]. Establishing a database of somatosensory thresholds measured at a number of body sites would facilitate future studies regarding the underlying mechanisms of neuropathic pain in SCI.

Third, an important factor that may play a part in the validity of threshold measures in persons with SCI is the presence of “abnormal” sensory experiences. As such, careful questioning of subjects during QST is recommended and may prove to be imperative for teasing out these abnormal events. Analyses of threshold values for qualitatively abnormal sensory experiences may shed light on mechanisms underlying different aspects of somatosensory function in persons with SCI and chronic pain more so than values measured only for “correct” sensory modality experiences. In the present study, we deliberately chose to use only those threshold values in which the subject reported the correct type of sensory experience (when they were asked to compare the sensation felt at areas at and below the LOI to the sensation felt above the LOI on the cheek). In this way, we eliminated data for sensory experiences that may have resulted from neural activity possibly conducted via abnormal or alternate pathways arising from reorganization within the central nervous system after the SCI. Although qualitatively different, or abnormal, sensations have sometimes been reported in previous studies in subjects with SCI [22–23], how such data have been treated for analysis purposes is often unclear. The decision to include or exclude such data from analysis may influence the relationship between threshold measures and clinical pain assessments in this patient population.

This study has several limitations. Inherent in the pain conditions following an SCI is great heterogeneity, which makes interpretations of results more difficult. As has been suggested recently [62], grouping subjects with chronic pain into those with “probable” and those with “definite” neuropathic pain may help limit variance between subjects and may prove essential in uncovering the mechanisms responsible for these group differences. In addition, the small sample of healthy, nondisabled control subjects and the small number of comparison test sites available for adequate matches to sites in persons with SCI and neuropathic pain presumably added to the variability of these measures, which likely reduced the power available for statistical tests. Therefore, the results presented here must be viewed with caution until replicated in another sample.

CONCLUSIONS

Based on the results of the present study, quantitative measures of sensory function appear to provide a reliable and accurate assessment of neurological dysfunction that may be related to the development and maintenance of neuropathic pain in persons with SCI. Use of QST as a diagnostic and/or outcome measurement strategy may provide a valuable adjunct for the assessment of clinical pain and may help determine the underlying mechanisms responsible for specific pain types in SCI.
ACKNOWLEDGMENTS

We would like to acknowledge Ms. Yenisel Cruz-Almeida, Mr. Jim Adcock, and Ms. Letitia Fisher for their assistance with recruitment and data collection and Dr. Alberto Martinez-Arizala for conducting ASIA examinations.

This material was based on work supported by the VA Rehabilitation Research and Development Service (merit review grants B3070R and B5023R), the State of Florida, and The Miami Project to Cure Paralysis.

The authors have declared that no competing interests exist.

REFERENCES


20. Eide PK, Jørøen E, Steinehjem AE. Somatosensory findings in patients with spinal cord injury and central dysaesthesia

FELIX and WIDERSTROM-NOGA. Quantitative sensory testing in SCI and neuropathic pain


Submitted for publication April 25, 2008. Accepted in revised form July 24, 2008.