

## ORIGINAL REPORT

# ASSESSING PAIN PERCEPTION USING THE PAINMATCHER® IN PATIENTS WITH WHIPLASH-ASSOCIATED DISORDERS

Lina Bunketorp Käll, PhD, RPT<sup>1</sup>, Jan Kowalski, BSc<sup>2</sup> and Elisabet Stener-Victorin, RPT<sup>3</sup>

From the <sup>1</sup>Institute of Neuroscience and Physiology, Department of Occupational Therapy, Audiology and Physiotherapy, <sup>3</sup>Department of Physiology, Institute of Neuroscience and Physiology, The Sahlgrenska Academy at Göteborg University, Göteborg and <sup>2</sup>Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden

**Objective:** To evaluate the Painmatcher®, in terms of reliability, and to explore the relationship between pain magnitude matching and pain threshold assessments in patients with whiplash-associated disorders. Also, to investigate gender differences in pain thresholds and explore the correlation between pain-related cognitions and pain threshold.

**Design:** A test-retest study.

**Subjects:** Forty-seven patients with whiplash-associated disorders.

**Method:** A visual analogue scale and a Painmatcher® (an instrument for comparing pain magnitude) were used to evaluate pain intensity. Pain threshold was assessed using the Painmatcher®. Pain-related cognitions were assessed using the Painometer® and the Tampa Scale for Kinesiophobia.

**Results:** The Painmatcher® demonstrated reliable pain magnitude matching scores, but the pain threshold assessment indicated a systematic disagreement. Women exhibited significantly lower pain thresholds than men ( $p < 0.01$ ). There was a weak but significant correlation between the pain intensity according to the visual analogue scale and the Painmatcher® ( $r = 0.46$ ) ( $p < 0.01$ ). There was a significant correlation between the emotional experience of pain and pain threshold ( $r = -0.33$ ) ( $p < 0.001$ ), but no significant correlation between fear of movement/(re)injury and pain threshold.

**Conclusion:** Measuring pain with the Painmatcher® is a reliable method, but may include a possible bias in threshold assessments and seems to be associated with unpleasantness.

**Key words:** Painmatcher®, whiplash, pain, perception, affective, kinesiophobia.

J Rehabil Med 2008; 40: 171–177

*Correspondence address:* Lina Bunketorp Käll, The Sahlgrenska Academy at Göteborg University, Institute of Neuroscience and Physiology, Department of Occupational Therapy, Audiology and Physiotherapy, Box 455, SE 405 30 Göteborg, Sweden. E-mail: lina.bunketorp@gu.se

Submitted June 13, 2007; accepted November 8, 2007

## INTRODUCTION

Pain is a common reason for seeking healthcare in society (1). Neck pain following motor vehicle accidents provokes scientific controversy and tremendous cost to society. The acute pain in the initial phase may be caused by minor sprains to the soft

tissues in the neck (2). The interpretation of pain assessment is, however, complex, and clinical observations are often contradictory to the old theory that the degree of pain experienced corresponds with the degree of injury (3, 4). The nociception of pain and the experience of pain are not one and the same (5, 6). This might indicate the presence of substantial individual differences in the processing of information related to noxious stimulation (7), and explain why patients with apparently the same extent of physical pathology perceive pain differently. The experience of pain is comprised of several dimensions, such as pain intensity, unpleasantness, fear, and anxiety, which are all independently affected by personal traits and psychological factors. Previous studies have shown that fear of pain is associated with exaggerated pain perception (8, 9).

Measuring pain is of utmost importance, but a difficult task in clinical management, especially in terms of the interpretation of patients' information, because there is no "gold standard" that provides the therapist with a method objectively to quantify this complex and subjective experience. In an effort to improve healthcare, different self-reporting measures have been developed. Pain is commonly assessed using rating scales and questionnaires, but can also be assessed using psychophysical devices. The visual analogue scale (VAS) (10) is a simple tool that measures the subjective aspect of the sensory pain experience. However, it does not give a multidimensional picture of the pain and is bounded by fixed end-points, which limits the range of measurements. Furthermore, the VAS is limited by the patient's memory, as the subject has to compare and grade the present pain intensity against the memory of previous pain experienced or the worst imaginable pain. The Painmatcher® (PM) is a newly developed instrument for comparing pain magnitude (11, 12) that allows the patient to match perceived pain in a certain region of the body to a physical sensation between the left thumb and index finger produced by the PM. Based on these psychophysical qualities the PM is a blinded instrument that has no predetermined visible upper limits. The PM can also measure pain thresholds in experimental settings (13).

