Background: Although the prevalence of sacroiliac joint (SIJ) pain is relatively high (15 – 30%), there is no unambiguous reference standard to diagnose SIJ pain. Pressure tenderness in the SIJ region is used for diagnostic purposes, but the clinimetric properties of this procedure remain to be determined.

Objectives: The aim of this study is to determine the reliability of pain pressure threshold (PPT) measurements in the SIJ region and the difference in PPTs in the SIJ region between healthy volunteers and PPTs in patients with SIJ pain.

Study Design: Prospective cohort study.

Setting: Outpatient pain clinic VU University Medical Center.

Methods: Forty-one healthy volunteers and 31 patients diagnosed with SIJ pain were included. PPTs were obtained from 5 measurement points in the region of the SIJ with a pressure pain algometer using a standardized protocol. The inter-rater reliability of this method was calculated by means of the Intraclass Correlation Coefficients (ICC) of individual assessment performed by 2 individual raters of SIJs of healthy volunteers on both sides. PPTs of healthy volunteers were compared to those of the affected side in patients with SIJ pain.

Results: PPT measurement showed moderate to good inter-rater reliability (ICC 0.6 – 0.82). The median PPTs of 5 points was comparable for both sides in healthy volunteers (right: 8.5 kg/cm² [IQR 6.0 – 10.0]; left 8.3 kg/cm² [5.8 – 10.0]). Median PPTs for the affected sides of patients with SIJ pain were significantly lower compared to the same side of healthy volunteers (right: 2.4 kg/cm² [IQR 2.2 – 3.2, n = 15]; left: 2.5 kg/cm² [2.3 – 3.2, n = 16]; P < 0.001 for both sides).

Limitations: Only the SIJ on one side of was measured in patients with SIJ pain, where both sides would be desirable.

Conclusions: Pressure pain algometry appears to be a reliable method to establish differences in PPTs between healthy volunteers and patients with SIJ pain. The diagnostic accuracy of this test should be investigated further.

Key words: Sacroiliac joint pain, pain pressure threshold, pressure algometry
The sacroiliac joint (SIJ) is a well-known source of low back pain, with a prevalence in a range of 10 to 62% depending on the diagnostic test used (1-3). However, there are no historical, physical, or radiological features which can definitely confirm the diagnosis of SIJ pain. According to the International Association for the Study of Pain (IASP) (4), SIJ pain arises from the anatomical area of the SIJ. It should be reproducible by performing specific pain provocation tests, or completely relieved by infiltration of the painful SIJ with local anesthetics. Unfortunately, there are many provocation tests described, but only a few have discriminative power to diagnose SIJ pain. Regarding the diagnostic blockades, the target structures are vague. To differentiate between intra- and periarticular structures can be complex for untrained physicians. Until now, little attention has been paid to the diagnostic power of the presence of pain in the region of the SIJ. According to the literature, pain arising from the SIJ, overlies the posterior aspect of the SIJ (5). However, radiation to the buttock, groin, and lower extremities has also been reported (5,6). Some investigators postulate that patients with presumed SIJ pain, point out the area adjacent to the posterior superior iliac spine (7). Based on the results of another study, the most intense pain area in patients with SIJ pain overlies the posterior margin of the SIJ (8). Also pressure in this area (i.e., the sacral sulcus tenderness) seems to be a very sensitive, although not a specific test for diagnosing SIJ pain (9). Although a pressure test is very simple, it is in fact not valid, simply because we cannot standardize the pressure exercised with our finger. Therefore in this study we wanted to standardize the pain pressure thresholds (PPTs) in the SIJ region by using a pressure algometer. The use of pressure algometry has already been shown to be reliable in healthy volunteers in different points of the body (10), but also in patients with non-specific low back pain (11) and in patients with pain in the lumbosacral and gluteal region (12).

METHODS

The study was performed in the outpatient pain clinic of the VU University Medical Center in Amsterdam. First, we established the inter-rater reliability of the pressure algometer, as well as the PPTs in healthy volunteers. For this purpose, we recruited healthy volunteers between the ages of 20 to 60, without a history of low back pain for at least 6 months prior to the assessment. The exclusion criteria were pregnancy, history of trauma and/or back surgery, and malignancy. Informed consent was provided before the procedure.

