The Journal of Pain, Vol 10, No 4 (April), 2009: pp 354-368
Available online at www.sciencedirect.com

Diagnostic Validity of Criteria for Sacroiliac Joint Pain: A Systematic Review

Karolina M. Szadek,* Peter van der Wurff,† Maurits W. van Tulder,‡ Wouter W. Zuurmond,§ and Roberto S. G. M. Perez‖

*Department of Anesthesiology, VU University Medical Center, and EMGO Institute for Research in Extramural Medicine, VU University Medical Center Amsterdam, Amsterdam, The Netherlands.
†Division of Perioperative Medicine and Emergency Care, Department of Anesthesiology and Pain Treatment, University Medical Center, Utrecht, The Netherlands, and Department of Physiotherapy, Military Rehabilitation Centre "Aardenburg," Doorn, The Netherlands.
‡Institute of Health Sciences, Faculty of Earth and Life Sciences, VU University Amsterdam and EMGO Institute, VU University Medical Center, Amsterdam, The Netherlands.
§Department of Anesthesiology, VU University Medical Center, Amsterdam, The Netherlands.
‖Department of Anesthesiology, VU University Medical Center, Amsterdam, The Netherlands, and EMGO Institute for Research in Extramural Medicine, VU University Medical Center Amsterdam, The Netherlands.

Abstract: A systematic literature review was conducted to determine the diagnostic validity of the criteria for sacroiliac (SI) joint pain as proposed by the International Association for the Study of Pain (IASP). Databases were searched up to September 2007. Quality of the studies was assessed using a Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool. Sensitivity, specificity, and diagnostic odds ratios (DOR) were calculated together with 95% confidence intervals (CI). Statistical pooling was conducted for results of provocative tests. Eighteen studies were included. Five studies examined the pattern of SI joint pain, whereas another 5 examined stressing test specific for SI joint pain. None of the studies evaluated the diagnostic validity of the SI joint infiltration or the diagnostic validity of the IASP criteria set as a whole. In all studies, the SI joint selective infiltration was used as a gold standard; however, the technique, medications, and required pain relief after the infiltration varied considerably between the studies. Taking the double infiltration technique as reference test, the pooled data of the thigh thrust test (DOR, 18.461; CI, 5.82 to 58.53), compression test (DOR, 3.88; CI, 1.7 to 8.9), and 3 or more positive stressing tests (DOR, 17.16; CI, 7.6 to 39) showed discriminative power for diagnosing SI joint pain.

***Perspective:** This review of clinical studies focused on the diagnostic validity of the IASP criteria for diagnosing SI joint pain. A meta-analysis showed that the thigh thrust test, the compression test, and 3 or more positive stressing tests have discriminative power for diagnosing SI joint pain. Because a gold standard for SI joint pain diagnosis is lacking, the diagnostic validity of tests related to the IASP criteria for SI joint pain should be regarded with care.*

Key words: *Sacroiliac joint pain, pain pattern, provocation test, joint infiltration, diagnostic vaidity, systemic review.*

Sacroiliac (SI) joint pain is considered to be a significant problem with a frequent occurrence in patients with low back pain. The prevalence of SI joint pain, as established on the basis of clinical evaluation, varies from 15% to 30% in patients with low back pain.[4,47,59] Despite this relatively high prevalence, it is difficult to make a definite diagnosis, as presenting symptoms may be similar to those of other causes of low back pain, and mimic, for instance, sciatica.[4,8] To make a clinical diagnosis, the International Association for the Study of Pain (IASP) has proposed a set of criteria for diagnosing SI joint pain that address mechanical disorders of this joint.[52] According to these criteria, SI joint pain refers to patients with pain in the area of the SI joint, which should be reproducible by performing

Address reprint requests to Dr Karolina M. Szadek, Department of Anesthesiology, VU University Medical Center, PO Box 7057, 1007 MB, Amsterdam, The Netherlands. E-mail: km.szadek@vumc.nl
1526-5900/$36.00

doi:10.1016/j.jpain.2008.09.014

specific pain provocation tests, or should be completely relieved by infiltration of the symptomatic SI joint with local anesthetics. These criteria, however, are not unambiguous. Regarding the first IASP criterion, the literature suggests that SI joint pain, though originating from the SI joint, can have a diffuse character and can refer to the buttock, groin, or/and lower extremity.[25,26,61,69,77] With respect to the second IASP criterion, there are several pain provocation tests described that intend to stretch, compress, or contract certain tissue structures related to the SI joint.[60] However, the reliability (except for the Gaenslen test and thigh thrust test) as well as validity of these tests in clinical practice is disputable.[70,71] The third IASP criterion, the selective infiltration of the SI joint, whereby the local anesthetic is injected in the joint cavity, plays a role of a reference standard.[12,22,25,47,59,61,63] This technique, however, is criticized because of missing data about specificity and sensitivity and therefore, considered not valid for diagnosing SI joint pain.[5] Moreover, according to the literature, SI joint painful pathology appears to involve not only intra-articular structures but also periarticular structures (for instance, ligaments and muscles).[14]

In the last few years, several reviews were published relating to SI joint pain.[5,10,11,14,20,23,24,31,35,36,51,60,65,70,71,80] Results of these reviews are contradictory with respect to methodological quality of the evaluated studies and diagnostic value of tests. Three of the most recent reviews suggested utilization of multiple diagnostic tests for physical examination instead of relying on the result of a single test.[35,60,65] The review by Hansen et al,[37] however, found that there is limited evidence for provocation tests and moderate evidence for diagnostic infiltration of the SI joint. On the other hand, Berthelot et al[5] argued in their review that both the provocation tests and diagnostic infiltration are unreliable for diagnosing SI joint pain. Furthermore, none of these reviews considered all of the criteria for SI joint pain as proposed by the IASP. Therefore, with the purpose of assessing whether these criteria are capable of discriminating among the individuals with and without SI joint pain, we performed a systematic review of the literature. The evaluation comprised the methodological assessment of the studies and the summary of the diagnostic validity by sensitivity, specificity, and diagnostic odds ratios (DORs) of diagnostic tests that could be ascribed to the IASP criteria.

Methods

Study Identification

To identify relevant literature, we conducted a comprehensive search in the following databases: PubMed, EMBASE, and CINAHL, from the beginning of these databases up to September 2007. The full search strategy was developed in collaboration with an experienced librarian (I.R.) using a filter outlined by Deville et al.[19] However, using this filter in a preliminary search resulted in only a few hits in the area of SI joint pain. Therefore, we broadened our search, as outlined in Appendix 1. We applied no restrictions with regard to year of publication or language. Case series and case reports as well as animal and cadaveric studies were excluded. Additionally, the reference sections of all articles selected for the review were scanned for potentially relevant articles that were not identified by the original search.

Article Retrieval

Eligibility of studies on the basis of title, key words, and abstract was determined by 2 reviewers (K.S. and P.v.d.W.), independently. If uncertainty remained, the full text was reviewed. Differences in judgment were resolved through a consensus procedure. Justifications for excluding studies were noted and discrepancies discussed. If no consensus was reached, a final decision was made by a third reviewer (R.P.). The publications were included according to the following criteria: (1) patients in a particular study were at least 18 years old, (2) suffering from nonspecific, non–pregnancy-related low back pain with or without radiation to the lower extremities or groin, (3) a diagnostic infiltration of the SI joint was compared with another diagnostic test, or (4) any diagnostic test was compared with 1 of the diagnostic criteria for SI joint pain according to the IASP.

Data Extraction

The same 2 reviewers performed data extraction independently, using a standardized questionnaire developed for this study. One of the reviewers is first author of 2 of the included studies.[68,69] He was not involved in any decision regarding data extraction or quality assessment of these studies. Data extraction and quality assessment of these 2 studies were done by only 1 reviewer (K.S.).

The following data were extracted: author, year of publication, country where the study was performed, the setting for patient recruitment, characteristics of the study population (age, gender, duration of the complaints), inclusion and exclusion criteria, the test(s) examined, reference test used, duration of symptoms and study results, whether recruitment was consecutive and/or data collection was performed prospectively (Table 1).

Assessment of Methodological Quality

The methodological quality of the studies was assessed by the 2 reviewers using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool.[75] The QUADAS items are defined as follows.

1. *Was the spectrum of patients representative of the patients who will receive the test in practice?*

This was considered positive if included subjects where older than 18 years, suffered from nonspecific, nonspondyloarthropathy, noninflammatory low back pain presumably stemming from the SI joint; there was no radicular radiation of pain to lower extremity along 1 or more dermatomes, accompanied by numbness and tingling, muscle weakness and loss of specific reflexes.