A useful instrument must have a high level of intra-individual agreement, especially in test-retest assessments. Unfortunately, there is widespread confusion and misuse of statistical methods in reliability studies investigating the level of agreement (14, 15). The test-retest repeatability of different ordinal rating scales, such as the VAS, has been examined using

correlational statistics, for example. The use of such statistics is not only misleading, but also an inappropriate approach to judge agreement, since they assess only the strength of the relationship between the measures and not the actual agreement (16). Svensson (17, 18) suggests that a specific ranking method approach suitable for all types of ordinal data should be used in evaluations of variables that are categorical or numerical and have an ordered structure. The reliability of the PM has been evaluated using this approach in patients with acute and chronic pain as well as in healthy individuals (11, 19, 20), but not in patients with subacute whiplash-associated disorder (WAD).

Chapman et al. (21) state that pain measurement is meaningful only if it is valid; that is a pain report must verifiably scale what it purports to measure. However, when measuring pain using various scales and devices the experience might be dominated by the sensory aspects of pain for one patient and by the emotional aspects of pain for another, or by a mix of both aspects for a third patient. Chapman et al. (21) state that perceived pain may be influenced by arousal, as indicated in a previous study (22). Even though the PM is completely harmless, the use of a psychophysical instrument that produces a noxious stimulus may be associated with unpleasantness, as previously shown (12). Since the PM may cause discomfort during assessment it might also be influenced by the fear of increased pain. This would cause difficulties in interpreting and quantifying the evaluation of the pain experience using the PM and suggests that studies investigating the relation between pain-related cognitions and perceived pain perception are needed. Furthermore, the authors of the aforementioned study (12) recommend that the ratio between the ongoing pain intensity and the pain threshold assessment using the PM needs to be investigated further.

This study aims to evaluate the PM in terms of reliability, and to explore the relationship between the magnitude matching and the pain threshold assessments using the PM in patients with subacute WAD. In addition, this study aims to investigate gender differences in pain thresholds and to explore the correlation between pain-related cognitions and pain threshold among patients.

## METHODS

The paired baseline data used in the present study were obtained from a randomized controlled trial (RCT) (23). The RCT was carried out at an interdisciplinary rehabilitation centre that specializes in WAD. Patients were included in the RCT based on the following criteria: subacute WAD (symptoms lasting for more than 6 weeks but less than 3 months) following a whiplash-type trauma to the neck. WAD was defined as a musculoligamentary sprain or strain of the cervical region, with no fractures and no dislocations of the cervical spine. The exclusion criteria in the study were: (i) unrelated diseases; (ii) additional injury that precluded completion of the questionnaire or would make evaluation difficult; (iii) previous severe neck pain for which the patient took more than one month of sick leave or received disability pension in the year preceding the accident; and (iv) inability to read and speak Swedish. This study was approved by the regional ethics review board.

## Patients

Of the total 212 patients who were sent a recruitment letter for the RCT, 63 patients (30%) with subacute WAD were willing to participate. Of the respondents, 6 were excluded because of unrelated diseases making evaluation difficult and 8 due to sick leave because of neck pain the year preceding the accident, among whom 6 had been on sick leave due to a previous whiplash trauma. Thus, 49 patients fulfilled the criteria and were included in the study and randomized to either a supervised individually tailored training group or a home training group. Two patients became ineligible and were excluded following randomization due to the detection of a severe, unrelated disorder and the receipt of a magnetic resonance imaging scan with evidence of severe degenerative lesions of the cervical spine prior to the trauma. Forty-seven patients were finally included in the study, 30 women (64%) and 17 men (36%) with a mean age of 31 years (age range 18–61 years). The whiplash trauma had occurred on average 64 days (range 42–121 days) prior to the start of the study. The patients who were not interested in participating in the study entered the regular treatment programme at the rehabilitation centre.

## Measurements

The sensory dimension of pain was assessed with a traditional 100-mm VAS with the end points 0 (no pain) and 100 (worst pain). The VAS is demonstrated to have a good reliability and validity (10, 24).

Affective dimensions of the pain were assessed using the Painometer® (POM). The POM is a self-administered, hard plastic, hand-held tool 20 cm long, 5 cm wide and 2.5 cm thick. The POM is a practical instrument that combines the VAS and parts of the McGill pain questionnaire (MPQ) composed of 78 word descriptors (25) into a single tool. On the back of the POM is a group of 23 sensory and affective words (POM-WDS). The patients were asked to mark all the descriptors in the group that described the affective dimension of their pain. Each of the descriptors was assigned a “quality of pain” intensity ranging from 1 to 5. The first 12 words describe the sensory component of pain – the sensation of pain, how the pain feels, and how much it hurts, such as cramping, dull, burning, sore, hurting and stabbing. The last 11 words describe the affective component of pain – the feelings and emotions surrounding pain and how much a person is bothered by the pain, such as annoying, sickening, terrifying, unbearable and troublesome. The intensity scores are hidden from the patient, but are listed on the documentation sheets. The intensity scores allow caregivers to assess whether the pain is mainly sensory or affective in nature. Summing the values of the chosen words is a quick method for comparing the intensity of these 2 components of pain. The POM has demonstrated good validity and reliability (26).