Two examiners measured the PPTs in the SIJ region on both sides. To that purpose, 5 measure points were marked on the skin on both sides of healthy volunteers based on descriptions of SIJ pain localizations found in the literature (5). The first examination point was marked on the skin 1 cm medially and caudally from the posterior superior iliac spine and 4 more 2 cm laterally, medially, cranially, and caudally from the first one. Anatomically, the second, i.e., lateral point, was located nearby the posterior superior iliac spine at the attachment of gluteus maximus muscle to the iliac crest. The third (2 cm cranially) and the fourth point (2 cm medially) overlaid the erector spinae muscle and the deeper located posterior sacroiliac ligament. The fifth, caudal point, was located at the attachment of gluteus maximus muscle to facies posterior of the sacrum and posterior sacroiliac ligament.

PPTs were measured using an analogue mechanical pressure algometer (Wagner Force Dial TM FDK 40, Greenwich, CT) dynamometer with a flat circular compression rubber tip (1 cm²). The dial ranges from 2.0 to 20.0 kgf (kilogram-force), in increments of 0.2 kgf. The value of the kilogram-force is equal to kilogram per square centimeter (kg/cm²). Pressures up to 2 kgf cannot be measured because of the start point of the applied force is pointed on the shield of the algometer, and the pointer holds the maximal applied force until the pointer is reset. Prior to the start of the study, the algometer was calibrated by the instrument maintenance department. The examiners were trained in performing the pain pressure assessments.

The PPTs were independently assessed by 2 examiners. The first examiner started with the right side and applied a constant axial force on each point until the participant reported pain. This was repeated at the left side. Specific instructions to the participants were to report the first sensation of pain. After 5 minutes, a second examiner repeated the measurements, without knowledge of the scores of the first assessor. After 10 participants, the order of examiners was changed. The participants received no information during the examination regarding the PPT recordings.

Subsequently, we included patients with SIJ pain in the age between 18 and 70 years to establish their PPTs, and compare them to the results obtained in healthy volunteers. Patients met the SIJ pain diagnostic criteria according to the IASP criteria (4). Exclusion criteria included ankylosing spondylitis, pregnancy, radicular
pain, recent trauma, recent back surgery, malignancy, and patient’s refusal.

A patient’s history was gathered by means of a standardized questionnaire. In the questionnaire, patients were asked to describe their radiation pattern of the pain and how the pain influenced their lives. Intensity of pain was measured with an 11 point numeric rating scale (NRS)-score with the anchors *0 = no pain and 10 = the worse imaginable pain (13). Physical examination was performed after the questionnaire was filled out.

Subsequently, the procedure for measuring the PPT as described for healthy volunteers was applied in patients with SIJ pain, except that the SIJ pain patients were only tested on the painful side, and only by one examiner. As in the group of healthy volunteers, we asked the patients to report the first sensation of pain provoked by the exerted pressure on the predefined point.

Before the start of the study verbal information was provided and written informed consent was acquired. The requirement for medical-ethical evaluation was waived by the medical ethics committee of our institution as this study fell out of the scope of the law on medical research.

Data Analysis

For data collection and analysis, SPSS database version 15 was used. The mean of the 5 PPTs was calculated for each side per participant. These mean values were tested with the Shapiro-Wilk test for a normal distribution. Since the data were not normally distributed, medians and inter quartile ranges (IQR) will be used. The median PPTs obtained in SIJ patients were compared with the median PPTs in healthy volunteers. Differences between assessments were tested using the Kruskal-Wallis test and Mann-Whitney test. In all cases, a P < 0.05 was used as a cutoff to indicate statistical significance.

The inter-rater reliability was analyzed using the Intraclass Correlation Coefficient for agreement (ICC agreement). We tested the ICC for agreement between the examiners, using a single measure, 2 way random model (14). The inter-rater reliability was estimated for each of the 10 measurement points, and for scores derived from the mean scores of the left and right side. The ICC can vary between 0 and 1.0, where 0 indicates no reliability and 1.0 indicates perfect reliability. The following cut-off points for the ICC were adopted: < 0.5 poor reliability, 0.5 – 0.75 moderate reliability, and > 0.75 good reliability (15).

Based on the agreement between the raters, the standard error of measurement (SEM) was calculated from the square root of the error variance, consisting of random error and systematic difference between examiners (14). The SEMagreement quantifies the precision of individual scores of the PPT, providing the absolute index of the measurement error expressed in the same units as the measurement of interest, in the case of this study kgf. Finally, the SEMagreement was used to calculate the smallest detectable change (MDC), i.e., the change between the consecutive measurements exceeding 95% confidence interval: (1.96*√2*SEMagreement) (11), in our case the measurement error calculated for 2 examiners (16). Bland-Altman plots were used to visualize the limits of agreement, systematic differences between the measurements performed by the 2 independent raters, and their distribution around zero.