2. *Were selection criteria clearly described?*

This was considered positive when both inclusion and exclusion criteria were clearly described, and it

Table 1. Study Characteristics for the Individual Studies That Investigated Diagnostic Accuracy of Provocative Tests for Subjects With Sacroiliac Joint Pain

Author, reference, country	*Sample size n = all subjects, n = cases*[1]*; age [mean (SD or range)]; gender % (F)*	*Recruitments: Setting (s), consecutive (c), prospective (p); period;*	*Inclusion criteria Exclusion criteria*	*Type of reference test; outcome measure*	*Type of Index test*
Broadhurst[7] 1998 Australia	n = 40, n = 40 [F =36 years; range, 18-72 M =35 years; range, 25-53 - (No SD)]; - 75 % F	s = university hospital, c = yes; p = yes; 2-year period	LBP below lumbosacral junction, associated with groin pain and absence of lumbar symptoms; pain with full weight bearing on 1 leg; worsening going down hill Previous LBP; fractures; infections; metastases; pregnancy; use of major tranquilizers; systemic disease	Single SIJB with 4 mL Lidocaine 1% or NaCl 0.9% with image intensification; 70% pain reduction of pain on VAS	Flexion abduction and external rotation (Patrick's sign), Posterior shear (thigh trust test), resisted abduction test.
Dreyfuss[22] 1996 USA	n = 88, n = 85 [45 years, (range, 18-87)]; 72% F	s = university hospital, spine center; c = yes, p = yes; period unknown;	LBP below L5 Radiation ±Pain pattern consistent with SIJP Unknown	Contrast-enhanced SIJB with 1.5 mL 2% lidocaine and 0.5 mL corticosteroids; 90% pain reduction on VAS	12 physical examination tests
Fortin1994[25] 1997[27] USA	n = 54, n = 16 [35 years (range 21-45)]; 38% F	s = Regional spinal diagnostic center; p = yes; c = unknown; period unknown;	LBP ≥2 weeks No preselection	Pain provocation arthrography with 1-2.7 mL contrast; intra-articular bupivacaine 0.75%; positive if patient's pain was provoked	pain patterns[25]; pain area pointing (Fortin finger test)[27]
Fukui 2002 Japan[32]	n = 28 [58 ±13 years]; gender unknown	s = University hospital; c = unknown p = unknown; 1994-2001	Pain in the SIJ region ±referred pain; Pain provocation by Patrick's and Gaenslen's test; 80% pain relief after SIJB Zygapophyseal joint pain; radicular pain	Contrast-enhanced intra-articular injection of 2 mL 1% mepivacaine and 2 mg dexamethasone; not predefined	pain patterns
Laslett 2003[45] 2005[44] Sweden	n = 62, n = 48; [42 years (range, 20-79)]; 67 % F	s = Private radiology practice; c = no; p = yes; 21 months	Buttock pain, ±lumbar or lower extremity symptoms, able to tolerate full physical examination Unwilling to participate; Only midline or symmetrical pain above L5; clear root compression signs; referred for special procedures except SI joint injections;	Contrast enhanced intra-articular double infiltration of <1.5 mL lidocaine and bupivacaine (corticosteroid); pain provocation during infiltration; required pain reduction 80%	McKenzie evaluation, SI joint tests, hip joint assessment[45]; distraction test, thigh thrust, Gaenslen's test, compression test, sacral thrust[44]
Maigne[47] 1996 France	n = 67, n = 54; [median = 45.3 years; IQR, 34.5-56.8]; 61% F	s = Public hospital, spine center; c = no, p = yes; period of recruitment unknown	Chronic LBP >50 days, radiation ±VAS >40 mm, failure of epidural or facet joint injections, pain and tenderness over the region of SIJ, 18-75 years old Previous spine surgery, disc herniation, chemonucleolysis, neurological changes, discopathy, lumbar spinal stenosis, spondylolisthesis	Contrast enhanced double SIJB with 2 mL l2% lidocaine and 0.5% bupivacaine; positive if pain reduction >75% for at least 2 hours	Clinical physical examination tests

(*Continued*)

Maigne[48] 1998 France	n = 39, n = 32; [53.4 ±18.5]; 56% F	s = University hospital; c = unknown; p = yes; period of recruitment unknown	Chronic low back pain, tenderness of the SI joint line, VAS ≥4 cm, age >18 years Surgery in history, lumbar disc narrowing, spondylolisthesis, spinal stenosis, herniated intervertebral disc, chemonucleolysis, pregnancy	1 mL contrast enhanced intra-articular SIJB with 2 mL lidocaine 2%; positive if pain relief 15 min after the injection was >75%	Quantitative radionuclide bone scan
Maigne[49] 2005 France	n = 61; n = 40; [48 ± 11 years]; 65% F	s = Public hospital, spine center; c = not; p = yes; 1996-2002	Chronic LBP ≥6 months, Unilateral pain, pain distribution within SI joint pain pattern, sacral sulcus palpation tenderness Pain radiation below the knee, lumbar cause of pain, work-related injury, lawsuit, psychiatric disorder	1 mL contrast enhanced intra-articular SIJB with 2 mL lidocaine 2%; Positive if pain relief was 75%.	sacroiliac bone scintigraphy
Manchikanti[50] 2001 USA	n = 120; [47.5 SEM = 1.16]; 66% F	s = nonuniversity, private practice; c = unclear; p = yes; period of recruitment unknown	Chronic LBP .6 months; age, 18-90 years Neurological deficits, definite diagnosis known, pain in the sacral region, tenderness over SIJ, positive provocative maneuvers	0.5-1 mL contrast enhanced intra-articular double SIJB with 0.5-1 mL lidocaine 2% or bupivacaine 0.5%; Positive effect undefined	Physical examination Single versus double blockade
Schwarzer[59] 1995 Australia	n = 100; n = 43 [median = 32.8, ICQ 28.7-40.9] 51% F	s = University hospital; c = yes; p = yes; April-October 1992	Pain below L5-S1, 18-80 years old Status post lumbar spinal surgery, exhibition of neurological signs	1 mL contrast enhanced intra-articular SIJB with 1 mL lidocaine 2%; Positive if pain relief was >75%	pain provocation during the infiltration; pain patterns
Slipman[62] 1996 USA	n = 50, n = 50; [range, 18-77]; 66% F	s = spine center; c = yes; p = yes; period of recruitment unknown	LBP, ±radiation to the lower extremity, sacral sulcus tenderness, 3 positive provocation tests, no improvement after physical therapy SA, urethritis, peripheral arthritis, psoriasis, early morning stiffness, inflammatory bowel disease, neurological deficit	0.5 mL contrast enhanced intra-articular SIJB with 2 mL lidocaine 2% and 1 mL corticosteroids or 3.0 mL lidocaine 2% or 3 mL lidocaine 1%; Positive if pain relief was ≥80%	Bone scan
Slipman[63] 1998 USA	n = 50, n = 50 [range, 18-77] 62 % F	s = University hospital, spine center; c = yes; p = no; period of recruitment unknown	LBP, ±radiation to the lower limb Previous spine surgery, SA, urethritis, peripheral arthritis, psoriasis, inflammatory bowel disease, neurological deficit	0.5 mL contrast enhanced SIJB with 2 mL lidocaine 2% en 1 mL betamethasone /or up to 3 mL lidocaine 2% or up to 3 mL lidocaine 1%; positive if pain relief was ≥80%	physical examination Patrick's test, pressure pain at the sacral sulcus; shear test, standing extension; Gaenslen's test and Yeoman test
Slipman[61] 2000 USA	n = 50; n = 50 [mean, 42.5; range, 20-75 years] F 64%	s = University hospital, spine center; c = yes; p = no; period of recruitment unknown	LBP or buttock pain, ±radiation to the lower limb, 3 positive provocation tests SA, urethritis, peripheral arthritis, psoriasis, inflammatory bowel disease, early morning stiffness, neurological or muscular deficit, spondylolisthesis, lumbar instability, lumbosacral radiculopathy	0.5 mL contrast enhanced SIJB with 2 mL lidocaine 2% Positive if pain relief was at least 80 %	pain referral zones

(*Continued*)

Table 1. Study Characteristics for the Individual Studies That Investigated Diagnostic Accuracy of Provocative Tests for Subjects With Sacroiliac Joint Pain