Secondly, a pain intensity matching device, PM (Cefar Medical AB, Lund, Sweden; Fig. 1) was used producing perceptual matching by gradually increasing electrical impulses between 2 fingers. The patient was instructed to hold the PM with the electrical stimulation unit between the right thumb and index finger with a firm grip. The PM causes an increasing electrical current that eventually becomes painful, and the patient is told to match the ongoing neck pain intensity to



Fig. 1. The Painmatcher® (Cefar Medical AB, Lund, Sweden).

the physical sensation between the fingers. The electrical stimulation is initiated by the patient and delivers electrical pulses with increasing intensity at random speeds. When the experienced neck pain corresponds in amplitude to the sensation of pain in the right hand, the patient was told to release the fingers from the PM, stopping the electrical stimulation unit, and a value was automatically recorded on a liquid crystal display screen (LCD-screen) and automatically saved in the memory.

The PM provides constant electrical stimulation controlled by a microprocessor, delivering rectangular pulses with a frequency of 10 Hz and amplitude of 15 mA. The pulse width gradually increases in 99 steps, with increments of 4  $\mu$ s, from zero to a maximum of 396  $\mu$ s. The value reached between 0–99 is directly related to the pulse width, and as soon as the electrical circuit is detected by releasing the fingers from the electrode device, the electrical stimulation stops. PM is based on a random variation of the time it takes to increase stimulation. That way, the measurement is determined only by what the patient feels, not how long the measurement takes. The PM meets international ethical norms, because the patient can interrupt the stimulation at any time. The measurement procedure is harmless, involves no intervention, and has no side-effects. The scores obtained on the PM might be concealed (in “Settings”) and appear then as a line and are not displayed for the patient until all 3 measurements are completed. The pain magnitude matching score was recorded 3 times, and the median value of these was used in the analysis. The PM has demonstrated good reliability (11).

*Pain threshold* was assessed using the PM. The patient was told to release the fingers from the PM when the first sensation of pain was experienced. The obtained score was used to define the pain threshold. The pain threshold score was recorded 3 times, and the median value of these was used in the analysis.

#### *Fear of movement/(re)injury*

The Tampa Scale for Kinesiophobia (TSK) was used to assess fear of movement/(re)injury. The scale was previously translated into Swedish in a forward and backward translation procedure by 2 other research groups (Linton S, personal communication and 27), but only one publication addresses the translation procedure (27). The TSK contains 17 statements developed to identify the fear of (re)injury due to movement or activities such as “It is not safe for a person with a condition like mine to be physically active”. Scores range from 1 (strongly disagree) to 4 (strongly agree). The scores on items 2, 4, 8 and 16 are reversed so that high scores on all items indicate high levels of fear. The total sum score ranges from 17 to 68. No publication addresses the reliability or validity of the original version of the TSK in American English. The Dutch version of the TSK appears to have sufficient reliability and validity (28). Recently, the reliability and validity of the Swedish version of the TSK has been established (16, 27).

#### *Procedure*

During baseline assessment in the RCT, the PM was used to evaluate intra-individual agreement in a test-retest assessment regarding pain magnitude matching and threshold assessment. In order to accustom the patients to the PM they tried out the instrument before the measurement started. They were informed that the PM is completely harmless and does not cause any tissue damage. The ratings were performed in the following order: PM – pain threshold (pain threshold was defined as the weakest amount of electrical stimulation the patient considered painful); PM – pain matching (pain matching was defined as the electrical stimulation the patient considered as intense as his/her own physical pain in the neck); and thereafter the patients rated the sensory dimension of pain according to VAS, the affective dimension of pain according to the POM, and completed the TSK. The sensory and affective dimension of pain was assessed between the PM test-retest assessments. The time interval between the test-retest PM procedure was approximately 5 minutes. The measurements were administered to the patients upon admission to the rehabilitation centre in a room with no outside disturbing factors.

#### *Statistical methods*

As the PM measurements are considered to be ordinal (based on perceived sensation), the PM values are presented as median and range. To estimate test-retest reliability (agreement), the same examiner administered the paired assessments on the same individual. The statistical method used to estimate reliability is introduced by Svensson (17, 18) and was chosen to preserve the non-metric, rank-invariant properties of the data in the test-retest analysis. The method provides estimates to identify and separately measure the level of systematic disagreement (group specific) and individual disagreement (unexplained by the systematic disagreement) between the 2 test occasions.

The empirical measure of the random part of the disagreement (individual dispersion) is the relative rank variance (RV). Possible values of RV range from zero to 1 and express the component of random disagreement, which is adjusted for systematic disagreement. The larger the random contribution is to the total disagreement, the higher the value of RV. An RV equal to zero indicates a lack of individual dispersion, which means that the method has good reliability and is responsive (i.e. powerful enough to capture true treatment effects). The presence of systematic disagreement is expressed by relative position (RP) and relative concentration (RC). Values of RP and RC range from –1 to 1 and a value close to zero indicates negligible systematic disagreement. The presence of RP (RP $\neq$ 0) means that the second of the 2 test occasions has systematically higher (+) or lower (–) ratings. A non-zero RC indicates that there is a difference in the 2 sets of assessments, meaning that one set of the paired data utilizes a smaller part of the range of the scale than the other set of data.

