Evaluation of variations in sensory and pain threshold assessments by electrocutaneous stimulation

Irène Lund, RPT,1 Thomas Lundeborg, MD Prof,2 Jan Kowalski, Statistician,1 Louise Sandberg, RPT,2 Cecilia Norrbrink Budh, RPT,3 and Elisabeth Svensson, Statistician Prof,4

1Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, SE-171 77 Sweden
2Rehabilitation Medicine Clinic, Karolinska University Hospital, Stockholm, SE-171 76 Sweden
3Spinalis Rehab Station, Karolinska University Hospital, Stockholm, SE-171 76 Sweden
4Department of Statistics (ESI), Örebro University, Örebro, SE 701-81 Sweden

Assessed sensory and pain thresholds can change consequently to disturbances associated with ongoing pain. Such assessments could be an additional method in the daily clinical evaluation of perceived pain. To study the test−retest variability within-day and between-day of such procedures a newly developed instrument producing electrocutaneous stimulation, PainMatcher (PM), was used to assess the electrical sensory thresholds (EST) and pain thresholds (EPT) in healthy volunteers and in patients with pain. The produced data were considered ordinal and analyzed with rank-invariant statistics with properties of analyzing systematic disagreement, bias, and individual variations.

The percentage agreements within ±1PM value for EST were in the two groups of healthy volunteers and patients in pain 94% and 92%, and for EPT assessments 49% and 78%, respectively. The variability in the EST assessments is possibly explained by a slight bias while the individual variations were negligible between the two occasions. The assessed EPT were unbiased in both groups while individual variations were significant among the healthy volunteers but negligible among the patients in pain. The EST was found to be increased in pain patients compared to healthy volunteers, p < 0.03, and the EPT decreased in pain patients compared to healthy volunteers, p < 0.001.

The results in this study indicate stable and reliable assessments of EST and EPT except for a possible bias. The threshold assessment procedure followed in this study may be a valuable tool in the clinical evaluation of sensory and pain assessments in pain patients.

Introduction

Since pain is regarded as one of the most common reasons to consult physical therapy (Main and Watson, 1999; Scudds, Scudds, and Simmonds, 2001) and since pain is listed as a frequent symptom in groupings of generic client problems (Guide to Physical Therapy Practice, 1997), the assessment of a patient’s perceived pain is of great importance to evaluate the outcome of physical therapy and rehabilitation programs.

Pain is a multidimensional individual experience and its communication is based on
self-reports and often assessed on rating scales, questionnaires, or by methods derived from psychophysical concepts such as threshold assessments (Borg, 1993; Gracely, 1999; Gracely, Lota, Walter, and Dubner, 1988).

The assessment of thresholds has so far not been included in clinical routines but mostly been restricted to neurodiagnostic and experimental procedures (Chesterton et al, 2002; Dotson, 1997; Waling, Sundelin, Ahlgren, and Järvholm, 2000). Thresholds for sensory detection and pain are tested by applying different modes of stimuli, e.g. mechanical, thermal and electrical. The stimulus is commonly applied with continuously increasing intensity, known as the method of limits. The threshold refers to the level of stimulus strength at which the subject first perceives the stimulation at all (sensory detection) and as painful respectively (Gracely, Lota, Walter, and Dubner, 1988; Sang, Mitchell, and Gracely, 2003). The event of reaching the perceived threshold is generally dependent on the excitability of the peripheral receptors or afferent nerves and on the spatio-temporal integration of the afferent discharge in the central nervous system (CNS). The pain threshold is usually assumed to represent the sensory discriminative component of pain. However, perceived pain may also be influenced by arousal (Chapman et al, 2001) or by psychosocial and cultural factors (Price, 2000; Turk, 1999).

Threshold levels have been seen to change as a consequence of ongoing pain, possibly reflecting neural plasticity where enhanced (hyperalgesia) or reduced pain (hypoalgesia) may occur, depending on the balance between these modulatory influences. Perceived thresholds could be affected by long-term noxious stimuli that sensitize peripheral neural structures. This sensitization may also produce plastic changes in central neural structures involved in pain perception, e.g., through wind-up or expansion of receptive fields of CNS neurons which further on could increase postsynaptic excitability and contribute to central sensitization (Coderre, Katz, Vacarino, and Melzack, 1993; Urban and Gebhart, 1999). The assessment of changed thresholds could therefore be a valuable additional tool in the daily clinical evaluation of pain.

