

Original article

# Diagnosis of Sacroiliac Joint Pain: Validity of individual provocation tests and composites of tests

Mark Laslett<sup>a,\*</sup>, Charles N. Aprill<sup>b</sup>, Barry McDonald<sup>c</sup>, Sharon B. Young<sup>d</sup>

<sup>a</sup>Department of Health and Society, Linköpings Universitet, Linköping, Sweden

<sup>b</sup>Magnolia Diagnostics, New Orleans, LA, USA

<sup>c</sup>Massey University, Institute of Information and Mathematical Sciences, Albany, New Zealand

<sup>d</sup>Mobile Spine and Rehabilitation Center, Mobile, AL, USA

Received 11 December 2002; received in revised form 1 September 2004; accepted 4 January 2005

## Abstract

Previous research indicates that physical examination cannot diagnose sacroiliac joint (SIJ) pathology. Earlier studies have not reported sensitivities and specificities of composites of provocation tests known to have acceptable inter-examiner reliability. This study examined the diagnostic power of pain provocation SIJ tests singly and in various combinations, in relation to an accepted criterion standard. In a blinded criterion-related validity design, 48 patients were examined by physiotherapists using pain provocation SIJ tests and received an injection of local anaesthetic into the SIJ. The tests were evaluated singly and in various combinations (composites) for diagnostic power. All patients with a positive response to diagnostic injection reported pain with at least one SIJ test. Sensitivity and specificity for three or more of six positive SIJ tests were 94% and 78%, respectively. Receiver operator characteristic curves and areas under the curve were constructed for various composites. The greatest area under the curve for any two of the best four tests was 0.842. In conclusion, composites of provocation SIJ tests are of value in clinical diagnosis of symptomatic SIJ. Three or more out of six tests or any two of four selected tests have the best predictive power in relation to results of intra-articular anaesthetic block injections. When all six provocation tests do not provoke familiar pain, the SIJ can be ruled out as a source of current LBP.

© 2005 Elsevier Ltd. All rights reserved.

**Keywords:** Sacroiliac joint; Low back pain; Physical examination; Diagnosis; Validity; Sensitivity; Specificity

## 1. Introduction

The sacroiliac joint (SIJ) can be a nociceptive source of low back pain (Fortin et al., 1994a, b; Bogduk, 1995). SIJ pain has no special distribution or features and is similar to symptoms arising from other lumbosacral structures. There are no provoking or relieving movements or positions that are unique or especially common to SIJ pain (Dreyfuss et al., 1996; Fortin et al., 1994a, b; Schwarzer et al., 1995; Maigne et al., 1996; Fortin and Falco, 1997). The clinical diagnosis of symptomatic SIJ

remains problematical, but the ability to make the diagnosis is an important objective. It may be presumed that treatment strategies for SIJ lesions should differ from strategies intended to relieve and treat pathologies of other structures such as disk, nerve root or facet joint pain. Without a readily accessible means of differentiating between these possible sources of pain, treatment strategies are perforce non-specific, and likely to have at best, modest efficacy.

At present, a current acceptable method of confirming or excluding the diagnosis of a symptomatic SIJ is fluoroscopically guided, contrast enhanced intra-articular anaesthetic block (Fortin et al., 1994b; Grieve, 1988; Merskey and Bogduk, 1994; Schwarzer et al., 1995; Sakamoto et al., 2001; Adams et al., 2002). While

\*Corresponding author. Department of Health and Society, Linköpings University, Auckland, New Zealand. Tel.: +649 626 0015.  
E-mail address: [mark.laslett@xtra.co.nz](mailto:mark.laslett@xtra.co.nz) (M. Laslett).

certain SIJ tests have been shown to have acceptable inter-rater reliability (Laslett and Williams, 1994; Kokmeyer et al., 2002), current evidence suggests that these tests alone cannot predict the results of a criterion standard such as diagnostic injection (Dreyfuss et al., 1996; Maigne et al., 1996; Slipman et al., 1998). These reports have not reported the sensitivity, specificity or likelihood ratios or provided data on the diagnostic power of individual or composites of provocation SIJ tests (Slipman et al., 1998). However, in a previous publication, the current authors have identified a composite of three provocation SIJ tests in the absence of centralization during repeated movement testing has clinically useful sensitivity, specificity and positive likelihood ratio (93%, 89% and 6.97%, respectively) (Laslett et al., 2003).

Conceptually, it seems reasonable to propose that stress testing of the SIJ should provoke pain of SIJ origin. However, clinical stress tests are unlikely to load the targeted structure alone. Herein lies the problem. When a test provokes familiar pain, the question arises if this is evidence of pathology within the targeted structure, or evidence of pathology in a different but nearby structure that is also stressed at the same time. However, if different stress tests of a structure provoke pain, greater diagnostic confidence may result. The use of composites of tests is common in musculoskeletal medicine. When a straight leg raise test provokes familiar leg pain, nerve root irritation from a herniated lumbar disc may be suspected. However, pain, paraesthesiae or skin anaesthesiae in a known segmental distribution, weakness of key muscles or reflexes must also be present before the diagnosis of a herniated lumbar disc can be made with any degree of confidence. Confirmation with computed tomography or magnetic resonance imaging completes the composite of tests for this diagnosis. The need for diagnostic research to investigate the added value of composites of tests within the diagnostic process has been emphasized (Deville et al., 2000; Grieve, 1988). This current study explored the utility of utilizing composites of SIJ provocation tests to predict the results of fluoroscopically guided, contrast enhanced SIJ blocks (diagnostic injection).