To explore the relationship between the pain magnitude matching score and the pain threshold scores measured with the PM, correlations were calculated with the Spearman rank-order correlation coefficient. The Mann-Whitney *U* test was used to compare gender differences at a group level with respect to the pain threshold measured with the PM. Also, the rank-invariant statistics were used to estimate the test-retest reliability by gender. The Spearman rank-order correlation coefficient was used to explore the relationship between: (i) the intensity of pain measured with a VAS and the pain magnitude matching score; (ii) the affective dimension of pain and the pain threshold measured with the PM; and (iii) fear of movement or/(re)injury and the pain threshold measured with the PM. For all analyses, two-tailed tests were performed, and a *p*-value of less than 0.05 was considered significant. Analyses were performed using the Statistical Package for Social Science (SPSS 13.0).

## RESULTS

To assist the interpretation of the PM test-retest assessments, the joint distribution of the paired pain magnitude matching scores and pain threshold scores are illustrated in Fig. 2.

Plots that coincide with the primary diagonal oriented from the lower-left to the upper-right corner means unchanged scores between the 2 assessments. In both figures, the main diagonal is oriented from the lower-left to the upper-right corner, which indicates unchanged assessments between 2 occasions. The levels of systematic disagreement (as a group) and random disagreement (individual) are illustrated in Table I.

The table presents the pattern of systematic (RP and RC) and random (RV) disagreement between the repeated assessments of the perceptual matching scores and pain threshold scores and their corresponding 95% confidence intervals (CIs). All RP and RC values for the perceptual matching scores were close to zero. The corresponding 95% CIs show no evidence of systematic (statistically significant=CIs excluding zero) disagreement. The RV value was very low indicating a slight

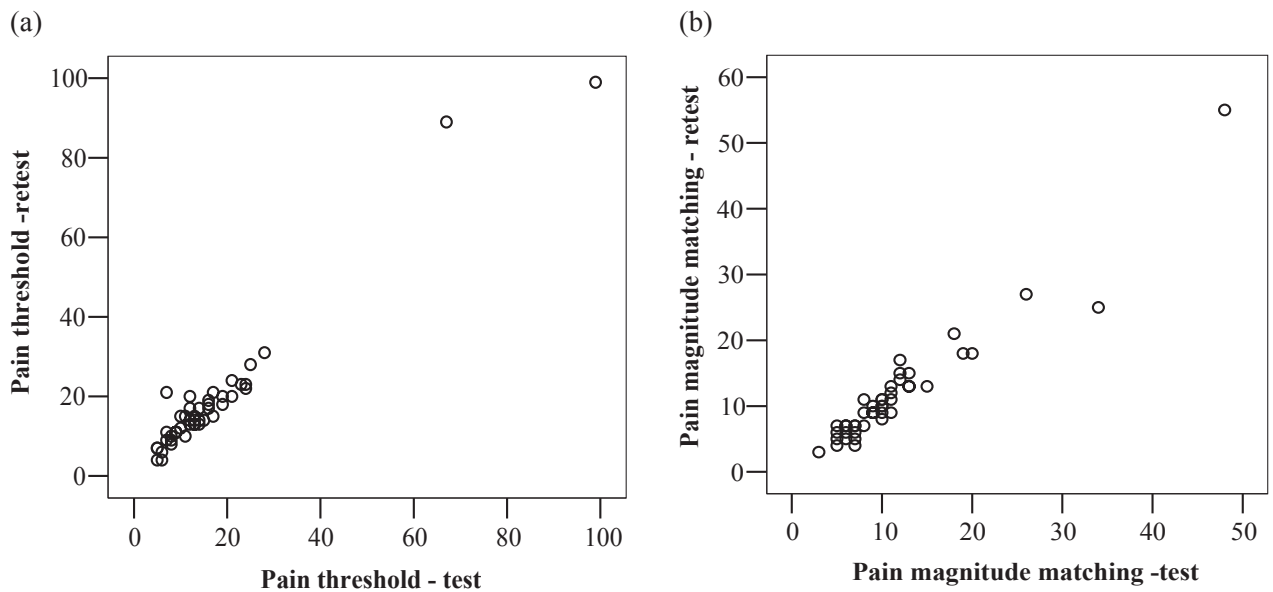


Fig. 2. (a) The joint distribution of the pain threshold test and retest assessment with the Painmatcher®. (b) The joint distribution of the pain magnitude matching test and retest assessment with the Painmatcher®.

dissonance and the corresponding 95% CI can be considered to be a negligible random disagreement. All RC and RV values for the threshold scores were close to zero and the corresponding 95% CIs show no evidence of systematic disagreement. The RP value of the threshold score was negative and more distant from zero (−0.16), indicating a slight systematic disagreement, which implies that the patients were more likely to report lower ratings on the second occasion.