Author, reference, country	*Sample size n = all subjects, n = cases[1]; age [mean (SD or range)]; gender % (F)*	*Recruitments: Setting (s), consecutive (c), prospective (p); period;*	*Inclusion criteria Exclusion criteria*	*Type of reference test; outcome measure*	*Type of Index test*
Vd Wurff[68,69] 2006 The Netherlands	n = 140; n = 60 [51 ± 13] 78% F	s = General hospital, pain department; c = p = yes; January 2001-April 2002	Chronic LBP ≥50 days, pain below L5 over the posterior aspect of SI joint unilaterally, ±leg pain, VAS >45 mm, age 18-80 years SA, leg-length discrepancies of >2 cm, Waddell score >2, tumors, recent lumbar spine fractures, disc abnormalities with nerve root compression clinical signs, osteoporosis, infection, clinically symptomatic cox-arthrosis, radicular pain with neurological signs, pregnancy, anticoagulants, liver and/or kidney failure	contrast (1 mL) enhanced double SIJB with 2 mL lidocaine 2% or bupivacaine 0.25%; Positive if pain relief was at least 50% for 1 hour after lidocaine or 4 hours after bupivacaine infiltration	pain mapping[69]; multitest regimen of 5 SI joint pain provocation tests[68]
Young[78] 2003 USA	n = 102, n = 81 [40.8; SD,12.1 years] 60.5% F	s = private radiology practice, c = unclear; p = yes; period of recruitment unknown	Chronic LBP Unwilling to participate, signs of nerve root compression, unable to tolerate clinical examinations	single disc, facet and/or SI joint arthrography/pain provocation/ injections; <1.5 mL local anesthetic; pain provocation during infiltration and pain reduction 80%	pain mapping

Abbreviations: F, female; M, male; SD, standard deviation; SEM, standard error of mean; LBP, low back pain; VAS, visual analog scale, ±, with or without; SIJB, SI joint blockade; SIJP, SI joint pain; SA, spondyloarthropathy.
NOTE. Publications are listed alphabetically by first author. Cases[1] = number of subjects actually undergoing the index test.

was clear whether the recruitment of subjects was conducted prospectively and the inclusion of subjects was consecutive.

3. *Is the reference standard likely to correctly classify the target condition?*

In the absence of a gold standard in diagnosing primary SI joint pain, the double infiltration of the SI joint was considered the best available reference test.[6] Moreover, the reference test used in a specific study was described without direct interpretation of its usefulness. Data from individual studies regarding this subject were extracted and recorded in Table 1.

4. *Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the 2 tests?*

This item was considered positive if the delay between the application of the index test and reference standard used in the study was reported and was not more than 7 days, or for SI joint infiltration with local anesthetics not shorter than 24 hours (bupivacaine, 0.5% $T_{0,5}$ = 1.5 to 5.5 hours, and lidocaine, 2% $T_{0,5}$ = 90 to 120 minutes).

5. *Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?*

This was considered positive if it was clearly described that all patients or a random selection of patients received verification of their disease status with the reference standard used in the study, regardless of the index test results. In the case of random selection, it would be clear that the randomization took place before the implementation of the index test.

6. *Did patients receive the same reference standard regardless of the index test result?*

This was evaluated positive if it was clear that all patients received the same reference standard used in the study regardless of the index test result. In the case of a random selection, this item would be scored as positive if it was clear that the randomization was performed before applying both index and reference test.

7. *Was the reference standard independent of the index test?*

This was scored positive if it was clear that the reference test used in the study was independent of the index test, positive/negative results of both tests were predefined and the index test did not form part of the reference standard. If the method of the assessment of index test and reference test would not be explained, this item would be scored as negative.

8. *Was the execution of the index test described in sufficient detail to permit replication of the test?*

This item was scored positive if the index test was sufficiently described to permit its reproduction, or a reference to the adequate description of the test was provided. Only mentioning the tests names would be scored negative, as some of the tests could be named erroneously.[43]

9. *Was the execution of the reference standard described in sufficient detail to permit its replication?*

This was scored positive if the reference standard was sufficiently described to permit its reproduction. This item would also be scored positive if a reference was provided to a full test description.

10. *Were the index test results interpreted without knowledge of the results of the reference standard?*
11. *Were the reference standard results interpreted without knowledge of the results of the index test?*

These 2 items were assessed positive if the study clearly stated that the results of both the index test and the reference standard, were interpreted in a blinded manner. If case of uncertainty, this item would be scored as "unclear" and negative by missing information.

12. *Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?*

This was considered positive if the observer was aware of the characteristics mentioned in the inclusion and exclusion criteria. If the diagnosis of SI joint pain was already known, this item was scored negative.

13. *Were uninterpretable/ intermediate test results reported?*

If it was clear that all test results, including uninterpretable/undetermined/ intermediate results were reported, this item was scored positive. Therefore, the number of included patients should match the number of the subjects receiving the index test. If it was not a case and not further explanation was available this was scored negative.

14. *Were withdrawals from the study explained?*

This was scored positive if it was clear what happened to all patients who entered the study, for example if a flow diagram of study participants was reported. If the data were not available, than the item would be scored negatively.

The 2 reviewers, blind to each other assessment, scored the criteria items: "yes" or "no" when studies satisfied or failed to meet the criteria, respectively, and "unclear" when information was lacking to decide whether the study satisfied or met that specific item. In the case of disagreement, the 2 reviewers tried to reach a consensus on each criterion, and in case of persisting disagreement, a third reviewer (R.P.) decided. Agreement between reviewers was quantified using the kappa (κ) coefficient with quadratic weighting (http://www.faculty.vassar.edu/lowry/kappa.html[58]). The strength of agreement was interpreted as poor (<0.20), fair (0.21 to 0.40), moderate (0.41 to 0.60), good (0.61 to 0.80) and very good (0.81 to 1.00).[2]

Pilot Study

To improve agreement between the 2 reviewers, this procedure was tested using 2 studies hat examined the diagnostic accuracy of Lachman's test in knee

instability.[18,46] The disagreement between the 2 reviewers was 18% (5/28); κ was 0.77 (95% confidence interval [CI], 0.58 to 0.96) for all criteria.

Data Analysis

The diagnostic validity of a test was assessed by its ability to correctly discriminate between subjects with and without SI joint pain. The ideal diagnostic test would always be positive in subjects with the disease, and negative in those without the disease. In such a case, the sensitivity and specificity of the test would be 100%. The effect size of sensitivity and specificity can be reported as a diagnostic odds ratio (DOR). The DOR informs how much greater the odds of having the disease are for individuals with a positive test result than for individuals with a negative test result.[33] Sensitivity and specificity were extracted from individual studies, and 2 $\times$ 2 contingency tables were reconstructed if possible. In 2 cases, this information was not available in the published paper and we requested the original data from the first authors.[68,78] For the statistical analysis, Meta-DiSc software was used.[79] Confidence intervals for sensitivity, specificity and DOR were calculated for each test and subsequently tested for heterogeneity. The heterogeneity and consistency of studies in the meta-analysis was assessed with χ^2 and I^2 statistics, respectively. A heterogeneity result of $P < .05$ was considered significant, given the low power of the test. The I^2 values have a continuous scale of 0% to 100%, with 0% defining no inconsistency and 25%, 50%, and 75% were assigned as limits of low, medium and high inconsistency.[38] Statistical pooling was conducted only for results of provocative tests, and only for the data extracted from the prospective studies, and 95% confidence intervals of the pooled DORs were calculated. Because of a limited number of studies, it was not possible to produce an ROC curve. Finally, a qualitative descriptive analysis was performed.

Results

Results of the Search

In a PubMed database search, we identified 616 potentially relevant articles. We excluded 587 studies on the basis of their titles and abstracts. Subsequently, we retrieved and reviewed 29 full reports for possible inclusion.[7,9,12,13,21,22,25-27,40,42,44,45,47-50,55-57,59,61-63,67-69,73,78] Ten studies were excluded because they examined either reliability or repeatability of the test.[9,12,13,40,42,55-57,67,73] Two studies were excluded because they considered asymptomatic individuals.[21,26] This resulted in 17 studies that fulfilled our inclusion criteria.[7,22,25,27,44,45,47-50,59,61-63,68,69,78] A comprehensive search in the EMBASE and CINAHL databases resulted in 493 and 13 hits, respectively. After eliminating duplicates from PubMed, 305 references remained, of which only 1 article was included.[32]

Agreement Between Assessors

Agreement between the 2 reviewers was high (224/234, 95.7%), with a κ (using quadratic weighting) of 0.93 (95% CI, 0.87 to 0.97) for all criteria. Most disagreements regarded questions 6 and 7 and were principally due to either reading errors or differences in interpretation ("yes" or "no" against "unclear"). All disagreements but 1 were resolved during a consensus meeting.