Results

Healthy Volunteers

Forty-one healthy volunteers (19 women and 22 men) in the mean age of 35.6 (SD 11.1) years were examined. The median PPT on the right side was 8.5 kgf (IQR 6.0 – 10.0) and on the left side 8.3 kgf (IQR 5.8 – 10.0). The difference between both sides was not statistically significant (P = 0.797).

The median PPT in women on the right side was 6.7 kgf (IQR 5.0 – 8.5), and on the left side 7.3 (IQR 5.1 – 8.3). In men, the median PPT was significantly higher (P = 0.004) for both sides than in women; 10.3 (IQR 7.3 – 11.4), and 9.2 (IQR 7.4 – 11.2) for the right and left side, respectively.

The reliability of the mean pressure was ICC 0.79 (95% CI 0.65 – 0.88) for the right side and ICC 0.82 (95% CI 0.67 – 0.9) for the left side. The reliability for each measured point varied between ICC 0.60 and 0.82. The lowest reliability was found for the fourth point and the highest for the second point for right as well as the left side. Overall, the reliability was higher for measurement points on the left side, except for the fourth point.

Thereafter, the SEM was calculated. For the left side, the SEM was 1.26 kgf, and for the right side it was 1.38 kgf. The MDC was for the left side 3.49 and 3.83 for the right side. Bland and Altman plots (Fig. 1) show no systematic bias (i.e., consistent higher or lower scores for one of the raters and limited influence of measurement magnitude).
SIJ Pain Patients

Thirty-one SIJ patients (11 men and 20 women), in the mean age of 55.6 years (SD 10.4 year) were included in a period of 3 months. According to age, there was no statistical difference between SIJ pain patients and healthy volunteers ($P = 0.517$). Nine patients (29.0%) pointed out the right side as painful, 8 (25.6%) the left side, and 14 (45.2%) had pain in both SIJ regions. Thirteen patients (41.9%) reported pain radiating to the buttock, in 5 patients (16.1%) the pain radiated to the groin, and 28 (90.3%) to the leg. Twenty-eight patients (90.3%) noted that the pain was continuously present.

Table 2 shows the activities influencing SIJ pain. With the exception of lying down, all activities increased the pain. The NRS-scores varied from 5 to 10, with a median of 8 (IQR 7 – 9).

The PPTs as measured in patients ranged from 2.0 to 7.2 kgf. The median was 2.5 (IQR 2.28 – 3.24). The median scores of the PPTs at each of the 5 predefined points are detailed in Table 2. In the first measurement point the highest median of the PPT was achieved (2.6 [IQR 2.4 – 3.6]), and the second measurement point showed the lowest median PPTs (2.4 [IQR 2.0 – 2.8]). Most patients indicated the third measurement point as the most painful, whereas the fourth point was never pointed out as the most painful by any patient (see Table 3).

The median PPT in patients with SIJ pain on the right side was 2.4 (2.2 – 3.2, $n = 15$), and 2.5 (2.3 – 3.2, $n = 16$) for the patients with the SIJ pain on the left side ($P = 0.417$). For men the median of the mean PPT was 3.2 (2.3 – 4.5) and for women 2.4 (IQR 2.3 – 2.6) ($P = 0.116$).
Compared to the median PPTs obtained in healthy volunteers, we found significantly lower PPTs for both sides ($P_{\text{left}} < 0.001$, $P_{\text{right}} < 0.001$); the difference was 5.6 kgf for the right side and 5.5 for the left side. Compared to healthy volunteers, the median PPTs of patients fell outside of the limits defined by the MDC.

### Discussion

The aim of this study was to evaluate the reliability of PPTs in the SIJ region for healthy volunteers and to compare PPTs between healthy volunteers and patients with SIJ pain. The inter-rater reliability of assessing PPTs was found to be moderate to good with no apparent systematic bias related to rater or magnitude of the measurement. For patients with SIJ pain, we found a significantly lower PPT in the SIJ region than in healthy volunteers.