Women exhibited significantly lower pain thresholds than the men ( $p < 0.01$ ). The median pain threshold was 13 (range 4–31) in the women compared with 19 (range 9–99) in the men. The pain magnitude matching scores measured with the PM was rated lower than the pain threshold in 37 patients (79%). The median pain magnitude matching score was 9 (range 3–55) and the median pain threshold score was 14 (range 4–31). According to the VAS assessment, the median pain intensity score reached 46 (range 10–88). There was a significant corre-

lation between the pain magnitude matching score and the pain threshold score measured with the PM ( $r = 0.62$ ) ( $p < 0.001$ ), where high pain magnitude corresponded to high ratings of pain threshold. There was a significant correlation between the pain intensity measured with the VAS and the magnitude matching by the PM ( $r = 0.46$ ) ( $p < 0.001$ ). There was also a significant correlation between the affective dimension of pain and pain threshold measured with the PM ( $r = -0.33$ ) ( $p < 0.001$ ) where high pain affect corresponded to high ratings of pain threshold. However, there was no significant correlation between fear of movement/(re)injury measured with the TSK and pain threshold ( $r = -0.01$ ) measured with the PM, suggesting that fear of movement/(re)injury was not related to pain threshold.

### DISCUSSION

The results demonstrate that the PM is a reliable device for pain magnitude matching assessments in patients with sub-acute WAD. However, the threshold assessment did show a presence of systematic disagreement (a change by group), i.e. the subjects were more likely to assess lower ratings at the second of the 2 test occasions. The RV was, however, low, which indicates a lack of individual dispersion, which means that the method is responsive (i.e. powerful enough to capture true treatment effects). The presence of systematic disagreement in the threshold assessment could be due to the limited time between the repeated measurements. The short interval between the two test occasions could have biased the result by leaving the patients with a residual uncomfortable paraesthesia-like sensation in the fingers caused by the perceptual stimuli generated by the PM. It would perhaps have been preferable to leave at least 30 min between the 2 test occasions, as in a

Table I. Test-retest reliability of the Painmatcher®: magnitude matching and pain threshold assessment. Systematic relative position (RP) and relative concentration (RC) and relative rank variance (RV) (random disagreement) between repeated assessments and their corresponding 95% confidence intervals (CIs) ( $n = 47$ )

|                                       | Perceptual matching | Pain threshold |
|---------------------------------------|---------------------|----------------|
| Systematic disagreement for the group |                     |                |
| In position                           |                     |                |
| RP (SE)                               | −0.03 (0.03)        | −0.16 (0.05)   |
| 95% CI                                | −0.09, 0.04         | −0.25, −0.07   |
| In concentration                      |                     |                |
| RC (SE)                               | 0.08 (0.06)         | −0.07 (0.06)   |
| 95% CI                                | −0.04, 0.19         | −0.18, 0.05    |
| Random individual disagreement        |                     |                |
| RV (SE)                               | 0.03 (0.01)         | 0.10 (0.06)    |
| 95% CI                                | 0.01, 0.05          | 0.00, 0.22     |

SE: standard error.

previous study (20). However, in the study by Lund et al. (20), the same result was obtained; the test-retest assessment of the patients in pain were observed as decreased pain threshold compared with the healthy participants. These findings are in accordance with previous studies (12, 29) and may signify an occasionally increased sensitivity to the painful stimuli (30) generated by the PM. This may have resulted in the patients rating their pain threshold lower the second time.

The patients were allowed to try out the PM several times prior to the testing procedure in order to become familiar with the device, which may have had an effect on the subsequent assessments. One could also question whether the short test interval in the present study made it easier for the patients to remember the intensity of the previous recording. Since a systematic difference implies higher or lower scores on the second test occasions, a learning effect is not likely. Lund et al. state in their study (20) that it is probably more difficult to determine the pain threshold than to distinguish no sensation from the least detectable sensation at all, such as the sensory threshold. It would perhaps have been preferable for the patients to become even more practised with the PM device before the measurements started. Yet, the PM pain threshold assessment had a lack of individual dispersion, which means that the method has good reliability and is responsive (i.e. powerful enough to capture true treatment effects).

An intriguing finding in the present study was that in 79% of the patients, the pain magnitude matching was rated lower than the pain threshold. This is in accordance with unexpected findings in a previous study (12), where almost half of the patients with acute oral pain had a lower PM score for ongoing pain than pain threshold. According to the definition of pain, the pain threshold is defined as the weakest amount of stimulation the patient considers painful, i.e. the lowest VAS score. The VAS assessments in the present study, however, revealed that the patients were suffering from pain in the neck (the median VAS score reached 46 (range 10–88)), which brings into question the interpretation of the pain magnitude matching results, which are supposed to be considered as intense as the patient's actual pain in the neck according to the manual. The seemingly contradictory findings in the present study might be explained by the fact that the patients' pain is experienced as enormously trying, even though the intensity of pain may be low, since it is always present. Besides, a previous study (6) has demonstrated that the sensory and affective dimensions of pain are 2 distinct components of pain, which may explain the variability in patients' responses to apparently the same extent of physical pathology. Based on this finding one could speculate whether patients' emotional experience may exceed the sensory dimension of the pain. For some individuals, the intensity of pain as measured by the PM might in that case be lower compared with the perceptual stimulus that is required to reach the pain threshold using the PM. The unexpected findings in the present study questions the clinical relevance of the PM and call for further studies investigating its usefulness.