The interpretations of pain data are complex, as the experience of pain is subjective and not uniformly, proportionally related to the extent of injury or stimulation (Coderre, Katz, Vacarino, and Melzack, 1993; Melzack, Coderre, Katz, and Vaccarino, 2001; Price, 2000). Therefore the relationship between the rated sensations in response to increasing intensity of electrotactile stimulation could be regarded as non-linear (Donaldson et al, 2003) but the data will have an ordered structure. Hence, in this study we will apply a statistical approach that is suitable for all types of data having an ordered structure (Svensson, 1993).

In order to study the measurement qualities of threshold assessment the aim of this study was to evaluate the reliability in terms of systematic and individual variability within and between days for electrical sensory and pain threshold assessments in healthy volunteers and pain patients. In addition, healthy volunteers were compared with pain patients regarding both sensory and pain threshold values.

Materials and methods

Participants

Healthy female volunteers, students from the physiotherapy program at the Karolinska Institutet, were consecutively recruited to test their perceived thresholds. Also female patients with pain present in different body areas and of different etiologies, classified as nociceptive, neuropathic, and chronic pain, were recruited from Karolinska University Hospital, the Rehabilitation medicine clinic and Spinalis to participate. The study was restricted to women primarily to avoid gender-related effects (Berkley and Holdcroft, 1999; Riley et al, 1998; Rollman, Lautenbacher, and Jones, 2000). The women gave their informed consent, and the study was approved by the Ethics Committee of Karolinska Hospital (dnr 01-169 and 03-162).

Procedure of assessing electrical thresholds

Electrotactile stimulation with continuously ascending intensity, by the PainMatcher instrument, was used to assess the electrical
sensory threshold (EST), namely, the least stimulation perceived at all by the subjects, and the electrical pain threshold (EPT), the least stimulation required to produce the first perception of pain.

Participants were informed that a paresthesia-like sensation would appear in the thumb and index-finger of the tested hand when the EST was reached, and a distinct sensation of pain, separated from perceived unpleasantness, at the EPT. They were also informed that the stimulation does not cause tissue damage at any level. To become familiar with the procedures, all participants were allowed to try out the equipment 3–4 times prior to the testing procedure.

Electrical stimulation device for threshold assessments

The electrical stimulation device (PainMatcher™, Cefar Medical AB, Lund, Sweden; see Figure 1), is a microprocessor that distributes constant current (15 mA) in mono-phasic rectangular pulses at random velocity with a frequency of 10 Hz to the electrodes. To create an electrically closed circuit the electrodes of the instrument are pressed with the thumb and second finger of one hand. The electrode placed under the index finger is the negatively charged electrode, the cathode. The contact area of the electrode (~6 cm²) and, hence, the resulting current density, is ensured by a certain minimum finger pressure against the electrodes. This is safeguarded by instructing participants to press the electrodes requiring a certain minimum, pre-determined pressure. Pressures below this pre-determined level result in no electrical output and eventually higher pressures will not affect the functionality of the stimulator. The stimulator is designed to compensate for skin resistance variations, in case of sweating or anxiety, up to 13 kΩ in order to produce a constant current.

Intensity is raised by increasing the pulse duration ranging from a minimum of 4 µs to a maximum of 396 µs in increments of 4 µs and a total of 99 steps. One measurement series from minimum to maximum intensity takes less than one minute. The electrical charge per pulse is extremely low, 5.9 µC.

When reaching the thresholds, EST and EPT respectively, the fingers are released from the electrodes and an open circuit is detected. The increase in the constant current generation is interrupted and a value between 0 and 99 (directly related to the pulse width) is then displayed on the LCD screen and automatically saved in the memory. The assessment procedure is based on individual responses with no visualized predetermined lower or upper limits, which

![Figure 1. Electrical stimulation device for threshold and matching assessments, PainMatcher™ (Cefar Medical AB, Lund, Sweden).](image)
blinds the subjects as well as the examiner to the assessment outcome.