## 2. Material and methods

The diagnosis of symptomatic SIJ pathology may mean that either SIJ structures contain the pain generating tissues, or that the SIJ functions or malfunctions in such a way as to cause pain. Throughout this report, references to symptomatic SIJ, SIJ pain or pathology are confined to meaning that the pain originates from the SIJ structures.

The study design is presented graphically in Fig. 1. Physiotherapists (ML and SBY) visited a private

radiology practice in New Orleans specializing in the diagnosis of spinal pain at regular intervals between January 1997 and August 1998 to carry out the clinical evaluations. Patients were not consecutive. Patients deemed likely by clinic staff to have SIJ pain were scheduled to receive the SIJ provocation tests on a day when an examining physiotherapist visited the clinic. The clinical examination and injection procedures were completed the same day. No other treatment was provided by the physiotherapist. The physiotherapists were blinded to the results of previous diagnostic injections and the results of previous imaging studies. Diagnostic injection was conducted blind from the results of SIJ provocation tests and the results of the physiotherapy examination. Results from the clinical and SIJ injection procedures were recorded on separate standardized data collection forms. Informed consent was sought prior to the clinical evaluations.

### 2.1. Inclusion criteria

Patients with buttock pain, with or without lumbar or lower extremity symptoms were invited to participate in the study. Patients were scheduled for the clinical evaluation in an opportunistic fashion with some patients being examined by the physical therapist at their initial visit to the clinic and others scheduled to return on a day when the physical therapist was present. Each patient had undergone imaging studies and had a variety of unsuccessful therapeutic interventions. They were referred for diagnostic evaluation and procedures by a variety of medical and allied health practitioners and a few were self-referred. Patients were drawn from the New Orleans metropolitan area, with some intrastate and interstate referrals.

### 2.2. Exclusion criteria

Patients were excluded from the study if they were unwilling to participate, had only midline or symmetrical pain above the level of L5, had clear signs of nerve root compression (complete motor or sensory deficit), or were referred for specific procedures excluding SIJ injection. Those deemed too frail to tolerate a full physical examination, were also excluded.

### 2.3. Background data collection

Patient data recorded included age, gender, occupation, employment status, pending litigation, duration of symptoms, aggravating/relieving factors and cause of current episode. The patient completed detailed pain drawings (Ohnmeiss et al., 1999; Beattie et al., 2000) and pain intensity was measured on a verbal analogue scale (VAS) (0 = "no pain" and 10 = "worst imaginable pain"). Disability was estimated with the

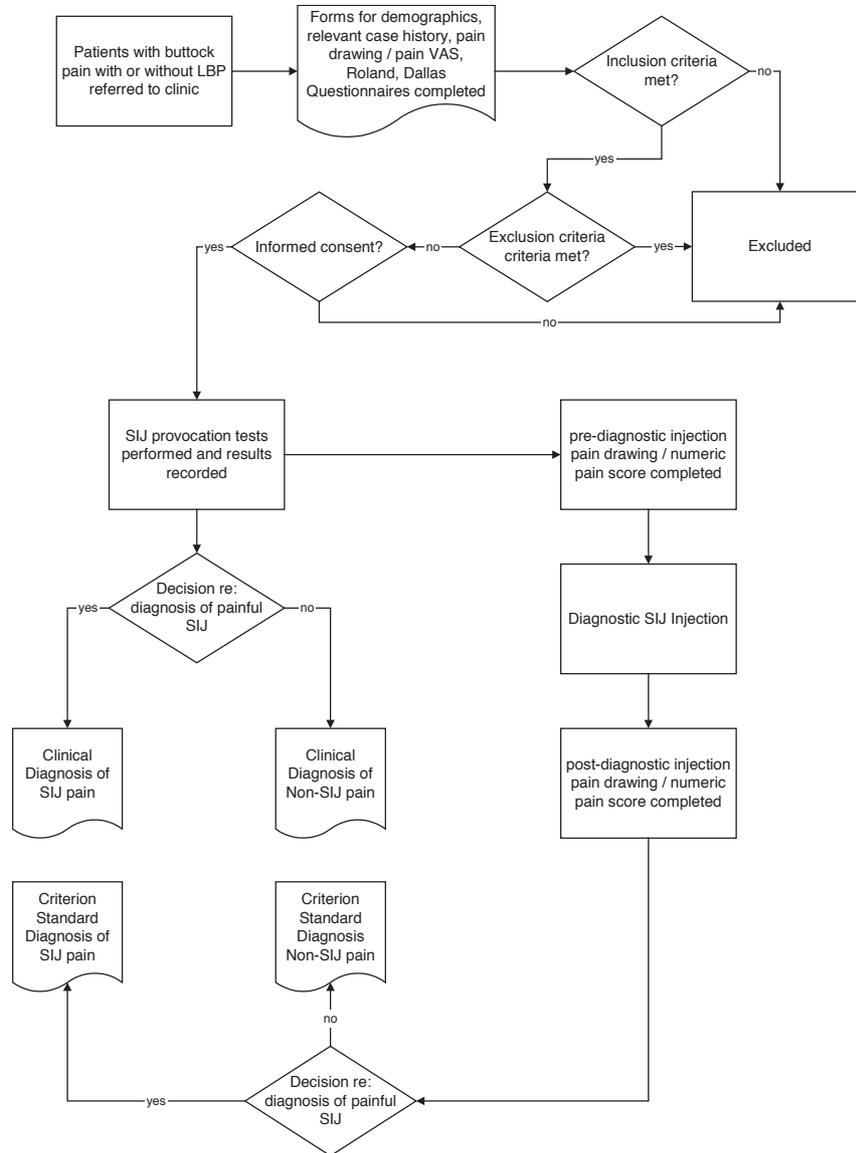


Fig. 1. Flow diagram of study protocol.