Study Characteristics

Study characteristics are presented in Table 1. Nine studies were conducted in the 1990s,[7,22,25,27,47,48,59,62,63] whereas the remaining 9 studies were published after 2000.[32,44,45,49,50,61,68,69,78] Six studies evaluated pain mapping, pain area, or pain referrals from the SI joint,[22,25,32,59,61,69] which could be ascribed to the evaluation of the first IASP criterion. Six studies assessed the accuracy of pain provocation or stressing tests,[7,22,44,47,63,68] the second IASP criterion. One study evaluated the role of different contributors to low back pain, based on physical examination and positive outcome of various diagnostic infiltrations.[50] Additionally, 10 studies evaluated other kinds of clinical examination as the value of clinical history and mobilization tests,[22] mechanical examination of the lumbar spine,[45,47,78] pain provocation arthrography,[25,27,59] and bone scintigraphy.[48,49,62]

The majority of the studies were carried out in university hospitals and/or spine centers.[7,22,25,27,32,47-49,59,61-63] Patient recruitment was prospective in all but 3 studies[32,61,63] and consecutive in a minority of studies.[7,22,59,61-63] In only 8 studies, the recruitment period was reported, which ranged from 7 months[59] to 7 years.[32]

All studies included patients with low back pain. In 7 studies, inclusion criteria required only chronic low back pain complaints.[47-50,68,69,78] The age of the patients ranged from 18 to 87 years. The overall involvement of women in the included studies was more than 60% and ranged from 38% to78%.

Methodological Quality Assessment

Results of the methodological quality assessment are presented in Table 2. Overall, the majority of the studies scored positive for at least 8 of 14 QUADAS items. All the studies included representative individuals. However, in 4 studies, the inclusion of patients was limited to pain below the level of L5-S1[22,59,68,69] and in another 2 to pain in the buttock.[44,45] Furthermore, the majority of the studies also comprised patients with pain radiating to the groin or lower extremity, emphasizing its nonradicular character.[7,22,32,44,45,47,61-63,68,69] In 3 studies, the selection criteria were not described[25,27] or were not sufficiently described.[78] In 9 studies, the period between the index and the reference test was either not reported or was not clearly described,[25,27,32,48-50,59,63,78] which could lead to a disease progression bias (item 4). In 6 studies, only a selected sample of patients was verified using a reference test, and as the selection was not random this could indicate a partial verification bias.[25,27,44,45,50,78] Furthermore, in 10 studies, it was not clear whether or not the patients received the same reference standard regardless of the index test result[25,32,44,45,47,50,61-63,78] (item 6). In 3 studies, it was not clear whether the reference standard was independent of the index test,[27,49,50] and,

Table 2. Methodological Quality Assessment of Studies That Investigated the Diagnostic Accuracy of Provocative Tests for Subjects With Sacroiliac Joint Pain

Author , year[1]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	κ[2]
Broadhurst[7]	Y	Y	N	Y	Y	Y	N	Y	Y	U	U	Y	Y	Y	,82
Dreyfuss[22]	Y	Y	N	Y	Y	Y	Y	Y	Y	U	U	Y	Y	Y	1
Fortin[25]	Y	N	N	U	N	N	N	Y	Y	U	U	Y	N	N	1
Fortin[27]	Y	N	N	U	N	Y	U	Y	U	Y	N	Y	N	N	1
Fukui[32]	Y	Y	N	U	Y	U	Y	Y	Y	U	U	U	N	N	,88
Laslett[44,45]	Y	Y	Y	Y	N	U	Y	Y	Y	U	N	Y	Y	Y	,82
Maigne[47]	Y	Y	Y	Y	Y	U	U	N	U	U	U	Y	N	N	,64
Maigne[48]	Y	Y	N	U	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	1
Maigne[49]	Y	Y	N	N	Y	Y	Y	Y	Y	N	N	Y	Y	Y	,43
Manchikanti[50]	Y	Y	Y	U	N	N	U	U	Y	U	U	Y	Y	Y	1
Schwarzer[59]	Y	Y	N	U	Y	Y	N	U	Y	U	U	Y	Y	Y	1
Slipman[62]	Y	Y	N	Y	Y	N	Y	Y	U	U	U	Y	Y	Y	1
Slipman[63]	Y	Y	N	U	Y	U	Y	N	U	Y	N	Y	Y	Y	1
Slipman[61]	Y	Y	N	Y	Y	U	Y	U	Y	U	U	U	N	N	1
Van der Wurff[68]	Y	Y	Y	Y	Y	Y	Y	N	Y	U	U	Y	N	U	,85
Van der Wurff[68]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	1
Young[78]	Y	N	Y	U	N	N	Y	Y	Y	Y	Y	U	N	U	1

NOTE. [1]Publications are listed alphabetically by author. Items were scored as follows: "Y" means adequate methods, "N" means inadequate methods, and "U" means an item was inadequately described and therefore a decision could not be made whether it satisfied the criteria or not. [2]The κ score between the 2 reviewers (K.S. and P.v.d.W.) over the 14 points of QUADAS. [3]The observed proportion of overall agreement between the 2 reviewers considering particular QUADAS items.

in another 3, part of the index tests formed part of the reference test.[7,25,59] Eight studies provided insufficient description of the index and/or reference test to permit its replication.[27,47,50,59,61-63,69] Only 2 studies clearly notified that both the index and reference test were assessed blindly to each other's results.[48,78] Another interpretation bias could concern 3 studies whereby we were uncertain about the availability of the clinical data during the interpretation of test results by the examiners.[32,61,78] Finally, in 7 studies, we found bias associated with the report of study results and withdrawals.[25,27,32,47,61,69,78]

The Reference Test

All of the included studies used contrast enhanced intra-articular (intracavital) injections with local anesthetics as a reference test. The volume of contrast agent in the studies was usually very small and used only for confirmation of the intra-articular position of the needle, whereas in 3 studies, SI joint arthrography was obtained.[25,27,59] In 6 studies, double injections were used, whereby on separate occasions lidocaine and bupivacaine were infiltrated.[44,45,47,50,68,69] In addition, in 6 studies corticosteroids were used together with local anesthetics.[22,32,44,45,62,63] The volume of injected medicines varied between studies and ranged between 1 mL[59] and 4 mL,[7] but in most of the studies 2 mL was used. However, in 2 studies, the volume of infiltrated medicines was adjusted to the volume of the joint itself, which was determined by a previous arthrography.[25,27]

A positive effect of the diagnostic infiltration was predefined in most of the studies but varied considerably between studies. In 2 studies, 50% pain reduction was considered positive if it was sustained for at least 1 or 4 hours after infiltration with lidocaine or bupivacaine, respectively.[68,69] In 1 study, the level of pain produced by provocation tests performed before and after infiltration was measured, and 70% pain reduction after the infiltration was considered positive.[7] In 4 studies, 75% postinfiltration pain relief was required,[47-49,59] whereas in 6 studies, 80% pain reduction was required[44,45,61-63,78] and in 1 study, 90% pain reduction was required.[22] These variations in cut-offs for positive effect, however, do not influence the test results, as shown in previous studies.[16,17] In the remaining 4 studies, a positive effect of the infiltration was not predefined.[25,27,32,50]

Intra-articular infiltration was tested for its ability to diagnose SI joint pain. Schwarzer et al[59] hypothesized that if the intra-articular infiltration would relieve none of the patients' pain, then the concept of SI joint pain would be refuted. The authors estimated that the prevalence of SI joint pain could be as high as 13% to 30% of patients with low back pain. Maigne et al[47] evaluated the prevalence of SI joint pain in low back pain patients, whereby instead of single infiltration, double infiltration was used. Obtaining comparable pain relief of both injections would discriminate patients with SI joint pain. The prevalence of SI joint pain in their study was 18.5%. In a randomized, placebo-controlled study,[7] in contrast to lidocaine injections, none of the patients achieved 70% or more pain relief after placebo injections.