The women did report significantly lower pain thresholds than the men in the present study. This is in accordance with

previous findings, as the literature suggests that women exhibit lower pain thresholds than men (31). However, the literature widely suggests that gender is only one factor that may influence the experience of pain. Although it has been extensively investigated, the precise physiological and psychological mechanism underlying the differences remains unclear (31). It is suggested that gender differences in pain perception may arise from differences in perceptual sensitivity or style, in cognitive and emotional ways of dealing with pain, and in social or occupational roles (32). Analgesic consumption and female hormones are other factors that may influence the perception of pain.

Two of the patients rated their pain threshold very high in the test-retest procedure, as indicated in Fig. 2a. Both were males and shared the same opinion that they did not experience the stimulus generated from the PM as painful (i.e. it did not give a valid measure of the perceived pain threshold). One of the men mentioned that he was a former car engineer and had changed many spark plugs and was thus used to electrical stimulation, which may have caused the test results.

In a previous study (20), the PM was used to assess pain thresholds in female healthy volunteers and in female patients with pain in different areas of the body and of different aetiologies. In healthy female volunteers the pain threshold median value was 15 (range 7–40) compared with 7 (range 2–27) in the female patients with different pain syndromes (20) and 13 (range 4–31) in the women with subacute WAD in the present study. These findings are somewhat contradictory to the results of previous studies (33–35) showing that patients with WAD exhibit significantly lower pain threshold than pain-free control patients. However, the experimental studies use different types of stimulus modalities for threshold assessments, such as pressure (mechanical), thermal (hot, cold) and electrocutaneous stimuli, which might explain the varying results. Besides, there might be differences in the way the patients were given instructions by the assessor regarding the PM measurement procedure.

Despite being confusing, high pain magnitude matching scores corresponded to high degrees of pain threshold measured by the PM in the present study. There is an emerging body of evidence that central hypersensitivity is responsible for ongoing pain in some patients with persistent WAD, and this may partly explain the occurrence of pain in the absence of detectable tissue damage (34, 35). The hypersensitivity may cause a generalized decrease in nociceptive thresholds (36). Considering the absence of tissue damage at the sites of testing, the authors suggest that the findings may indicate secondary hyperalgesia, which in turn may signify central sensitizations of nociceptive pathways (36). On the basis of these earlier findings, it is thought that high pain magnitude matching in the present study, i.e. high pain intensity would correlate with a low pain threshold and not high pain threshold as the analysis revealed.

In addition to hypersensitivity to pain, many chronic pain syndromes are associated with reduced endogenous inhibition of pain, which implies that an individual's processing of pain-

related information changes with the onset of the syndrome (7). As time passes, the transitional phase between the acute and the chronic pain occurs during the "subacute" phase in which cognitive and affective components of the pain perception become more prominent (37). It is still unclear which factors initiate the vicious circle leading to "late whiplash syndrome" (38). Since the patients in the present study were in the subacute phase of WAD (i.e. presented with symptoms lasting for more than 6 weeks but less than 3 months) a conceivable explanation of the significant correlation between the pain intensity matching scores and pain threshold scores could be that the patients still had a significant endogenous pain-inhibitory capacity. Another explanation could be that, even though a patient's experienced intensity of whiplash associated pain is high, the PM assessment is combined with such a high degree of unpleasantness, which in fact, hinders the patient from reaching and matching the actual pain in the neck to the unpleasant sensation between the fingers. In this case, the patient might estimate both the magnitude matching score and pain threshold score lower than a patient who does not experience as much discomfort during the PM assessment. This would theoretically partly explain the significant correlation between the magnitude matching and pain threshold assessment.

A previous study (7) suggests that individual differences in the endogenous modulation of pain are crucial for understanding variability in pain responses. It would have been interesting to include healthy volunteers in the present study for group comparisons of pain thresholds. A previous study (39) suggests that psychological factors such as fear and anxiety can affect pain threshold levels. The result in the present study does not support these findings, since no significant correlation between fear of movement/(re)injury and pain threshold was found.

There was a significant correlation between the pain intensity score measured with the VAS and the PM magnitude matching score, suggesting that there is an association between these measures in subacute whiplash-associated neck pain. Yet, correlational statistics assess only the strength of the relationship between different measures and not the agreement, and gives no measure of interchangeability (14). The method suggested by Svensson (17, 18) is therefore chosen to estimate the test-retest reliability of the PM, since it makes it possible to identify and measure systematic disagreement (bias) separately from individual disagreement between 2 assessments.