Roland–Morris questionnaire (Roland and Morris, 1983; Jensen et al., 1992) and the Dallas Pain and Disability questionnaire (Lawlis et al., 1989).

#### 2.4. Operational definitions

*The familiar symptom:* The familiar symptom is the pain or other symptoms (such as aching, burning, paraesthesiae or numbness) identified on a pain drawing, verified by the patient as being the complaint that has led the patient to seek diagnosis and treatment. During a diagnostic test the familiar symptoms must be distinguished from other symptoms produced by the test, and may be produced, increased, decreased or abolished.

*Choice of SIJ tests to evaluate:* Tests based on palpation for positional faults or movement dysfunc-

tions were not considered for inclusion in the current study, since adequate inter-examiner reliability has not been demonstrated in earlier studies (Potter and Rothstein, 1985; McCombe et al., 1989; Meijne et al., 1999). However, one study found that a selection of pain provocation tests were found to have acceptable reliability (Cohen's Kappa >0.04) (Laslett and Williams, 1994) and these were considered as suitable procedures for evaluation of diagnostic validity.

*Positive provocation SIJ test:* A provocation SIJ test that produces or increases familiar symptoms.

*Negative provocation SIJ test:* A provocation SIJ test that does not produce or increase familiar symptoms.

*Positive SIJ injection:* Slow injection of solutions provokes familiar pain, and instillation of a small volume of local anaesthetic (less than 1.5 cc) resulted in 80% or more relief of the pain for duration of effect

of the anaesthetic agent. Anaesthetic effect was assessed by change in pre- and post-injection numeric pain rating scales. Patients reporting a concordant pain response and at least 80% relief of their familiar pain were scheduled for a confirmatory block. Lidocaine was used in the initial injection and Bupivacaine was used in the confirmatory block to eliminate the need for a sham injection (Barnsley et al., 1993).

*Negative SIJ injection:* Diagnostic injections were considered indeterminate when there was a concordant pain response but insufficient pain relief, or when substantial pain relief was reported in the absence of provocation of familiar pain. Indeterminate responses were considered negative for statistical analysis. Injections not causing concordant pain provocation or analgesic response were deemed negative.

### 2.5. Clinical evaluation

The clinical evaluation was carried out by physiotherapists with 25 years (ML) and 17 years (SBY) experience in orthopaedic examinations of spinal pain patients and included a standard history and structured physical examination lasting between 30 min and 1 h. The structured physical examination included a McKenzie examination of the lumbar spine (McKenzie, 1981), SIJ provocation tests (Laslett and Williams, 1994), and a hip joint assessment (Cyriax, 1975).

*The sacroiliac pain provocation tests:* The tests employed in this study were: distraction (Fig. 2), right sided thigh thrust (Fig. 3), right sided Gaenslen's test (Fig. 4), compression (Fig. 5) and sacral thrust (Fig. 6) and have acceptable inter-rater reliability (Kokmeyer et al., 2002) and have been described previously (Cyriax,

1975; Laslett and Williams, 1994; Maigne et al., 1996; Laslett et al., 2003).

### 2.6. Radiology examination

The technique used for fluoroscopically guided contrast enhanced SIJ arthrography has been previously described (Fortin et al., 1994b; Schwarzer et al., 1995). The radiologist examiner (CA) has over 20 years experience in diagnostic spinal injection procedures, including SIJ injection. The SIJ injection was given within 30 min of completion of the physiotherapy clinical examination. Pain drawings and numeric pain rating scales for pain intensity were acquired prior to and 30–60 min following diagnostic injection.

During this study, corticosteroid was introduced into the joint as a therapeutic procedure when the initial injection of contrast and Lidocaine provoked familiar symptoms, as was normal practice at the clinic. The radiologist documented the procedure(s) performed, radiographic findings and conclusions. Pain provocation and analgesic responses to SIJ injection were recorded.

### 2.7. Data reduction and analysis

Statistical calculations were performed using statistical software Minitab (version 13.31 Minitab Inc. ©2000), and CIA (version 2.0.4 © Trevor N Bryant, 2000 University of Southampton) (Bryant, 2000). Two by two contingency tables were constructed and sensitivity, specificity, positive and negative predictive values, and likelihood ratios with 95% confidence intervals were calculated for each test independently and for composites of SIJ tests. Sensitivity, specificity,



Fig. 2. Distraction provocation SIJ test.



Fig. 3. Thigh thrust SIJ provocation test.



Fig. 4. Gaenslen's provocation SIJ test (right sided test).

positive and negative predictive values have been calculated using the Wilson method (Altman et al., 2000; Bryant, 2000). Likelihood ratios have been calculated using the score method (Altman et al., 2000). Receiver operator characteristic (ROC) curves are an overall measure of diagnostic efficacy (Altman et al., 2000, pp. 111–116). These curves combine sensitivity and specificity, and the area under the curve (AUC) is a summary measure of achieved discrimination, with perfect discrimination represented by an AUC of 1.0, and scores equal to or less than 0.5 are equivalent to or

worse than can be expected by random chance. The closer the AUC approaches 1.0, the better discriminatory power the diagnostic test has in relation to the criterion or reference standard.

### 3. Results

Sixty-two patients agreed to participate and were examined by both radiologist and physical therapist. Of these patients, three were unable to tolerate the physical



Fig. 5. Compression provocation SIJ test.



Fig. 6. Sacral thrust provocation SIJ test.

examination, two were pain free on the day of the clinical assessment, seven had no SIJ injection, and two had a bony obstruction causing a technical failure to inject the SIJ. These patients were excluded from the study. Forty-eight patients satisfied all inclusion criteria. Twenty-seven patients received the clinical assessment at their first clinic visit, 21 patients at the second.