Study design

Evaluation of the test–retest variability in EST and EPT assessments were made in two consecutive days in the healthy volunteers. Four daily assessments were performed with a 30-minute rest between each to avoid carry-over effects. The experiments were carried out in the early afternoon, on the same premises, and by the same investigator on both days.

Threshold assessments were performed at one of the participants’ regular visits to the clinic, and were assessed twice within 30 minutes on the same day.

Statistical methods

The mean value and standard deviation (SD) were calculated for age. The data from the EST and EPT assessments, the PainMatcher\(^{\text{\textcopyright}}\) (PM values) were described as the median and range of the PM values since they are based on perceived sensation and therefore regarded to be ordinal.

In the healthy volunteers the evaluation of the test–retest, between-day variability of the EST and EPT assessments on two consecutive days was based on the median values of the four repeated assessments of each individual within each day, and the within-day variability was based on the two first assessments made on the first day. The test–retest data from the patients were collected within the same day.

The within- and between-day variability in the threshold assessments were evaluated by means of the joint frequency distributions of paired data in contingency tables and by the marginal frequency distributions of data from the two comparing occasions. The main diagonal of identical values in the contingency table is oriented from the lower-left to the upper-right corner, which corresponds to direction of the main diagonal of scatter plots of continuous data. The percentage agreement (PA) and the percentage agreement within one PM value (PA ± 1PM) were calculated. In order to identify the type of disagreement, a detailed evaluation of the variability in repeated assessments was performed by means of a rank-based method that takes account of the non-metric properties of ordinal data that was used (Svensson, 1993, 1998a, 1998b, 2001; Svensson, and Starmark, 2002). This method makes it possible to identify and measure the systematic disagreement in assessments (bias) separately from the additional individual component of variability in repeated assessments. A sign of systematic disagreement, a bias, between two assessments appears as different, marginal frequency distributions of the threshold values between the two occasions (Svensson, 1993; Svensson, 1998a). The difference between the probabilities for the first set of marginal distributions being shifted towards higher or lower values relative to the second set defines the measure of relative position. The empirical measure of relative position is denoted RP, and was estimated from the observed threshold values. Possible values of RP range from −1 to 1, where RP = 0 means lack of systematic change (bias) between the two assessments. The level of bias was graphically illustrated by plotting the two sets of cumulative proportions of the observed threshold values against each other beginning at the point (0, 0) yielding a type of relative operating characteristic (ROC) curve. As mentioned, in case of bias, the two marginal frequency distributions differ, and so do the two sets of cumulative proportions, which mean that the ROC curve will deviate from a straight line.

The two sets of marginal distributions reveal only the systematic part of variability between repeated assessments (Svensson, 1993). The joint distribution of paired data in the contingency table contains information about an additional individual variability that is unexplained by the marginal frequencies. In order to calculate this additional individual variation, bias is adjusted for by transforming the pairs of data in the table to pairs of rank values, where the ranks are tied on the pairs. The square of the rank differences is the basis of the calculation of the relative rank variance, RV, which is a measure of the individual variation. The higher the values of RV, the more dispersed are the observations. Possible values of RV range from RV = 0, lack of individual variations, to RV = 1 which means a variability of the same magnitude as from independent unpaired assessments (Svensson, 1993).

The comprehensive evaluation including the marginal distributions, the ROC curve and the
RP value, and the frequency distribution of pairs of data and the RV value is valuable for understanding the sources of variation and for improving the reliability of assessments. Reliable data are unbiased, which means RP = 0, and are distributed close to the main diagonal of the contingency table, which means high values, about 80–90%, of PA (and PA ± 1PM) and a RV value close to zero. An important consequence of bias as evident from a non-zero RP value and with a corresponding 95% confidence interval (CI) that does not cover zero (Altman, 1991) is that the PA will never reach 100%. However, as the RV value is a measure of additional variability after adjustment for bias, the RV value could reach zero both in biased and unbiased repeated assessments. Therefore, a small value of RV in the presence of bias indicates that a high reliability is possible when the source of bias is accounted for.