Diagnostic Value of the Tests

From the results of 7 studies, we could not construct 2 × 2 contingency tables because in the design of these studies a control group was lacking.[25,27,32,47,61,63,69] The majority of these studies examined pain mapping or pain referral patterns of the SI joint.[25,27,32,61,69] Most frequently reported was the area that overlies the posterior aspect of the SI joint, which is consistent with the first

IASP criterion. However, the occurrence of radiation to the buttock, groin and lower limb was not considered uncommon.[22,25,32,59,61,69] According to 1 study, the presence of pain in SI joint region or buttock characterizes a high sensitivity but very low specificity and DOR is lower than 1[22] (see Table 3), indicating no diagnostic utility for SI joint pain. Some investigators postulate that patients with presumed SI joint pain point out the area adjacent to the superior posterior iliac spine.[22,27] Based on the results of 1 study, this clinical test has reasonable specificity and sensitivity[22] with a DOR of 2.745 (95% CI, 0.99 to 7.63). Furthermore, based on the results of 1 study, the most intense pain area in patients with SI joint pain overlies the posterior margin of the SI joint.[69] This finding was consistent with the pain referral map, as determined by Fortin at al.[25]

With regard to the diagnostic validity of the second IASP criterion, in 1 study the calculation of sensitivity and specificity differed from our own calculation.[7] Taking a positive response to a local anesthetic to be true positive, a recalculation for resisted abduction, Patrick's sign and thigh thrust test yielded specificities of 87%, 77%, and 80%, respectively, and a sensitivity of 100% for all these tests. From the study of Dreyfuss et al,[22] we extracted only the results in which 2 examiners (physician and chiropractor) agreed completely in their findings. For the study of Laslett et al[44] the 2 × 2 contingency table was reconstructed for the distraction, compression, Gaenslen, thigh thrust, and sacral thrust test. From the study of Young et al,[78] we retrieved data of noncentralization of the pain and 3 or more positive provocation tests. Because these data were not consistent with the number of positive *(22)* and negative *(35)* responders to the SI joint infiltration as reported, we requested the raw data from the first author. Furthermore, we also requested the raw data for individual tests results from the study of van der Wurff et al.[68]

We tested the results of 5 individual provocation tests: compression, distraction, thigh thrust, Gaenslen's test, and Patrick's sign. The analysis of heterogeneity and consistency for the thigh thrust test was shown to be significant and inconsistent. The result of the study of Dreyfuss et al[22] was proved to be an "outlier" during the plot inspection. A comprehensive analysis did not clarify this outlier. The only explanation that we found were variations in the protocol of the reference test. The differences concern single versus double infiltrations, and the thresholds for a positive reference standard, ranging from 50% to 90% pain relief. Consequently, we chose for the subgroup analysis of the trial results concerning the thigh thrust test, which used the double SI joint infiltrations. The pooled sensitivity, specificity, and DOR of the thigh thrust test and the compression test showed that these tests have a discriminative power and are presented in Table 4. Results of Gaenslen's test, distraction test, and Patrick's sign could not be pooled due to the heterogeneity between the trials. Finally, the results of 4 studies that examined the accuracy of composition of provocation test for SI joint pain were pooled,[44,45,68,78] showing good diagnostic validity and discriminative power for SI joint pain (see Table 4).

Concerning the third IASP criterion, there are no data about the diagnostic validity of the complete pain relief after the selective infiltration of the SI joint. None of the included studies evaluated the IASP criteria set as a whole for SI joint pain.

Discussion

The purpose of the present review was to evaluate diagnostic validity of tests that could be ascribed to the IASP criteria for diagnosing SI joint pain. The first criterion, the presence of pain in the SI joint region, tests such as pain mapping or pain referral patterns, have an ability to correctly identify patients with SI joint pain. However, they fail in discriminating patients without SI joint pain. Furthermore, pain originating from the SI joint can extend to the buttock, groin, and even to the lower extremity.[22,25,32,59,61,69] In addition, it is doubtful whether the presence of pain in the region overlying the SI joint is exclusive for SI joint structures. It is well known that other anatomical structures in the low back, such as the disc and facet joints, are also capable of producing referred pain in the buttock region.[4,41] In this respect, we are facing a major clinical problem to differentiate between SI joint pain and other pain sources related to the lumbar and buttock region.

With regard to the second IASP criterion, 2 individual pain provocation tests—the compression and thigh thrust test—are helpful in diagnosing SI joint pain. Patients with a positive thigh thrust test or compression test are more likely to have SI joint pain. Subsequently, studies validating a comprehensive set of stressing tests proved good diagnostic validity of a threshold of 3 positive tests for diagnosing SI joint pain. Using a threshold of 3 or more positive stressing tests, the DOR of 3 positive provocation test is high in patients with SI joint pain (DOR, 17.2). However, when applying pain provocation tests, it is nearly impossible to define which structures actually are stressed.[47,54] Even structures such as the iliolumbar ligament or piriformis muscle cannot be excluded as potential source for this pain, since they are functionally related.[3,53] Consequently, it is very difficult to distinguish whether the provoked pain is exclusively intra-articular, or related to capsular ligaments.

Concerning the third IASP criterion, all of the trials included in our review used the selective infiltration as a reference test; however, the diagnostic validity of this test has received a little attention. Although no particular injection technique is recommended by the IASP, it is generally accepted to perform an intra-articular infiltration of a small volume of contrast medium to localize the joint, followed by a little amount of a local anesthetic. With this intra-articular technique, one has the ability to correctly determine symptomatic from asymptomatic SI joint patients.[7,47,59] It is surprising, however, that the selective infiltration targets only the joint cavity. Taking the basic anatomy of the SI joint into account, one may assume that the complaints originating from the SI joint could also involve neighboring SI joint ligaments. On the other hand, injecting even a very small volume of a local anesthetic into the joint cavity does not prevent leakage

Table 3. Diagnostic Validity of Tests

		CI			CI			CI	
Index test Author (ref.)	Sen	low	high	Sp	low	high	DOR	low	high
Broadhurst[7]									
Resisted abduction	1.0	0.81	1.0	0.87	0.66	0.97			
Patrick's sign	1.0	0.77	1.0	0.77	0.56	0.91			
Thigh thrust test	1.0	0.78	1.0	0.80	0.59	0.93			
Dreyfuss[22]									
SI joint pain	0.85	0.81	0.92	0.08	0.03	0.15	0.5	0.13	1.99
Groin pain	0.19	0.11	0.29	0.63	0.55	0.73	0.4	0.14	1.13
Buttock pain	0.80	0.74	0.88	0.14	0.07	0.23	0.66	0.21	2.16
Sitting position	0.03	0.01	0.08	0.9	0.87	0.96	0.28	0.04	2.12
PSIS pointing	0.76	0.65	0.85	0.47	0.35	0.57	2.75	1.0	7.52
Gillet	0.43	0.28	0.57	0.68	0.56	0.8	1.59	0.49	5.21
Thigh thrust	0.36	0.26	0.47	0.50	0.39	0.62	0.56	0.22	1.46
Patrick's	0.69	0.6	0.79	0.16	0.08	0.25	0.41	0.14	1.23
Gaenslen's	0.71	0.62	0.81	0.26	0.16	0.36	0.87	0.31	2.44
Sacral thrust	0.53	0.44	0.64	0.29	0.17	0.44	0.47	0.16	1.41
Spring test	0.75	0.64	0.85	0.35	0.22	0.47	1.6	0.49	5.23
"Sacral sulcus" tenderness	0.95	0.90	0.98	0.09	0.04	0.13	1.85	0.34	9.86
Laslett[44]									
Distraction test	0.6	0.36	0.8	0.81	0.65	0.91	5.95	1.6	22.15
Compression	0.69	0.44	0.86	0.69	0.5	0.84	4.84	1.33	17.67
Thigh thrust test	0.88	0.64	0.97	0.69	0.5	0.84	15.4	2.93	80.95
Gaenslen's test r	0.53	0.3	0.75	0.71	0.53	0.86	2.56	0.74	8.89
Gaenslen's test l	0.5	0.27	0.73	0.77	0.6	0.89	3	0.85	10.63
Sacral thrust	0.63	0.39	0.82	0.75	0.58	0.87	5	1.38	18.17
1 or more positive tests	1.00	0.84	1.00	0.44	0.59	0.44			
2 or more positive tests	0.94	0.76	0.99	0.66	0.57	0.68	28.36	4.11	187.18
3 or more positive tests	0.94	0.74	0.99	0.78	0.69	0.807	53.57	7.3	362.16
4 or more positive tests	0.60	0.40	0.76	0.81	0.72	0.89	6.5	1.72	24.71
5 or more positive tests	0.27	0.122	0.41	0.88	0.81	0.94	2.54	0.58	11.16
6 positive tests	0.07	0.012	0.201	0.88	0.85	0.94	0.5	0.07	3.78
Composition of 4 tests 1 or more +	1.00	0.84	1.0	0.47	0.39	0.47			
Composition of 4 tests 2 or more +	0.88	0.69	0.96	0.78	0.69	0.82	25.0	4.92	121.5
Composition of 4 tests 3 or more +	0.63	0.44	0.77	0.84	0.75	0.92	9.0	2.31	35.13
Composition of 4 tests 4 or more +	0.80	0.41	0.96	0.84	0.78	0.87	21.6	2.50	170.54
Laslett[45]									
3 or more positive tests n = 43	0.91	0.67	0.98	0.78	0.7	0.81	35.71	4.75	246.7
Positive clinical examination n = 34	0.91	0.69	0.98	0.87	0.76	0.91	66.67	7.3	540.23
Maigne[48]									
Radionuclide bone scanning	0.46	0.19	0.75	0.95	0.74	1.0	15.43	1.56	152.35
Maigne[49]									
Bone scintigraphy	0.43	0.24	0.62	0.65	0.55	0.76	1.42	0.39	5.22
Manchikanti[50]									
Single versus double SIJB	1.0	0.16	1.0	0.78	0.52	0.94	16.11	0.65	401.34
Schwarzer[59]									
Similar and exact pain reproduction during infiltration	0.85	0.55	0.98	0.47	0.28	0.66	4.81	0.91	25.53
Slipman[62]									
Radionuclide bone scan	0.13	0.36	2.98	1	0.82	1	6.38	0.33	125.5
Van der Wurff[68]									
3 or more positive tests	0.85	0.73	0.93	0.79	0.68	0.85	21.36	5.7	79.44
Distraction test	0.26	0.11	0.46	0.73	0.54	0.87	0.93	0.29	2.95
Compression test	0.6	0.39	0.78	0.7	0.51	0.84	3.3	1.15	9.73
Thigh Thrust test	0.93	0.76	0.99	0.64	0.45	0.8	21.9	4.4	108.9
Gaenslen's test	0.63	0.42	0.81	0.79	0.61	0.91	6.3	2.01	19.8
Patrick's sign	0.63	0.42	0.81	0.76	0.58	0.89	5.31	1.74	16.2
Young[78]									
No centralization of pain	0.9	0.79	0.97	0.2	0.13	0.25	2.59	0.54	12.06
3 or more positive tests	0.77	0.56	0.91	0.7	0.51	0.85	7.78	2.34	25.85