There were no significant differences between positive and negative responders to diagnostic injection with regards to age, gender, working status, Dallas and Roland questionnaire results or pain intensity prior to

examination. [Table 1](#) presents basic demographic, and disability data for all included patients.

Of the 48 patients satisfying inclusion criteria 16 patients had positive SIJ injections. There were no adverse effects reported by patients from either the physical examination or SIJ injection, other than temporary local soreness at the injection site or increase in discomfort from the clinical examination.

The provocation SIJ tests provoked familiar pain in those patients confirmed by diagnostic injection as having painful SIJ pathology more commonly than

Table 1  
Patient characteristics ( $n = 48$ )

	Number	%		
Female	32	66.7		
Male	16	33.3		
	Mean	Median	SD	Range
Age (years)	42.1	42.0	12.3	20–79
Symptom duration (months)	31.8	22	38.8	2–156
Off work (months)	17.8	19	33.4	2–84
Questionnaire	Mean %	Median %	SD %	Range %
Roland–Morris ( $N = 42$ )	75.7	82.5	21.6	22–100
Dallas pain and disability ( $N = 48$ )				
Daily activities interference	61.2	66.0	18.0	23–93
Work leisure interference	66.2	75.0	22.6	14–95
Anxiety/depression	54.3	55.0	27.2	0–100
Social interference	48.7	50.0	24.8	0–85

Table 2  
Prevalence, sensitivity, specificity and likelihood ratios for individual SIJ provocation tests

	Distraction	Compression	Thigh thrust	Gaenslen's (right)	Gaenslen's (left)	Sacral thrust
Prevalence of positive test	31.9%	43.8%	50.0	37.0%	31.1%	37.5%
Sensitivity	0.60	0.69	0.88	0.53	0.50	0.63
95% CI	0.36, 0.80	0.44, 0.86	0.64, 0.97	0.30, 0.75	0.27, 0.73	0.39, 0.82
Specificity	0.81	0.69	0.69	0.71	0.77	0.75
95% CI	0.65, 0.91	0.51	0.82	0.53, 0.84	0.60, 0.89	0.58, 0.87
PPV	0.60	0.52	0.58	0.47	0.50	0.56
95% CI	0.36, 0.80	0.32, 0.72	0.39, 0.76	0.26, 0.69	0.27, 0.73	0.34, 0.75
NPV	0.81	0.82	0.92	0.76	0.77	0.80
95% CI	0.65, 0.91	0.63, 0.92	0.74, 0.98	0.58, 0.88	0.60, 0.89	0.63, 0.91
+LR	3.20	2.20	2.80	1.84	2.21	2.50
95% CI	1.42, 7.31	1.18, 4.09	1.66, 4.98	0.87, 3.74	0.95, 5.00	1.23, 5.09
–LR	0.49	0.46	0.18	0.66	0.65	0.50
95% CI	0.24, 0.83	0.20, 0.87	0.05, 0.55	0.34, 1.09	0.34, 1.03	0.24, 0.87

Notes: PPV = positive predictive value, NPV = negative predictive value, +LR = likelihood ratio for positive test, –LR = likelihood ratio for negative test, 95% CI = 95% confidence interval.

those with negative injections. However, false positive tests were common. Prevalence of positive tests in the sample ranged from 29.2% to 50.0%. Sensitivity, specificity, positive and negative predictive values and likelihood ratios for each individual test are presented in Table 2.

One approach to combining tests is simply to count the number of positives. Two by two contingency tables for the results of composites of all six SIJ tests (0, 1 or more, 2 or more and so on) versus the results of diagnostic injection and are presented in Table 3. Sensitivity, specificity, positive and negative predictive values and likelihood ratios were calculated and are presented in Table 4. The optimum composite rule was to identify the SIJ as the pain generator if there were

three or more positive tests, with estimated sensitivity of 93.8%, specificity of 78.1%, and AUC of 0.842 (s.e. 0.042).

On the other hand, looking at specific combinations of tests, it was found that the distraction test had the highest single positive predictive value (PPV) and AUC, the thigh thrust, compression and sacral thrust tests improved the overall diagnostic ability (as measured by improvement in AUC). The Gaenslen's tests did not improve the AUC value. This implies that Gaenslen's tests did not contribute positively and may be omitted from the diagnostic process without compromising diagnostic confidence. The optimal rule was to perform the distraction, thigh thrust, compression and sacral thrust tests but stopping when there are two positives.

Table 3  
Two by two contingency tables for results of composites of six SIJ tests and diagnostic injection

Composite of SIJ provocation tests	Positive SIJ injection	Negative SIJ injection	Totals
1 or more positive tests (+)	<b>16</b>	<b>18</b>	34
Zero positive tests (–)	<b>0</b>	<b>14</b>	14
Totals	16	32	48
2 or more positive tests (+)	<b>15</b>	<b>11</b>	26
Less than 2 tests (–)	<b>1</b>	<b>21</b>	22
Totals	16	32	48
3 or more positive tests (+)	<b>15</b>	<b>7</b>	22
Less than 3 positive tests (–)	<b>1</b>	<b>25</b>	26
Totals	16	32	48
4 or more positive tests (+)	<b>9</b>	<b>6</b>	15
Less than 4 positive tests (–)	<b>6</b>	<b>26</b>	33
Totals	15	32	47*
5 or more positive tests (+)	<b>4</b>	<b>4</b>	8
Less than 5 positive tests (–)	<b>11</b>	<b>28</b>	40
Totals	15	32	47*
6 positive tests (+)	<b>1</b>	<b>4</b>	5
Less than 6 positive tests (–)	<b>14</b>	<b>28</b>	43
Totals	15	32	47*

\*One patient without a distraction test but positive on the other three.