There was a significant correlation between the affective dimension of pain and the pain threshold scores measured with the PM in the present study. This result is somewhat surprising, since high pain affect would supposedly correspond to low ratings of pain threshold and vice versa. However, in the controlled trial considering acute oral pain (12), the degree of pain unpleasantness was higher than the pain intensity in the healthy subjects during the PM measurement. The authors suggest that the degree of unpleasantness experienced from the PM assessment may explain why almost half of the subjects ranked the ongoing oral pain lower than the pain threshold with the PM. The authors conclude that the PM producing noxious stimulus may thus be associated with unpleasantness (12). Ac-

ordingly, it may thus be a measure of pain intensity coloured by the emotional experience of pain. The unpleasant feeling during measurement may result in a lower pain magnitude matching score than was expected.

Since the PM seems to be associated with unpleasantness during the measurement procedure, it might not give a valid measure of the ongoing neck pain intensity. However, construct validity of the clinical pain report is known to be challenging, since the pain problem often interacts with psychological factors and with the patient's physical and social environment (21). Besides, Chapman et al. (21) state that there is rarely a measure of the tissue trauma that causes the pain and the patient nevertheless rarely if ever receives training in scaling pain. Moreover, in clinical pain measurement context may be an important factor (21). Yet, compared with the use of VAS the PM has the advantage that the patient is blinded to the measurement and it is not bounded by fixed end-points. Besides, the patients do not have to rely on memories of lack of pain and worst pain ever experienced, as in the VAS assessment. The PM has been found to be as responsive as the VAS and a numeric rating scale (11), i.e. it might detect individual changes in pain intensity. There is, however, an obvious need for research directed at assessing construct validity for pain measurements like the PM in different patient populations. It would be interesting to investigate which pain measurement tool best mirrors the pain experience according to the patients and to evaluate further the strengths and weakness of the PM and determine ways of improving instrument and its clinical relevance.

In conclusion, pain intensity matching assessment using the PM is reliable. The pain threshold scores tend to decrease in repeated assessments, indicating a systematic disagreement. However, the validity of the instrument as a pain magnitude matching device is questionable since its use seems to be associated with unpleasantness. This presents difficulties in interpreting the results of pain measurement using the PM, and further study into its usefulness is therefore required. The women in this study reported significantly lower pain thresholds than the men and no significant correlation between fear of movement/(re)injury and pain threshold was found.

#### ACKNOWLEDGEMENTS

We would like to thank the Swedish Insurance Society, the Swedish Association of Insurance Medicine, and Capiro for financial support.

#### REFERENCES

1. Main CJ, Watson PJ. Psychological aspects of pain. *Man Ther* 1999; 4: 203-215.
2. Radanov BP, Sturzenegger M, Di Stefano G. Long-term outcome after whiplash injury. A 2-year follow-up considering features of injury mechanism and somatic, radiologic, and psychosocial findings. *Medicine (Baltimore)* 1995; 74: 281-297.
3. Price DD. Psychological and neural mechanism of the affective dimension of pain. *Science* 2000; 288: 1769-1772.
4. Melzack R,Coderre TJ, Katz J, Vaccarino AL. Central neuro-