Table 4. Pooled Results of the Diagnostic Validity of Tests

	Sensitivity (CI)	*Specificity (CI)*	*DOR (CI)*
Compression test	0.628 (0.47-0.77)	0.692 (0.57-0.80)	3.885 (1.7-8.9)
Thigh thrust test	0.907 (0.78-0.97)	0.662 (0.53-0.77)	18.461 (5.82-58.53)
3 or more positive provocation tests	0.850 (0.75-0.92)	0.764 (0.68-0.84)	17.162 (7.6-39.0)

to the neighboring nerve structures and ligaments.[30] Consequently, there is a possibility that using this technique, more structures are targeted than the intra-synovial space. Furthermore, according to the literature, infiltration of interosseous ligament or the L4-S3 nerves, have, respectively, diagnostic and prognostic ability in patients with SI joint pain.[15,77] At present, the use of SI joint injections for the treatment of chronic complaints of nonspecific origin from this joint is not recommended by the European guidelines because of limited evidence of their efficacy.[1] However, the problem might be that treatment effects depend on the correct differentiation between the structures contributing to SI joint pain and are likely to rely on blocking of pain-signaling structures. Reports in the literature evaluating this subject are contradictory. Early cadaveric studies reporting dorsal as well as ventral innervation by the lumbar and sacral nerves[39,64] are not confirmed by a more recent report refuting ventral SI joint innervation.[34] Subsequently, (immuno-) histological studies of SI joint ligaments showed the presence of sensory nerves in the ventral capsular ligament,[28,66] the dorsal ligamentous tissue adjacent to the posterior superior iliac spine,[29,72] and in the interosseous ligament.[66] These findings, however, only concern ligamentous structures, thus structures surrounding the space targeted by the diagnostic infiltration. Taking the limitations of the diagnostic infiltration into account, the diagnostic validity of other tests handled in our review is controversial.

There are some discrepancies in the methodological assessment of the trials between our systematic review and other reviews in which the QUADAS tool has been used.[35,37,60] Although the κ of our quality assessment within review groups seems high, it is questionable if reliability across review groups is also high. The most rational explanation for discrepancies between groups may be the different interpretation of the QUADAS item list. Although the scoring of the QUADAS tool has been widely described,[74] it is necessary to incorporate the definition of the items in the review.[76] Also, the reliability of this instrument needs further testing. Furthermore, in the present review, the 4th QUADAS item was considered positive if the delay between the application of the index test and reference test used in the study was reported and was not more than 7 days, and, additionally, for SI joint infiltration with local anesthetics not shorter than 24 hours. For our definition, we considered the half-life of the anesthetics as the minimum period between the first and second injections. The maximal period between the index and the reference test is arbitrary because to our knowledge there is no literature concerning this point. However, in our opinion, for an individual study, a short-as-possible period between the index test and the reference test should be chosen to limit the influence of confounders, such as additional therapy. Consequently, we think that choosing a strict time interval between both tests would make our results more consistent.

The question arises about the relevance of the results of our review and whether they can contribute to the improvement of the current practice. Included studies were performed mainly in university hospitals and/or spine treatment units; their participants may differ from those in primary care, with respect to pain severity, chronicity and complexity of complaints and coexisting pathology. However, reaching the threshold of 3 positive tests, it is possible to select a certain subgroup within chronic low back pain patients, in which the diagnosis of SI joint pain could be involved. Still, attention should be paid to patients in whom singular test provokes familiar pain in SI joint, especially when using the compression or thigh trust test. Positive response to the individual pain provocation tests for SI joint could indicate a need for further diagnostics, which could involve intra-articular infiltrations with local anesthetics. However, using this technique, practitioners must be aware of its limitations.

In view of the fact that a gold standard for SI joint pain diagnosis is lacking, the diagnostic validity of other tests related to the IASP criteria for SI joint pain should be regarded with care. Taking this limitation into account, best evidence suggests that patients whereby at least 3 SI joint-selective stressing tests reproduce the patient's pain could be regarded to have SI joint pain. Reproduction of the patient's pain with compression or thigh thrust test indicates the need for further diagnostics. Relying only on the presence of pain in the SI joint region could lead to over-diagnosing of SI joint pain and is therefore not recommended. Intra-articular injections appear to be valid, but their diagnostic validity should be studied further. Moreover, information with respect to the IASP criteria as a whole set is lacking, and further research is needed in this area. Finally, sources of SI joint pain other than intra-articular, for example, peri-articular ligaments, should be considered for further evaluation.

Acknowledgments

We express our greatest gratitude to Ingrid Riphagen (Medical Library, VU University Medical Center) for her instructions and advice for the development of the literature search strategy. The work for this paper was performed in collaboration with the Pain Management and Research Center, University Hospital Maastricht, The Netherlands.

References

1. Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, Mannion AF, Reis S, Staal JB, Ursin H, Zanoli G: European guidelines for the management of chronic nonspecific low back pain. Eur Spine J 15(Suppl 2): S192-S300, 2006

2. Altman DG: Practical Statistics for medical research, New York, Chapman and Hall, 1999, pp 403-409.

3. Benzon HT, Katz JA, Benzon HA, Iqbal MS: Piriformis syndrome: Anatomic considerations, a new injection technique, and a review of the literature. Anesthesiology 98:1442-1448, 2003

4. Bernard TN, Kirkaldy-Willis WH: Recognizing specific characteristics of nonspecific low-back-pain. Clin Orthop Rel Res 1987;266-280, 1987

5. Berthelot JM, Labat JJ, Le Goff B, Gouin F, Maugars Y: Provocative sacroiliac joint maneuvers and sacroiliac joint block are unreliable for diagnosing sacroiliac joint pain. Joint Bone Spine 73:17-23, 2006

6. Bogduk N: International Spinal Injection Society guidelines for the performance of spinal injection procedures, I: Zygapophysial joint blocks. Clin J Pain 13:285-302, 1997

7. Broadhurst NA, Bond MJ: Pain provocation tests for the assessment of sacroiliac joint dysfunction. J Spinal Disord 11:341-345, 1998

8. Buijs E, Visser L, Groen G: Sciatica and the sacroiliac joint: a forgotten concept. Br J Anaesth 99:713-716, 2007

9. Carmichael JP: Inter- and intra-examiner reliability of palpation for sacroiliac joint dysfunction. J Manip Physiol Ther 10:164-171, 1987