Bold numerals indicate the actual data in the 2 × 2 tables.

Table 4  
Sensitivity, specificity, positive and negative predictive values and likelihood ratios (95% confidence intervals) for composites from one to six SIJ tests

Statistic	1 or more positive tests	2 or more positive tests	3 or more positive tests	4 or more positive tests	5 or more positive tests
Sensitivity	1.00 (0.81, 1.00)	0.93 (0.72, 0.99)	0.94 (0.72, 0.99)	0.60 (0.36, 0.80)	0.27 (0.11, 0.52)
Specificity	0.44 (0.28, 0.61)	0.66 (0.48, 0.80)	0.78 (0.61, 0.89)	0.81 (0.65, 0.91)	0.88 (0.72, 0.95)
PPV	0.47 (0.32, 0.63)	0.58 (0.39, 0.75)	0.68 (0.47, 0.84)	0.60 (0.36, 0.80)	0.50 (0.22, 0.79)
NPV	1.00 (0.79, 1.00)	0.96 (0.78, 0.99)	0.96 (0.81, 0.99)	0.81 (0.65, 0.91)	0.72 (0.56, 0.84)
+LR	1.78 (1.41, 2.54)	2.73 (1.72, 4.64)	4.29 (2.34, 8.58)	3.20 (1.42, 7.31)	2.13 (0.64, 6.83)
–LR	0.00 (0.00, 0.46)	0.10 (0.02, 0.45)	0.80 (0.14, 0.37)	0.49 (0.24, 0.83)	0.84 (0.54, 1.11)

Notes: PPV = positive predictive value, NPV = negative predictive value, +LR = likelihood ratio for positive test, –LR = likelihood ratio for negative test.

This resulted in an AUC of (0.819, s.e. 0.054) with sensitivity of 0.88 and specificity of 0.78. Table 5 presents two by two contingency tables for the four tests that positively contribute to making the diagnosis (distraction, thigh thrust, compression and sacral thrust). Table 6 presents sensitivity, specificity, positive and negative predictive values and likelihood ratios for two positives of these four tests.

#### 4. Discussion

All patients with SIJ pathology identified by injection had at least one positive test. Only one patient out of 16 with SIJ pain had a single positive test with 15 having two or more positive SIJ tests. Consequently, one reasonable clinical rule is that when all provocation SIJ tests are negative, symptomatic SIJ pathology can be ruled out. The thigh thrust test is the most sensitive test and the distraction test is most specific.

Three or more of the six tests produce the highest likelihood ratio (4.29), but removal of Gaenslen's test from the examination and application of the rule "any two positive tests" of the remaining four tests produces almost as good a result (likelihood ratio = 4.0). Because the thigh thrust and distraction tests have the highest individual sensitivity and specificity, respectively (see Table 3), performance of these tests first seems reasonable. If both tests provoke familiar pain, no further testing is indicated. If one test is positive, the compression test is applied and if positive, a painful SIJ is likely and no further testing is required. If compression is not painful the sacral thrust test is applied. If this is painful, SIJ pathology is likely, whereas if it is not painful, SIJ pain is unlikely. Not only does this rule avoid subjecting patients to unnecessary tests, but also would in most cases permit a diagnosis even if one or more tests were not completed. Fig. 7 presents a diagnostic algorithm for this reasoning process.

Table 5

Two by two contingency tables for results of composites of SIJ tests out of distraction, thigh thrust, compression and sacral thrust versus diagnostic injection results

Composite of SIJ provocation tests	Positive SIJ injection	Negative SIJ injection	Totals
1 or more positive tests (+)	<b>16</b>	<b>17</b>	32
Zero positive tests (–)	<b>0</b>	<b>15</b>	15
Totals	16	32	48
2 or more positive tests (+)	<b>14</b>	<b>7</b>	16
Less than 2 tests (–)	<b>2</b>	<b>25</b>	27
Totals	16	32	48
3 or more positive tests (+)	<b>10</b>	<b>5</b>	15
Less than 3 positive tests (–)	<b>6</b>	<b>27</b>	8
Totals	16	32	48
4 positive tests (+)	<b>4</b>	<b>5</b>	9
Less than 4 positive tests (–)	<b>1</b>	<b>27</b>	38
Totals	15	32	47*

\*One patient without a distraction test but positive on the other three. Bold numerals indicate the actual data in the 2 × 2 tables.

Table 6

Sensitivity, specificity, positive and negative predictive values and likelihood ratios for two positive tests of distraction, thigh thrust, compression and sacral thrust

Statistic	Estimate	95% confidence interval
Sensitivity	0.88	0.64, 0.97
Specificity	0.78	0.61, 0.89
PPV	0.67	0.45, 0.83
NPV	0.93	0.77, 0.98
+LR	4.00	2.13, 8.08
–LR	0.16	0.04, 0.47

Notes: PPV = positive predictive value, NPV = negative predictive value. +LR = likelihood ratio for positive test, –LR = likelihood ratio for negative test.

When severe pain occurs with all body movements (e.g. acute disc prolapse, fractures, etc.), pain is provoked by any test including the provocation SIJ tests. In these circumstances interpretation of the SIJ tests is inappropriate. In our opinion, where another source of pain is known to be a major source of pain, the interpretation of the SIJ tests as evidence of a symptomatic SIJ should be avoided or entertained only with scepticism.