- plasticity and pathological pain. *Ann N Y Acad Sci* 2001; 933: 157–174.
5. Fernandez E, Turk DC. Sensory and affective components of pain: separation and synthesis. *Psychol Bull* 1992; 112: 205–217.
  6. Bunketorp L, Lindh M, Carlsson J, Stener-Victorin E. The perception of pain and pain-related cognitions in subacute whiplash-associated disorders: its influence on prolonged disability. *Disabil Rehabil* 2006; 28: 271–279.
  7. Edwards RR. Individual differences in endogenous pain modulation as a risk factor for chronic pain. *Neurology* 2005; 65: 437–443.
  8. Crombez G, Vlaeyen JW, Heuts PH, Lysens R. Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain* 1999; 80: 329–339.
  9. van den Hout JH, Vlaeyen JW, Houben RM, Soeters AP, Peters ML. The effects of failure feedback and pain-related fear on pain report, pain tolerance, and pain avoidance in chronic low back pain patients. *Pain* 2001; 92: 247–257.
  10. Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. *Res Nurs Health* 1990; 13: 227–236.
  11. Lundeberg T, Lund I, Dahlin L, Borg E, Gustafsson C, Sandin L, et al. Reliability and responsiveness of three different pain assessments. *J Rehabil Med* 2001; 33: 279–283.
  12. Alstergren P, Forstrom J. Acute oral pain intensity and pain threshold assessed by intensity matching to pain induced by electrical stimuli. *J Orofac Pain* 2003; 17: 151–159.
  13. Chesterton LS, Barlas P, Foster NE, Lundeberg T, Wright CC, Baxter GD. Sensory stimulation (TENS): effects of parameter manipulation on mechanical pain thresholds in healthy human subjects. *Pain* 2002; 99: 253–262.
  14. Svensson E. Guidelines to statistical evaluation of data from rating scales and questionnaires. *J Rehabil Med* 2001; 33: 47–48.
  15. Altman DG, editor. *Practical statistics for medical research*. London: Chapman & Hall/CRC; 1991, p. 406–409.
  16. Bunketorp L, Carlsson J, Kowalski J, Stener-Victorin E. Evaluating the reliability of multi-item scales: a non-parametric approach to the ordered categorical structure of data collected with the Swedish version of the Tampa Scale for Kinesiophobia and the Self-Efficacy Scale. *J Rehabil Med* 2005; 37: 330–334.
  17. Svensson E. Application of a rank-invariant method to evaluate reliability of ordered categorical assessment. *J Epidemiol Biostat* 1998; 4: 403–409.
  18. Svensson E, editor. *Analysis of systematic and random differences between paired ordinal categorical data (dissertation)*. Göteborg: Göteborg University; 1993.
  19. Stener-Victorin E, Kowalski J, Lundeberg T. A new highly reliable instrument for the assessment of pre- and postoperative gynecological pain. *Anesth Analg* 2002; 95: 151–157, table of contents.
  20. Lund I, Lundeberg T, Kowalski J, Sandberg L, Budh CN, Svensson E. Evaluation of variations in sensory and pain threshold assessments by electrocutaneous stimulation. *Physiother Theory Pract* 2005; 21: 81–92.
  21. Chapman CR, Donaldson GW, Nakamura Y, Jacobson RC, Bradshaw DH, Gavrin J. A psychophysiological causal model of pain report validity. *J Pain* 2002; 3: 143–155.
  22. Lehofer M, Liebmann PM, Moser M, Schauenstein K. Nervousness and pain sensitivity: I. A positive correlation. *Psychiatry Res* 1998; 79: 51–53.
  23. Bunketorp L, Lindh M, Carlsson J, Stener-Victorin E. The effectiveness of a supervised physical training model tailored to the individual needs of patients with whiplash-associated disorders – a randomized controlled trial. *Clin Rehabil* 2006; 20: 201–217.
  24. Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. *Pain* 1983; 16: 87–101.
  25. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975; 1: 277–299.
  26. Gaston-Johansson F. Measurement of pain: the psychometric properties of the Pain-O-Meter, a simple, inexpensive pain assessment tool that could change health care practices. *J Pain Symptom Manage* 1996; 12: 172–181.
  27. Lundberg MKE, Styf J, Carlsson GC. A psychometric evaluation of the Tampa Scale for Kinesiophobia – from a physiotherapeutic perspective. *Physiother Theory Pract* 2004; 20: 121–133.
  28. Vlaeyen JW, Kole-Snijders AM, Boeren RG, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain* 1995; 62: 363–372.
  29. Hendiani JA, Westlund KN, Lawand N, Goel N, Lisse J, McNearney T. Mechanical sensation and pain thresholds in patients with chronic arthropathies. *J Pain* 2003; 4: 203–211.
  30. Voerman VF, van Egmond J, Crul BJ. Elevated detection thresholds for mechanical stimuli in chronic pain patients: support for a central mechanism. *Arch Phys Med Rehabil* 2000; 81: 430–435.
  31. Chesterton LS, Barlas P, Foster NE, Baxter GD, Wright CC. Gender differences in pressure pain threshold in healthy humans. *Pain* 2003; 101: 259–266.
  32. Lamberg L. Venus orbits closer to pain than Mars, Rx for one sex may not benefit the other. *JAMA* 1998; 280: 120–124.
  33. Sterling M, Jull G, Vicenzino B, Kenardy J. Sensory hypersensitivity occurs soon after whiplash injury and is associated with poor recovery. *Pain* 2003; 104: 509–517.
  34. Moog M, Quintner J, Hall T, Zusman M. The late whiplash syndrome: a psychophysical study. *Eur J Pain* 2002; 6: 283–294.
  35. Banic B, Petersen-Felix S, Andersen OK, Radanov BP, Villiger PM, Arendt-Nielsen L, et al. Evidence for spinal cord hypersensitivity in chronic pain after whiplash injury and in fibromyalgia. *Pain* 2004; 107: 7–15.
  36. Sheather-Reid RB, Cohen ML. Psychophysical evidence for a neuropathic component of chronic neck pain. *Pain* 1998; 75: 341–347.
  37. Taub N, Worsowics G, Gnatz S, Cifu D. Pain rehabilitation. 1. Definitions and diagnosis of pain. *Arch Phys Med Rehabil* 1998; 79: 49–53.
  38. Pearce JM. Polemics of chronic whiplash injury. *Neurology* 1994; 44: 1993–1997.
  39. Rhudy JL, Meagher MW. Fear and anxiety: divergent effects on human pain thresholds. *Pain* 2000; 84: 65–75.