10. Cattley P, Winyard J, Trevaskis J, Eaton S: Validity and reliability of clinical tests for the sacroiliac joint: A review of literature. Australas Chiropr Osteopathy 10:80, 2002

11. Chen YC, Fredericson M, Smuck M: Sacroiliac joint pain syndrome in active patients: A look behind the pain. Physician Sports Med 30:30-37, 2002

12. Chou LH, Slipman CW, Bhagia SM, Tsaur L, Bhat AL, Isaac Z, Gilchrist R, El Abd OH, Lenrow DA: Inciting events initiating injection-proven sacroiliac joint syndrome. Pain Med 5:26-32, 2004

13. Cibulka MT, Koldehoff R: Clinical usefulness of a cluster of sacroiliac joint tests in patients with and without low back pain. J Orthop Sports Phys Ther 29:83-89, 1999

14. Cohen SP: Sacroiliac joint pain: A comprehensive review of anatomy, diagnosis, and treatment. Anesth Analg 101: 1440-1453, 2005

15. Cohen SP, Abdi S: Lateral branch blocks as a treatment for sacroiliac joint pain: A pilot study. Reg Anesth Pain Med 28:113-119, 2003

16. Cohen SP, Bajwa ZH, Kraemer JJ, Dragovich A, Williams KA, Stream J, Sireci A, McKnight G, Hurley RW: Factors predicting success and failure for cervical facet radiofrequency denervation: A multi-center analysis. Reg Anesth Pain Med 32:495-503, 2007

17. Cohen SP, Stojanovic MP, Crooks M, Kim P, Schmidt RK, Shields CH, Croll S, Hurley RW: Lumbar zygapophysial (facet) joint radiofrequency denervation success as a function of pain relief during diagnostic medial branch blocks: A multicenter analysis. Spine J 8:498-504, 2008

18. Cooperman JM, Riddle DL, Rothstein JM: Reliability and validity of judgments of the integrity of the anterior cruciate ligament of the knee using the Lachman's test. Phys Ther 70: 225-233, 1990

19. Deville WLJM, Bezemer PD, Bouter LM: Publications on diagnostic test evaluation in family medicine journals: An optimal search strategy. J Clin Epidemiol 53:65-69, 2000

20. Dreyfuss P, Dreyer SJ, Cole A, Mayo K: Sacroiliac joint pain. J Am Acad Orthop Surg 12:255-265, 2004

21. Dreyfuss P, Dryer S, Griffin J, Hoffman J, Walsh N: Positive sacroiliac screening tests in asymptomatic adults. Spine 19: 1138-1143, 1994

22. Dreyfuss P, Michaelsen M, Pauza K, McLarty J, Bogduk N: The value of medical history and physical examination in diagnosing sacroiliac joint pain. Spine 21: 2594-2602, 1996

23. Foley BS, Buschbacher RM: Sacroiliac joint pain: Anatomy, biomechanics, diagnosis, and treatment. Am J Phys Med Rehabil 85:997-1006, 2006

24. Forst SL, Wheeler MT, Fortin JD, Vilensky JA: The sacroiliac joint: Anatomy, physiology and clinical significance. Pain Physician 9:61-67, 2006

25. Fortin JD, Aprill CN, Ponthieux B, Pier J: Sacroiliac joint-pain referral maps upon applying a new injection arthrography technique: 2. Clinical evaluation. Spine 19:1483-1489, 1994

26. Fortin JD, Dwyer AP, West S, Pier J: Sacroiliac joint-pain referral maps upon applying a new injection arthrography technique: 1. Asymptomatic volunteers. Spine 19:1475-1482, 1994

27. Fortin JD, Falco FJ: The Fortin finger test: An indicator of sacroiliac pain. Am J Orthop 26:477-480, 1997

28. Fortin JD, Kissling RO, O'Connor BL, Vilensky JA: Sacroiliac joint innervation and pain. Am J Orthop 28:687-690, 1999

29. Fortin JD, Vilensky JA, Merkel GJ: Can the sacroiliac joint cause sciatica? Pain Physician 6:269-271, 2003

30. Fortin JD, Washington WJ, Falco FJ: Three pathways between the sacroiliac joint and neural structures. AJNR Am J Neuroradiol 20:1429-1434, 1999

31. Freburger JK, Riddle DL: Using published evidence to guide the examination of the sacroiliac joint region. Phys Ther 81:1135-1143, 2001

32. Fukui S, Nosaka S: Pain patterns originating from the sacroiliac joints. J Anesth 16:245-247, 2002

33. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM: The diagnostic odds ratio: A single indicator of test performance. J Clin Epidemiol 56:1129-1135, 2003

34. Grob KR, Neuhuber WL, Kissling RO: Innervation of the human sacroiliacal joint. Z Rheumatol 54:117-122, 199.

35. Hancock MJ, Maher CG, Latimer J, Spindler MF, McAuley JH, Laslett M, Bogduk N: Systematic review of tests to identify the disc, SIJ or facet joint as the source of low back pain. Eur Spine J 10:165-184, 2007

36. Hansen HC, Helm S: Sacroiliac joint pain and dysfunction. Pain Physician 6:179-189, 2003

37. Hansen HC, McKenzie-Brown AM, Cohen SP, Swicegood JR, Colson JD, Manchikanti L: Sacroiliac joint interventions: A systematic review. Pain Physician 10: 165-184, 2007

38. Higgins JP, Thompson SG, Deeks JJ, Altman DG: Measuring inconsistency in meta-analyses. BMJ 327:557-560, 2003

39. Ikeda R: Innervation of the sacroiliac joint: Macroscopical and histological studies. Nippon Ika Daigaku Zasshi 58: 587-596, 1991

40. Katz V, Schofferman J, Reynolds J: The sacroiliac joint: A potential cause of pain after lumbar fusion to the sacrum. J Spinal Disord Tech 16:96-99, 2003

41. Kirkaldy-Willis WH, Hill RJ: A more precise diagnosis for low-back pain. Spine 4:102-109, 1979

42. Kokmeyer DJ, van der Wurff P, Aufdemkampe G, Fickenscher TCM: The reliability of multitest regimens with sacroiliac pain provocation tests. J Manip Physiol Ther 25: 42-48, 2002

43. Laslett M: Pain provocation tests for diagnosis of sacroiliac joint pain. Aust J Physiother 52:229, 2006

44. Laslett M, Aprill CN, McDonald B, Young SB: Diagnosis of sacroiliac joint pain: Validity of individual provocation tests and composites of tests. Manip Ther 10:207-218, 2005

45. Laslett M, Young SB, Aprill CN, McDonald B: Diagnosing painful sacroiliac joints: A validity study of a McKenzie evaluation and sacroiliac provocation tests. Aust J Physiother 49: 89-97, 2003

46. Lee JK, Yao L, Phelps CT, Wirth CR, Czajka J, Lozman J: Anterior cruciate ligament tears: MR imaging compared with arthroscopy and clinical tests. Radiology 166:861-864, 1988

47. 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 21: 1889-1892, 1996

48. Maigne JY, Boulahdour H, Chatellier G: Value of quantitative radionuclide bone scanning in the diagnosis of sacroiliac joint syndrome in 32 patients with low back pain. Eur Spine J 7:328-331, 1998

49. Maigne JY, Planchon CA: Sacroiliac joint pain after lumbar fusion: A study with anesthetic blocks. Eur Spine J 14: 654-658, 2005

50. Manchikanti L, Singh V, Pampati V, Damron KS, Barnhill RC, Beyer C, Cash KA: Evaluation of the relative contributions of various structures in chronic low back pain. Pain Physician 4:308-316, 2001

51. McKenzie-Brown AM, Shah RV, Sehgal N, Everett CR: A systematic review of sacroiliac joint interventions. Pain Physician 8:115-125, 2005

52. Merskey H, Bogduk N: Classification of chronic pain: Descriptions of chronic pain syndromes and definitions of pain terms, Seattle, WA, IASP Press, 1994. pp 190-191

53. Pool-Goudzwaard AL, Kleinrensink GJ, Snijders CJ, Entius C, Stoeckart R: The sacroiliac part of the iliolumbar ligament. J Anat 199:457-463, 2001

54. Potter NA, Rothstein JM: Intertester reliability for selected clinical tests of the sacroiliac joint. Phys Ther 65: 1671-1675, 1985

55. Pulisetti D, Ebraheim NA: CT-guided sacroiliac joint injections. J Spinal Disord 12:310-312, 1999

56. Riddle DL, Freburger JK: Evaluation of the presence of sacroiliac joint region dysfunction using a combination of tests: A multicenter intertester reliability study. Phys Ther 82:772-781, 2002