No single study can satisfy all criteria recommended by advisory groups (Deyo et al., 1994) and this study is no exception. One threat to external validity within this study is that the patients in this study were more chronic and disabled than those usually seen in primary care or most secondary referral environments. While generalizability of the study results must be questioned, it is our anecdotal experience that most primary care and secondary referral patient populations are less difficult to examine and analyse, and these results understate

rather than overstate the diagnostic power of the provocation SIJ tests. Additionally, the effects of preceding provocation tests may confound interpretation of single test results. Progressive increases or decreases in pain responses to the second, third or fourth tests cannot be ruled out as a confounding factor. A different study design would be required to eliminate this confounder, such as allowing a specified rest period between tests, or applying only a single test to each individual patient before the diagnostic injection. However, the latter design would not permit evaluation of groups or sequences of tests.

The criterion standard for the diagnosis of painful lumbar facet joint is comparative anaesthetic or placebo controlled blocks and this is widely accepted and utilized in studies (Dreyfuss et al., 2003). However, standards used in recent studies of SIJ pain diagnosis are diverse. The International Association for the Study of Pain (IASP) has proposed criteria for making the diagnosis of symptomatic SIJ and are: (1) pain is present in the region of the SIJ, (2) stressing the SIJ by clinical tests that are selective for the joint reproduces the patient's pain, (3) selectively infiltrating the putatively symptomatic joint with local anaesthetic completely relieves the patient of pain (Merskey and Bogduk, 1994). In recent diagnostic studies of SIJ pain there are variations on a general theme. Fortin et al. (1994a) used patterns of pain distribution, provocation of pain during SIJ injection and a single anaesthetic block. Schwarzer et al. (1995) used a single injection in patients with pain 'centred' below L5/S1 and a 75% reduction in pain following injection of local anaesthetic. Dreyfuss et al. (1996) used a single injection of local anaesthetic and cortico-steroid, noted pain provocation and required more than 90% reduction in the 'main pain' as distinct from a change in

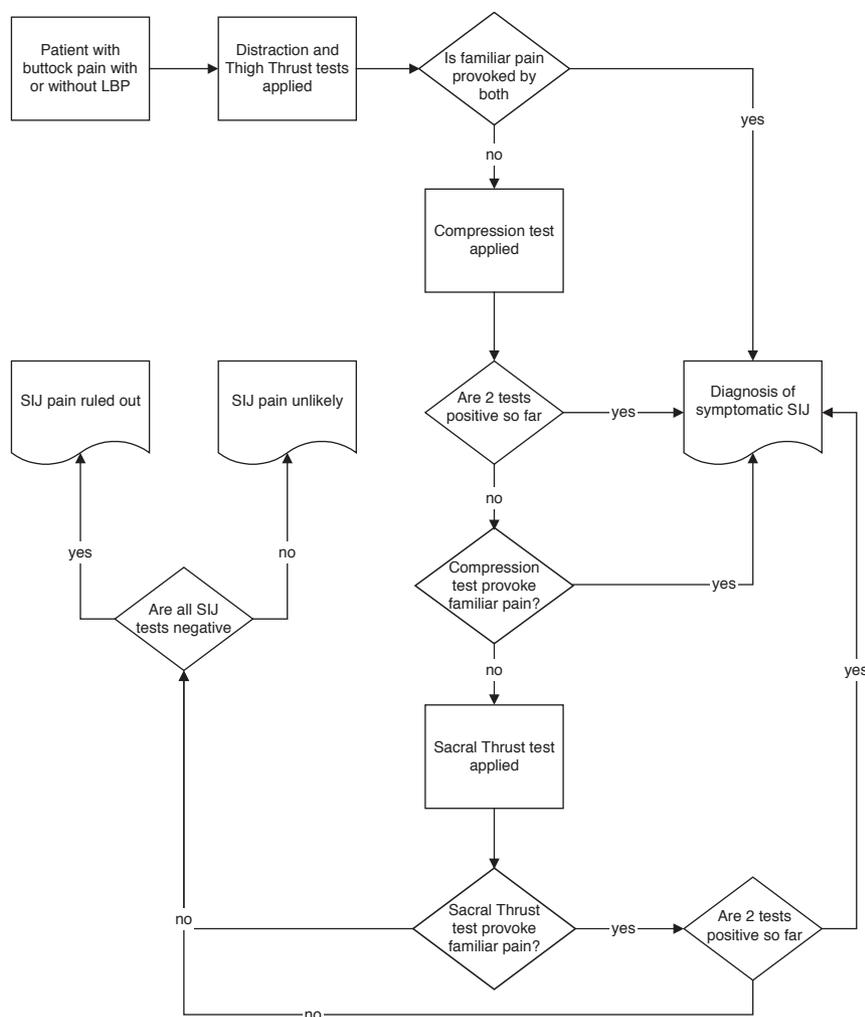


Fig. 7. Diagnostic algorithm for SIJ pain using provocation SIJ tests: distraction, thigh thrust, compression and sacral thrust.

VAS assessment of pain generally. Maigne et al. (1996) used comparative double blocks in patients selected by pain drawing as likely to have SIJ pain and at least 75% reduction on a general pain VAS. Slipman et al. (1998) used an 80% reduction on a general pain VAS following single anaesthetic injection in consecutive LBP patients. In an earlier presentation of a subset of patients from the current study, we utilized double comparative blocks and 80% or more reduction in a verbal analogue scale of pain intensity and provocation of pain during SIJ injection as the criterion standard (Laslett et al., 2003). In the current analysis, a single diagnostic injection under fluoroscopic control and contrast enhancement that provoked familiar pain was used and resulted in 80% or more relief of pain as measured by a verbal analogue scale of pain intensity. Where familiar pain was provoked during injection, corticosteroid was injected in addition to local anaesthetic and the patient scheduled for a confirmatory, comparative block. It is noted that in a recent publication (Bogduk and McGuirk, 2002, p. 174), double comparative blocks

are recommended for confirmation of the diagnosis. The data collection for the current paper was between 1996 and 1998, at the time when mixtures of standards were common. Future criterion standard validity studies should use the standard recommended by Bogduk and McGuirk without inclusion of corticosteroid in the initial screening injection.