The differences in assessed EST and EPT between the groups of healthy volunteers and patients in pain were analyzed by the Wilcoxon-Mann-Whitney U test. A p-value less than 0.05 was regarded as significant. STATISTICA 6.0 was used to calculate descriptive statistics and the Wilcoxon-Mann-Whitney U test. The software package SYSS-RAN 1.0 for Matlab 6 was used to calculate RP, RV, and corresponding 95% confidence intervals (CI).

Results

Test—retest variability – healthy volunteers

The test—retest within-day assessments were completed by 48 healthy volunteers (mean age 22.5; SD 2.6 years). The electrical sensory threshold (EST) median values were 3, (Range, 1 to 8) during the first two assessments on the first day and the electrical pain threshold (EPT) median values were on the same occasions 15 (Range, 7 to 40).

The test—retest between-day assessments were completed by 35 of the 48 (73%) healthy volunteers, (mean age 22.2; SD 2.5 years) as 13 of the 48 subjects dropped out gradually. The main reason for dropping out was the subject’s experience of similar results in the repeatedly assessed sensory and pain threshold levels of the first day, making them unmotivated to participate on the second day of assessments.

The between-day variability in EST assessments is shown in Figure 2A. The EST median values were 4 (Range, 2 to 6) and 3 (Range, 1 to 6) in the two consecutive days. There was a 57% total agreement and 94% of the test—retest retest values agreed within ±1PM value. The ROC curve, Figure 2C, shows evidence of a slight bias towards lower EST values on the second day which is also evident from the negative RP value, −0.13. The 95% CI from −0.27 to 0.02 covers the zero value and is wide, which means that there is insufficient evidence to confirm a significant bias in this study. According to Figure 2A the additional individual variation is small, RV 0.05 (95% CI, 0.02 to 0.12), which means that the observed variability could be explained by the systematic change towards lower EST values on the second day assessments.

Figure 2B shows the between-day variation in EPT values among the healthy volunteers. The median EPT value was 15 in both days and ranged from 8 to 29. The percentage agreement, PA, was 20% and the PA ± 1PM value was 49%. The ROC curve in Figure 2D indicates lack of systematic change in EPT; RP = 0.02. The 95% CI (−0.12 to 0.16) is wide but covers zero value symmetrically. Hence the main explanation of the between-day variations in EPT is the significant individual variation, RV = 0.18 (95% CI, 0.06 to 0.30).

Test—retest variability – patients

The test—retest within-day assessments was performed by 36 patients (mean age 41.1; SD 12.5 years) with duration of pain ranging from 0 to more than 12 months in patients with nociceptive pain and for more than 12 months in patients classified as neuropathic and chronic pain.

Figure 3A shows the within-day variability in EST assessments in pain patients. The median values were equal (4; Range, 3 to 8), the percentage agreement, PA, was 53% and PA ± 1PM was 92%. The ROC curve in Figure 3C indicates a possible bias towards higher EST values on the second occasion which is also indicated by the RP of 0.10. The 95% CI is wide and covers zero
value of RP, (95% CI, −0.05 to 0.26). The individual variability is small but non-zero, RV, 0.07 with 95% CI 0.01 to 0.14. The variability could mainly be explained by the possible bias.

Figure 3B shows that the within-day variability in EPT ranged from 3 to 27 on both occasions, and the median value was 7. The PA was 39% and the PA ± 1PM was 78%.

From the ROC curve, Figure 3D, and the RP, −0.03 one can conclude lack of bias; (95% CI, −0.09 to 0.03). Figure 3B shows negligible individual variations; RV 0.01 (95% CI, 0 to 0.03). No obvious explanation to the observed variability is evident from this evaluation.

Comparison of EST and EPT between healthy subjects and patients

The median values of the first two assessments of EST during day 1 was 3 (Range, 2 to 6) in the healthy volunteers and the corresponding value was 4 (Range, 3 to 8) in the patients. The EST was regarded higher in the patients than in the healthy volunteers (p = 0.03). The median EPT was higher in the healthy volunteers than in the patients 15 (Range, 8 to 28) and 7 (Range, 4 to 27) respectively (p < 0.001), Figure 4.