57. Robinson HS, Brox JI, Robinson R, Bjelland E, Solem S, Telje T: The reliability of selected motion- and pain provocation tests for the sacroiliac joint. Manip Ther 12:72-79, 2007

58. Rubinstein SM, Pool JJ, van Tulder MW, Riphagen II, de Vet HC: A systematic review of the diagnostic accuracy of provocative tests of the neck for diagnosing cervical radiculopathy. Eur Spine J 16:307-319, 2007

59. Schwarzer AC, Aprill CN, Bogduk N: The sacroiliac joint in chronic low-back-pain. Spine 20:31-37, 1995

60. Simpson R, Gemmell H: Accuracy of spinal orthopaedic tests: A systematic review. Chiropr Osteopat 14:26, 2006

61. Slipman CW, Jackson HB, Lipetz JS, Chan KT, Lenrow D, Vresilovic EJ: Sacroiliac joint pain referral zones. Arch Phys Med Rehabil 81:334-338, 2000

62. Slipman CW, Sterenfeld EB, Chou LH, Herzog R, Vresilovic E: The value of radionuclide imaging in the diagnosis of sacroiliac joint syndrome. Spine 21:2251-2254, 1996

63. 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. Arch Phys Med Rehabil 79:288-292, 1998

64. Solonen KA: The sacroiliac joint in the light of anatomical, roentgenological and clinical studies. Acta Orthop Scand 27:1-127, 1957

65. Stuber KJ: Specificity, sensitivity, and predictive values of clinical tests of the sacroiliac joint: A systematic review of the literature. J Can Chiropr Assoc 51:30-41, 2007

66. Szadek KM, Hoogland PV, Zuurmond WW, de Lange JJ, Perez RS: Nociceptive nerve fibers in the sacroiliac joint in humans. Regul Anesth Pain Med 33:36-43, 2008

67. Toussaint R, Gawlik CS, Rehder U, Ruther W: Sacroiliac joint diagnostics in the Hamburg construction workers study. J Manip Physiol Ther 22:139-143, 1999

68. van der Wurff P, Buijs EJ, Groen GJ: A multitest regimen of pain provocation tests as an aid to reduce unnecessary minimally invasive sacroiliac joint procedures. Arch Phys Med Rehabil 87:10-14, 2006

69. van der Wurff P, Buijs EJ, Groen GJ: Intensity mapping of pain referral areas in sacroiliac joint pain patients. J Manip Physiol Ther 29:190-195, 2006

70. van der Wurff P, Hagmeijer RHM, Meyne W: Clinical tests of the sacroiliac joint: A systematic methodological review, 1: Reliability. Manip Ther 5:30-36, 2000

71. van der Wurff P, Meyne W, Hagmeijer RHM: Clinical tests of the sacroiliac joint: A systematic methodological review, 2: Validity. Manip Ther 5:89-96, 2000

72. Vilensky JA, O'Connor BL, Fortin JD, Merkel GJ, Jimenez AM, Scofield BA, Kleiner JB: Histologic analysis of neural elements in the human sacroiliac joint. Spine 27: 1202-1207, 2002

73. Weiner DK, Sakamoto S, Perera S, Breuer P: Chronic low back pain in older adults: Prevalence, reliability, and validity of physical examination findings. J Am Geriatr Soc 54:11-20, 2006

74. Whiting P, Rutjes AW, Dinnes J, Reitsma JB, Bossuyt PM, Kleijnen J: Development and validation of methods for assessing the quality of diagnostic accuracy studies. Health Technol Assess 8:1-234, 2004

75. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J: The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 10: 25, 2003

76. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J: Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Med Res Methodol 6:9, 2006

77. Yin W, Willard F, Carreiro J, Dreyfuss P: Sensory stimulation-guided sacroiliac joint radiofrequency neurotomy: Technique based on neuroanatomy of the dorsal sacral plexus. Spine 28:2419-2425, 2003

78. Young S, Aprill C, Laslett M: Correlation of clinical examination characteristics with three sources of chronic low back pain. Spine J 3:460-465, 2003

79. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A: Meta-DiSc: A software for meta-analysis of test accuracy data. BMC Med Res Methodol 12:31, 2006

80. Zelle BA, Gruen GS, Brown S, George S: Sacroiliac joint dysfunction: Evaluation and management. Clin J Pain 21: 446-455, 2005

Appendix 1. Search Strategy Used to Identify Studies on Diagnostic Tests for SI Joint Pain

PubMed

Search conducted on August 29, 2007. ("Sacroiliac Joint"[MeSH] OR "Lumbosacral Region"[MeSH] OR sij[tiab] OR ((si[tiab] OR sacroiliac[tiab]) AND joint*[tiab]) OR (lumbosacral[tiab] AND region[tiab]))

AND

("Pain"[MeSH] OR "Pain Measurement"[MeSH] OR "Pain Clinics"[MeSH] OR "Low Back Pain"[MeSH] OR "Back Pain"[MeSH] OR "Myofascial Pain Syndromes"[MeSH] OR "Pain, Referred"[MeSH] OR "Pain Threshold"[MeSH] OR "Pain, Intractable"[MeSH] OR "Pelvic Pain"[MeSH] OR "Pain, Postoperative"[MeSH] OR "Somatosensory Disorders"[MeSH] OR "Neuralgia"[MeSH] OR "Arthralgia"[MeSH] OR "Somatoform Disorders"[MeSH] OR pain[tiab])

AND

("Injections"[Mesh:NoExp] OR "Injections, Intra-Articular"[Mesh] OR "Injections, Spinal"[Mesh:NoExp] OR inject*[tiab] OR preinjection*[tiab] OR infiltrat*[tiab] OR (pain[tiab] AND (pattern*[tiab] OR mapping*[tiab])) OR algometr*[tiab] OR Pain measurement[mesh] OR provocation[tw] OR block*[tw] OR test[tiab] OR tests[tiab] OR standard[tiab] OR standards*[tiab])

AND

(Diagnosis[mesh] OR Diagnosis[sh] OR Etiology[sh] OR diagnos*[tw] OR specificity[Title/Abstract] OR "Sensitivity and Specificity"[MeSH])

NOT

(animals[mesh] NOT humans[mesh])

NOT

case reports[pt].

EMBASE

Search conducted on September 9, 2007.

Sacroiliac-joint/de OR Lumbosacral-spine/de OR sij:ti,ab OR ((si:ti,ab OR sacroiliac:ti,ab) AND joint*:ti,ab) OR (lumbosacral:ti,ab AND region:ti,ab)

AND

Backache/exp OR Pain-assessment/exp OR Pain-Clinics/exp OR pelvis-pain-syndrome/de OR Neuralgia/exp OR Arthralgia/exp OR musculoskeletal-pain/de OR pain:ti,ab OR arthralgia:ti,ab OR neuralgia:ti,ab OR backache:ti,ab OR somatosensory-disorder/exp OR 'somatoform disorder'/de OR 'psychosomatic disorder'/de OR psychogenic-pain/de

AND

injection/exp OR 'intraarticular drug administration'/exp OR inject*:ti,ab OR preinjection*:ti,ab OR infiltrat*:ti,ab OR (pain:ti,ab AND (pattern*:ti,ab OR mapping*:ti,ab)) OR algometr*:ti,ab OR 'Pain assessment'/exp OR 'sensory system examination'/de OR provocation:ti,ab OR block*:ti,ab OR test:ti,ab OR tests:ti,ab OR standard:ti,ab OR standards*:ti,ab OR 'nerve block'/de OR 'physical examination'/de

AND

Diagnos* OR etiolog* OR Diagnosis/de OR 'diagnostic accuracy'/de OR 'diagnostic value'/de OR 'differential diagnosis'/de OR 'diagnostic test'/de OR 'physical examination'/de OR 'Pain assessment'/exp OR 'sensory system examination'/de OR 'gold standard'/exp

AND

'human'/de.

CINAHL

Search conducted on September 9, 2007.

[(sacroiliac joint) or (MM "Sacroiliac Joint")] and [(low back pain) or (MM "Low Back Pain") or (MM "Back Pain") or (arthralgia) or (MH "Arthralgia+")] and [(injection) or (MM "Injections, Intraarticular") or (block) or (MH "Nerve Block") or (MM "Pain Measurement") or (MM "Algometry") or (diagnostic test) or (MH "Diagnostic Tests, Routine")] and [(diagnosis) or (MM "Diagnosis") or (MM "Diagnosis, Differential") or (MM "Sensitivity and Specificity/MT/ST")] not (MM "Animals") not (MM "Case Studies").