Although false positive rates for SIJ injections have not been previously reported, a rate of 7.7% may be calculated from data presented from one study (Schwarzer et al., 1995) and 20.5% from another (Maigne et al., 1996). In this current analysis, the 16 patients reporting a positive response to a single anaesthetic injection and 12 proceeded on to receive a second injection. All of these patients reported a positive anaesthetic response, confirming the diagnosis of SIJ pathology with a false positive rate of zero. Of the four initial responders who did not receive a confirmatory block, three derived such pain relief from the initial block that a confirmatory block was inappropriate. It is assumed that ablation of pain following the initial block was a consequence of the

introduction of corticosteroid during the initial procedure. One patient did not return for the scheduled confirmatory block for unknown reasons.

In a worst-case scenario using the comparative confirmatory blocks as a criterion standard, we can propose that all four cases not returning for a confirmatory block would have returned a negative response to that procedure and the initial block deemed false positive. Estimations (with 95% confidence intervals) for sensitivity, specificity, positive/negative predictive values would be 83.3 (55.2, 95.3), 69.4 (53.1, 82.0), 47.6 (28.3, 67.7), 92.6 (76.6, 97.9) percent, respectively, and positive/negative likelihood ratios would be 2.72 (1.57, 4.74) and 0.24 (0.66, 0.87), respectively. The estimations from this scenario still exceed what can be expected by random chance at the lower 95% confidence limit. However, in this scenario the false positive rate (95% confidence intervals) would have been 30.6% (15.1, 45.6).

Although diagnostic injection is the only available criterion standard against which clinical tests can reasonably be evaluated for validity, it is acknowledged that false negative and false positive responses to injection are possible. Where there is a defect in the articular capsule, leakage of anaesthetic into adjacent areas may occur (Fortin et al., 1994a; Schwarzer et al., 1995), and pain relief may be a reflection of an anaesthetic affect of these structures rather than the SIJ structures. This possibility and the unknown effect of psychosocial influences on pain responses to invasive diagnostic procedures may contribute to the false positive and negative rates. No attempt to estimate these influences was attempted during this study. In addition, intra-articular injection of anaesthetic has the potential to ablate SIJ pain when originating within the joint cavity, but is unlikely to have an anaesthetic effect on SIJ structures external to the joint (Grieve, 1988). (Maigne et al., 1996; Laslett et al., 2003) Where SIJ structures external to the joint cavity are actual pain generators, an intra-articular injection of local anaesthetic into and confined to the joint space will produce a false negative diagnostic result, whereas the clinical examination may possibly correctly identify the periarticular and unanaesthetized SIJ structures as pain generators.

The patients entered into this study were not consecutive. The physiotherapists performing the clinical examination were not residents in New Orleans where the diagnostic injections were being carried out and could visit only intermittently over a 19-month period. Consequently, no estimate of prevalence should be inferred from the data presented in this report. Additionally, calculation of predictive values or false positive rates in a sample, at least partially selected for possible SIJ involvement, are not generalizable to other patient populations.

The results of this study are in contrast to the results of earlier similar studies (Schwarzer et al., 1995; Dreyfuss et al., 1996; Maigne et al., 1996; Slipman et al., 1998), and the conclusion from a meta-analysis of studies of clinical tests for painful SIJs (van der Wurff et al., 2000). However, there is support for the use of pain provocation tests in SIJ diagnosis (Laslett et al., 2003) and these tests are preferred over palpation tests for mobility or position (Freburger and Riddle, 2001). It is difficult to account for the differences between our results and the results from other studies. However, some of the explanation may lie in differences in application of the examination technique. There is evidence that physiotherapists apply different degrees of force when utilizing SIJ provocation tests (Levin et al., 1998, 2001) and this may be one of several factors influencing results.

## 5. Conclusion

Provocation SIJ tests have significant diagnostic utility. Six provocation tests were selected on the basis of previously demonstrated acceptable inter-examiner reliability. Two of four positive tests (distraction, compression, thigh thrust or sacral thrust) or three or more of the full set of six tests are the best predictors of a positive intra-articular SIJ block. When all six SIJ provocation tests are negative, painful SIJ pathology may be ruled out.

## Acknowledgements

Thanks to Duncan Reid, Wayne Hing, and the Auckland University of Technology Multimedia Unit for assistance with photographs.

Travel and Louisiana licensing costs for Mrs. Young were funded by The McKenzie Institute International.

## References

- Adams MA, Bogduk N, Dolan P. Biomechanics of low back pain. London: Churchill-Livingstone; 2002.
- Altman DG, Machin D, Bryant TN, Gardner MJ. Statistics with confidence. 2nd ed. Bristol: British Medical Journal; 2000.
- Barnsley L, Lord S, Bogduk N. Comparative local anaesthetic blocks in the diagnosis of cervical zygapophysial joint pain. *Pain* 1993;99–106.
- Beattie PF, Meyers SP, Stratford P, Millard RW, Hollenberg GM. Associations between patient report of symptoms and anatomic impairment visible on magnetic resonance imaging. *Spine* 2000; 25(7):819–28.
- Bogduk N. The anatomical basis for spinal pain syndromes. *Journal of Manipulative and Physiological Therapeutics* 1995;18(9):603–5.
- Bogduk N, McGuirk B. Medical management of acute and chronic low back pain, vol. 13. Amsterdam: Elsevier Science BV; 2002.