As noted above, in healthy volunteers, we found a moderate to good inter-rater-reliability for manual algometry in the SIJ region, which is comparable with the results in previous studies (17-19). These findings seem to be consistent across different studies, despite differences in design and target locations. For instance, a study by Antonaci et al (17) found an ICC of 0.75 in head and neck algometry, in another study by Aweid et al (20), an ICC of 0.77 – 0.90 on the medial tibia was found, and in a study coordinated by Nussbaum and Downes (18) showed an ICC of at least 0.9 in pressure algometry in the biceps brachii muscle. However, it should be noted that measuring the ICC for multiple points led to a more reliable PPT pattern as compared to single PPT point assessment; therefore, it is advisable to assess a combination of measurement points in the SIJ region.

### Table 2

Patients were asked to answer if the activity worsens the pain or make it less painful. Except lying down, all activities were noted to worsen the pain. All patients noted that a long time standing worsened the pain.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Number of patients (percentage)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Worse</td>
</tr>
<tr>
<td>Lay down</td>
<td>5 (16.1%)</td>
</tr>
<tr>
<td>Standing up</td>
<td>26 (83.9%)</td>
</tr>
<tr>
<td>Long time standing</td>
<td>31 (100%)</td>
</tr>
<tr>
<td>Sitting</td>
<td>20 (64.5%)</td>
</tr>
<tr>
<td>Walking</td>
<td>26 (83.9%)</td>
</tr>
<tr>
<td>Coughing/sneezing</td>
<td>20 (64.5%)</td>
</tr>
<tr>
<td>Oral medication</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

### Table 3

The median at each predefined point is shown. The highest median of the PPT was achieved (2.6 [IQR 2.4 – 3.6]), and the second measure point showed the lowest median PPTs (2.4 [IQR 2.0 – 2.8]). Most patients (40%) noted that the third point was the most painful.

<table>
<thead>
<tr>
<th>Point</th>
<th>Median in kgf (IQR)</th>
<th>Frequency of most painful point</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.6 (2.4 – 3.6)</td>
<td>4 (13.3%)</td>
</tr>
<tr>
<td>2</td>
<td>2.4 (2.0 – 2.8)</td>
<td>9 (30.0%)</td>
</tr>
<tr>
<td>3</td>
<td>2.4 (2.0 – 3.0)</td>
<td>12 (40.0%)</td>
</tr>
<tr>
<td>4</td>
<td>2.5 (2.0 – 3.0)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>5</td>
<td>2.4 (2.2 – 3.1)</td>
<td>5 (16.7%)</td>
</tr>
</tbody>
</table>

After measuring PPTs in healthy volunteers, we tested the PPTs in SIJ pain patients. According to the literature, low PPTs found in the low back region are associated with low back pain (21). We found that the PPT values determined in patients with SIJ pain are much lower than the MDC for measuring PPTs in the SIJ region, and our results are not a consequence of a measurement error.

As there are no clear cut-off criteria for PPTs of the SIJ, we have searched the literature for clinical relevance of PPTs in general. According to a study by Fischer (10), who bilaterally measured PPTs at several points at the different parts of the body (i.e., the deltoid muscle, the upper trapezius, paraspinal at level L4, and at the gluteus medius), a difference greater than 2.0 kg/cm² between the measurement sides in the same patient should be considered as a pathological finding. In our study we did not measure both sides in most patients.
with SIJ pain. Instead, we compared the PPTs with the results found in healthy volunteers. As the differences between healthy volunteers and patients were higher than 2.0 kgf, our data may provide an indication of the clinical relevance of our findings. Although different cut off levels should be determined and used for different body areas (22), it is found that in general PPTs equaling 3 kgf or less could be considered as abnormally low (23), which was the case in our study regarding the findings in SIJ pain patients. Nonetheless, considering the difference in design (comparing PPTs with healthy volunteers instead of comparing with the other side in the patient), the clinical relevance of our results need to be interpreted with caution.

The differences between SIJ pain patients and healthy volunteers may be explained by the fact that SIJ pain may not only originate from the articulation and its cartilage (24), but also from periarticular ligaments (24,25). Nociceptive peptides such as Substance P in the periarticular tissue may provide an indication of the nociceptive role of these structures. Substance P has repeatedly been associated with the transmission of pain and “neurogenic inflammation” (26). In the study by Fortin et al (26), it was shown that substance P is present in the dorsal ligaments also, and that pain in the dorsal ligaments can originate as SIJ pain. Also Cohen (27) reported that SIJ painful pathology not only involves intra-articular causes, but also periarticular causes like ligamentous injury, enthesopathy, or myofascial pain.