**Discussion**

The results of this study showed that the percentage agreement within ±1PM value ranged in the two groups of healthy volunteers and patients in pain between 92% and 94% in EST and between 49% and 78% in the EPT assessments, respectively. Possible explanation to the variability in the test—retest assessments of the EST could be a slight systematic disagreement, bias, between the two occasions. Since the

<table>
<thead>
<tr>
<th>EST</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PM values day 1</strong></td>
</tr>
<tr>
<td>1 2 3 4 5 6</td>
</tr>
<tr>
<td>Tot CP</td>
</tr>
<tr>
<td>PM values day 2</td>
</tr>
<tr>
<td>Tot CP</td>
</tr>
</tbody>
</table>

*Figure 2. A-D: Test—retest between day variability of assessed electrical sensory threshold (EST) and electrical pain threshold (EPT) in healthy volunteers; 2 A-B, Joint distribution of paired data from day 1 and day 2 assessing EST (2A) and EPT (2B) shown as PainMatcher® (PM) values, n = 35. Tot = total frequency; CP = cumulative proportion; 2 C-D, The cumulative proportion are shown in the relative operating characteristic (ROC) curve of paired assessments of EST (2C) and EPT (2D), n = 35.*
Table 1: PM values day 1 and day 2

<table>
<thead>
<tr>
<th>PM values day 1</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
<th>26</th>
<th>27</th>
<th>28</th>
<th>29</th>
<th>Tot</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM values day 2</td>
<td>29</td>
<td>28</td>
<td>27</td>
<td>26</td>
<td>25</td>
<td>24</td>
<td>23</td>
<td>22</td>
<td>21</td>
<td>20</td>
<td>19</td>
<td>18</td>
<td>17</td>
<td>16</td>
<td>15</td>
<td>14</td>
<td>13</td>
<td>12</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>8</td>
<td>1</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.94</td>
<td>0.91</td>
<td>0.89</td>
<td>0.83</td>
<td>0.77</td>
<td>0.74</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Tot 2 1 0 5 0 3 4 5 2 6 1 1 0 2 1 0 1 0 1 0 0 0 35

CP 0.66 0.69 0.99 0.23 0.31 0.43 0.54 0.63 0.80 0.83 0.86 0.86 0.91 0.94 0.94 0.97 0.97 1.00 1.00 1.00 1.00

Figure 2. Continued.
measure of systematic disagreement indicated different directions (i.e., towards lower and higher values respectively), a learning effect is not the likely explanation. The individual variations, after adjusting for possible bias, were negligible. The EPT assessments were unbiased in both groups but individual variations were significant among healthy volunteers and negligible among the patients in pain. The results in this study indicate stable and reliable assessments of electrical sensory and pain thresholds except for a possible bias.

The pain threshold assessments implies the identification of a point on a continuum of increasing stimulus intensity that distinguishes non-painful from painful experience suggesting that the perception of pain has to be evaluated in contrast to a back-ground noise of other sensations (Borg, 1993). It is probably more difficult to determine the pain threshold than
to distinguish no sensation from the least detectable sensation at all, such as at the sensory threshold. Despite the possible insecurity in finding the exact pain threshold level twice, the results from the contingency tables in this study identify just a few individuals that differs five to seven categories in the healthy volunteers and five to six categories among the pain patients in the paired assessments. The patients’ estimations of the pain threshold seemed less variable compared to the healthy volunteers, possibly due to patients’ greater confidence in identifying pain sensations. The demonstrated variability may also be explained by the fact that perceived pain is highly personal and subjective. Therefore, a greater agreement than that observed is perhaps not possible to find, provided assessments are made under the same circumstances. The results of this study support the belief that

<table>
<thead>
<tr>
<th>PM values 1st assessment</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>Tot</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>1.00</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
<td>9</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Tot</td>
<td>0</td>
<td>1</td>
<td>13</td>
<td>14</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>0</td>
<td>0.03</td>
<td>0.39</td>
<td>0.78</td>
<td>0.94</td>
<td>0.97</td>
<td>0.97</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. A-D, Test—retest within day variability of assessed electrical sensory threshold (EST) and electrical pain threshold (EPT) in patients in pain; 3A-B, Joint distribution of paired data from 1st and 2nd assessments of EST (3A) and EPT (3B) shown as PainMatcher® (PM) values, n = 36. Tot = total frequency; CP = cumulative proportion; 3 C-D, The cumulative proportion are shown in the relative operating characteristic (ROC) curve of paired assessments of EST (3C) and EPT (3D), n = 36.
the sensory threshold is easier to assess than the pain threshold.