- Bryant TN. Confidence interval analysis for windows. (2.0.0): BMJ Books, 2000.
- Cyriax, J., 6th ed. Textbook of orthopaedic medicine. Volume one: diagnosis of soft tissue lesions. London: Balliere Tindall; 1975. p. 546–54.
- Deville WLJM, van der Windt DAWM, Dzaferagic A, Bezemer PD, Bouter LM. The test of lasegue: systematic review of the accuracy in diagnosing herniated discs. *Spine* 2000;25(9):1140–7.
- Deyo RA, Haselkorn J, Hoffman R, Kent DL. Designing studies of diagnostic tests for low back pain or radiculopathy. *Spine* 1994;19(18 (Suppl)):2057S–65S.
- Dreyfuss PH, Michaelsen M, Pauza K, McLarty J, Bogduk N. The value of history and physical examination in diagnosing sacroiliac joint pain. *Spine* 1996(21):2594–602.
- Dreyfuss PH, Dreyer SJ, Vaccaro A. Lumbar zygapophysial (facet) joint injections. *The Spine Journal* 2003;3(Suppl):50–9.
- Fortin JD, Falco FJ. The Fortin finger test: an indicator of sacroiliac pain [see comments]. *American Journal of Orthopedics* 1997;26(7):477–80.
- Fortin JD, Aprill C, Pontieux RT, Pier J. Sacroiliac joint: pain referral maps upon applying a new injection/arthrography technique. Part II: clinical evaluation. *Spine* 1994;19(13):1483–9.
- Fortin JD, Dwyer AP, West S, Pier J. Sacroiliac joint: pain referral maps upon applying a new injection/arthrography technique. Part I: asymptomatic volunteers. *Spine* 1994;19:1475–82.
- Freburger JK, Riddle DL. Using published evidence to guide the examination of the sacroiliac joint region. *Physical Therapy* 2001;81(5):1135–42.
- Grieve GP. Diagnosis. *Physiotherapy Practice* 1988;4:73–7.
- Jensen MP, Strom SE, Turner JA, Romano JM. Validity of the sickness impact profile Roland scale as a measure of dysfunction in chronic pain patients. *Pain* 1992;50:157–62.
- Kokmeyer DJ, van der Wurff P, Aufdemkampe G, Fickenscher TCM. The reliability of multitest regimens with sacroiliac pain provocation tests. *Journal of Manipulative and Physiological Therapeutics* 2002;25(1):42–8.
- Laslett M, Williams M. The reliability of selected pain provocation tests for sacroiliac joint pathology. *Spine* 1994;19(11):1243–9.
- Laslett M, Young SB, Aprill CN, McDonald B. Diagnosing painful sacroiliac joints: a validity study of a McKenzie evaluation and sacroiliac joint provocation tests. *Australian Journal of Physiotherapy* 2003;49:89–97.
- Lawlis GF, Cuencas R, Selby D, McCoy CE. The development of the Dallas pain questionnaire. An assessment of the impact of spinal pain on behavior. *Spine* 1989;14(5):511–6.
- Levin U, Nilsson-Wikmar L, Stenstrom CH, Lundeberg T. Reproducibility of manual pressure force on provocation of the sacroiliac joint. *Physiotherapy Research International* 1998;3(1): 1–14.
- Levin U, Nilsson-Wikmar L, Harms-Ringdahl K, Stenstrom CH. Variability of forces applied by experienced physiotherapists during provocation of the sacroiliac joint. *Clinical Biomechanics* 2001; 16:300–6.
- Maigne JY, Aivaliklis A, Pfefer F. Results of sacroiliac joint double block and value of sacroiliac pain provocation tests in 54 patients with low back pain. *Spine* 1996;21(16):1889–92.
- McCombe PF, Fairbank JCT, Cockersole BC, Pynsent PB. Reproducibility of physical signs in low back pain. *Spine* 1989;14(9): 908–18.
- McKenzie RA. The lumbar spine: mechanical diagnosis and therapy. Waikanae: Spinal Publications Ltd.; 1981.
- Meijne W, van Neerbos K, Aufdemkampe G, van der Wurff P. Intraexaminer and interexaminer reliability of the Gillet test. *Journal of Manipulative and Physiological Therapeutics* 1999; 22(1):4–9.
- Mersky H, Bogduk N. Classification of chronic pain: descriptions of chronic pain syndromes and definitions of pain terms. 2nd ed. Seattle: IASP Press; 1994.
- Ohnmeiss DD, Vanharanta H, Ekholm J. Relationship of pain drawings to invasive tests assessing intervertebral disc pathology. *European Spine Journal* 1999;8(2):126–31.
- Potter NA, Rothstein JM. Intertester reliability for selected clinical tests of the sacroiliac joint. *Physical Therapy* 1985;65(11): 1671–5.
- Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine* 1983;8(2):141–50.
- Sakamoto N, Yamashita T, Takebayashi T, Sekine M, Ishii S. An electrophysiologic study of mechanoreceptors in the sacroiliac joint and adjacent tissues. *Spine* 2001;26:E468–71.
- Schwarzer AC, Aprill C, Bogduk N. The sacroiliac joint in chronic low back pain. *Spine* 1995;20(1):31–7.
- Slipman CW, Sterenfeld EB, Chou LH, Herzog R, Vresilovic E. The predictive value of provocative sacroiliac joint stress maneuvers in the diagnosis of sacroiliac joint syndrome. *Archives of Physical Medicine and Rehabilitation* 1998;79(3):288–92.
- van der Wurff P, Meyne W, Hagmeijer RHM. Clinical tests of the sacroiliac joint: a systematic methodological review. Part 2: validity. *Manual Therapy* 2000;5(2):89–96.