Also, an interesting point of discussion is the fact that women in the healthy volunteer group demonstrated significantly lower PPTs compared to men for all measured points. Gender difference in pain perception and sensitivity to experimental pain stimuli has already been well established, whereby lower PPTs in women were shown for different noxious stimuli and anatomical regions (28). This difference was not found in the SIJ pain patients, which can probably be explained by very low PPTs for both male and female patients in the SIJ pain group, limiting the possibility of finding differences within this group. Also influence of a sample size cannot be ruled out at this point.

This study has some limitations. The first limitation is the sample size, which may have led to type I or type II errors. The second limitation of this study is the use of the Wagner Force Dial FPK algometer, which measures a pressure from 2.0 kgf upwards. This means that all patients have at least a PPT of 2.0. The examiner occasionally experienced that some patients already noted their PPT before the pointer moved up above a PPT of 2.0 kgf. This would mean that the results presented in this study could represent an underestimation of the actual difference in PPTs between pain patients and healthy volunteers. For clinical practice, it would be recommended to use an algometer with a lower detection threshold.

Although the PPT measurement protocols in healthy volunteers and in patients with SIJ pain were identical, the demographic characteristics differ between both groups, which can be seen as a third limitation. The age of the healthy volunteers is lower (35.6 years [SD 11.1]) than the age of the SIJ patients (55.6 [SD 10.4]), (P = 0.517). Although this difference is not statistically significant, age should not be ruled out as a possible confounder at this point, as a previous study showed that pain tolerance decreases with increasing age (29). Furthermore, psychological parameters influencing pain were not assessed in this patient group, but may have influenced pain perception. From different studies it is known that state and trait anxiety as well as catastrophizing have a strong relationship with experienced severity of pain (30).

Based on these data, the next step would be to determine the value of pressure algometry in SIJ pain patients. The PPTs for the SIJ have apparent face validity since it can be used to measure a parameter (pressure pain) that is considered to be indicative for SIJ pain in clinical practice. Also, the procedure appears to have discriminative validity in a sense that differences between healthy volunteers and SIJ pain patients can be observed. However, other indices of validity, such as construct, concurrent, criterion, and prognostic validity, need to be established in future research. This would include testing the diagnostic accuracy of the instrument compared to other available instruments (reference standard) and to assess change within the context of the course of the disease of treatment. Unfortunately, a gold standard for SIJ pain assessment is not available to compare PPTs with.

Different methods have been proposed as a gold standard, but sensitivity and specificity remain limited. According to the IASP criteria, positive reactions on physical provocation tests and diagnostic blockades could determine SIJ pain. Though a variety of physical tests have been evaluated, only a combination of tests appears to be useful to pinpoint the SIJ as the origin of the pain (31). Best evidence suggests that the tight trust test, the compression test, or a combination of 3 tests have discriminative power for diagnosing SIJ pain (32).
The value of diagnostic blockades has been studied substantially, however, in a best evidence review by Rubinstein and Van Tulder (33), only moderate evidence for diagnostic blockades in the evaluation of spinal pain was found. A systematic review by Simopoulos et al (34) showed a good diagnostic accuracy of diagnostic blockades, but the diagnostic value of SIJ infiltration with local anesthetics remains controversial in light of the potential for false-positive and false-negative results (35).

The clear relationship observed in our study between the presence of pain as measured with the NRS (which was moderate to severe in our patient sample) and low PPTs provides some indication for concurrent validity between both assessments, but this needs to be evaluated further in a more specific manner.

In addition, patients with established other causes of low back pain should be tested for their PPT in the SIJ area, in order to assess the discriminative ability of this test. Also patients with sacral herniated nuclei pulposi (HNP) should be investigated. The SIJ is mainly innervated by the sacral rami dorsales (36) and a HNP in the area of these roots may cause pain in the same region as SIJ pain.

In addition, structures other than the (periarticular) SIJ can give rise to pain in the referral area of the SIJ, such as ligamentous injury, enthesopathy, fractures, and myofascial pain (18). For clinical assessment, radiological and laboratory screening may be required to exclude possible red flags in the SIJ joint region.

**Conclusion**

In conclusion, pressure pain algometry appears to be a reliable method to establish differences in PPTs between healthy volunteers and patients with SIJ pain. Further research has to be performed to identify the possibilities of pressure pain algometry in investigating and diagnosing SIJ pain.

**References**


12. Altman DG. Practical Statistics for Medical Research. 1999. (Need publisher and city.)