It is possible that (a) fewer categories to choose between during assessments, and (b) opportunities for participants to test the procedure prior to tests, could increase participants’ ability to judge the pain threshold with greater certainty.
Another finding of this study was that there were different threshold levels between the healthy subjects and the patients in the within-day test/retest. The responses of the patients in pain were observed as increased sensory threshold and decreased pain threshold in comparison with the healthy participants, which was similar to earlier findings (Alstergren and Förström, 2003; Hendiani et al, 2003). In the pain patients this could indicate a decreased ability to discriminate non-painful stimulation, described as an adaptation due to pathological changes in somatosensory processing following pain and also an increased sensitivity to painful stimuli as compared to healthy participants (Voerman, van Egmond, and Crul, 2000). Test–retest variability in the patients was, in this study, only evaluated within the same day in order to avoid effects influenced by varied clinical pain.

Different ages of the participants in the two groups could have influenced the results. However, the pain threshold has in general been described as increased in elderly adults, and the pain tolerance as decreased. On the other hand, age-dependent responses have been reported as dependent on the stimulation modality (Gibson and Helme, 2001).

While threshold assessments mostly are reported from research projects in physical therapy the clinical applicability of the psychophysiological assessments of sensory threshold and pain threshold are not yet fully described though it could be an additional valuable tool when applied into clinical routine for evaluation of pain interventions in physical therapy.

Figure 4. EST and EPT assessments in healthy subjects and pain patients shown as PainMatcher (PM) values. The median values of the first two assessments during day 1 in healthy volunteers were compared with the median values of the two assessments in patients in pain, \( n = 48 \) and 36, respectively.
therapy. Detection concepts including threshold assessments are useful in quickly ascertaining changes in pain sensitivity and have been found to be fairly reliable and produce generalizable assessments (Price, 1993; Sang, Mitchell, and Gracely, 2003). Furthermore, electrical stimulation is reported as producing an easily detectable pain sensation, easy to control and generating reproducible value (Gracely, 1999). The results of this study support these findings. Moreover, electrical stimulation of the skin is the least likely to involve any danger of tissue damage in comparison with other modes of evoking pain (Lahoda, Stacher, and Bauer, 1977).

In this type of study a common measure of reliability is the Cohen’s kappa that is a measure of agreement beyond the chance-expected agreement (Altman, 1991). However, summary measures like kappa and percentage agreement produce no detailed information concerning the art of disagreement in the paired ordinal data. Furthermore, the approach by Svensson used in this study provides a comprehensive analysis of an observed disagreement and makes it possible to identify and measure systematic disagreement (bias) separately from individual disagreement between two assessments.

The described procedure of assessing thresholds is experienced as simple to use and gives the subject the possibility of controlling the stimulation duration. The simplicity could also make it possible to apply in clinical evaluative work as well as in clinical and experimental studies.

In conclusion, the results of this study indicate stable and reliable assessments of electrical sensory and pain thresholds except for a possible bias. Therefore, it appears that the threshold assessment procedure followed in this study may be a valuable clinical tool when integrated into physical therapy examination for evaluation of pain-alleviating interventions. However, a larger study is needed to confirm the findings of measurement qualities in this type of threshold assessments.

Acknowledgement

This work was supported by grants from Karolinska Institutet’s foundation.

References


Svensson E 1993 Analysis of systematic and random differences between paired ordinal categorical data. Stocholm, Almqvist & Wiksell International, Sweden


Svensson E 1998b Ordinal invariant measures for individual and group changes in ordered categorical data. Statistics in Medicine 17: 2923–2936


