The diagnostic value of three sacroiliac joint pain provocation tests for sacroiliitis identified by magnetic resonance imaging

B Arnbak1,2, AG Jurik1,2,3, RK Jensen1,2, B Schiøttz-Christensen1,2, P van der Wurff4,5, TS Jensen1,2

1Research Department, Spine Centre of Southern Denmark, Hospital Lillebaelt, Middelfart, Denmark, 2Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark, 3Department of Radiology, Aarhus University Hospital, Aarhus, Denmark, 4Research and Development, Military Rehabilitation Centre Aardenburg, Doorn, The Netherlands, and 5Department of Physiotherapy, HU University of Applied Sciences, Utrecht, The Netherlands

Objectives: The aim of the current study was to investigate the diagnostic value of three sacroiliac (SI) joint pain provocation tests for sacroiliitis identified by magnetic resonance imaging (MRI) and stratified by gender.

Method: Patients without clinical signs of nerve root compression were selected from a cohort of patients with persistent low back pain referred to an outpatient spine clinic. Data from Gaenslen’s test, the thigh thrust test, and the long dorsal sacroilia ligament test and sacroiliitis identified by MRI were analysed.

Results: The median age of the 454 included patients was 33 (range 18–40) years and 241 (53%) were women. The prevalence of SI joints with sacroiliitis was 5%. In the whole study group, only the thigh trust test was associated with sacroiliitis, the area under the receiver operating characteristic (ROC) curve (AUC) was 0.58 [95% confidence interval (CI) 0.51–0.65], sensitivity 31% (95% CI 18–47), and specificity 85% (95% CI 82–87). In men, sacroiliitis was associated with all the SI joint tests assessed and multi-test regimens, with the greatest AUC found for at least one positive out of three tests [AUC 0.68 (95% CI 0.56–0.80), sensitivity 56% (95% CI 31–79), and specificity 81% (95% CI 77–85)]. In women, no significant associations were observed between the SI joint tests and sacroiliitis.

Conclusions: Only in men were the SI joint tests found to be associated with sacroiliitis identified by MRI. Although, the diagnostic value was relatively low, the results indicate that the use of SI joint tests for sacroiliitis may be optimized by gender-separate analyses.

Spondyloarthritis is a relatively common inflammatory disorder with a frequency of 1–2% in the European population (1). However, because of its insidious onset and delayed appearance of radiographic changes, the diagnosis of spondyloarthritis is often delayed 5–10 years from the onset of symptoms (2). During this period, undiagnosed patients may experience severe pain and stiffness and have potentially progressive loss of function, contributing to reduced health-related quality of life (3). As new treatment possibilities have been shown to reduce both symptoms and inflammation, efforts have been made to improve early referral and diagnosis of spondyloarthritis.

Although sacroiliac (SI) joint pain provocation tests have been shown to correlate with SI joint dysfunction (4), and their use is recommended in clinical guidelines for identification of low back pain (LBP) originating from the SI joint (5), the value of SI joint tests for the purpose of identifying patients with sacroiliitis has been sparsely investigated (6). Thus, to date, no physical examination tests have been included in referral recommendations for spondyloarthritis. The current recommendations are based primarily on pain characteristics, extraspinal symptoms, and analysis of human leucocyte antigen (HLA)-B27 (3, 7, 8). Sacroiliitis is considered the keystone in the diagnosis of spondyloarthritis (9, 10) and magnetic resonance imaging (MRI) has been proven to be superior to scintigraphy (11, 12), computed tomography (11, 13), and conventional radiography (12, 13) in diagnosing early sacroiliitis. However, to reduce the number of unnecessary MRI scans, efficient and feasible identification of patients with possible sacroiliitis is needed. As physical examination tests are low-cost and simple to perform, it would be ideal if SI joint provocation tests could be used to identify patients at risk of having sacroiliitis. Consequently, there is a need for further investigation of the validity of SI joint provocation tests for sacroiliitis.

The majority of studies on the diagnostic performance of SI joint pain provocation tests have been performed on mixed gender samples but generally with an over-representation of women (4). However, the anatomy of the pelvic ring, along with its load distribution, varies between genders (14), and thus it is possible that the SI joint pain provocation tests perform differently in men.
and women. Therefore, the aim of the current study was to analyse the diagnostic value of three SI joint pain provocation tests for sacroiliitis in a mixed gender sample of patients with LBP and stratified by gender.

Method

Setting

The Spines of Southern Denmark (SSD) cohort consisted of young patients with persistent LBP referred to the Spine Centre of Southern Denmark, where their data were collected with the purpose of investigating the early diagnosis of spondyloarthritis. The Spine Centre is an outpatient, non-surgical unit specializing in managing non-inflammatory back pain in a secondary care public hospital setting. The unit performs multidisciplinary assessment of patients with spinal pain after referral from general practitioners, chiropractors, and medical specialists in primary care.

Patients

During the study period, from March 2011 to October 2013, 16 clinicians (physiotherapists, chiropractors, and medical doctors) were allocated to a multidisciplinary ‘project team’. Based on the appointment availability in the project team and access to the MRI scanner, the booking secretaries randomly allocated patients to the project team in a consecutive manner. Patients aged 18 to 40 years referred with LBP as the primary complaint were allocated to the project. The clinicians in the project team excluded patients from the project who did not understand Danish or had contraindications for MRI. A study inclusion/exclusion flow chart is presented in Figure 1.

Of the patients referred with LBP in the study period, 1619 patients aged 18–40 years were randomly selected and invited to participate in the study. Other clinical teams in the Spine Centre attended to the remaining patients according to normal clinical procedures at the Spine Centre. One hundred and sixty were excluded before, and 422 after, the first visit; thus, in total, 1037

![Figure 1. Study inclusion/exclusion flow chart. LBP, Low back pain; MRI, magnetic resonance imaging; SSD, Spines of Southern Denmark.](www.scandjrhumatol.dk)
patients were included in the SSD cohort (Figure 1). For the current analysis, only patients without clinical signs of nerve root compression and with no missing data on the three SI joint pain provocation tests were included. Signs of nerve root involvement were defined as leg pain on a pain drawing, in addition to at least one of the following clinical findings associated with the site of the leg pain: muscle weakness, altered sensation to touch or pinprick, impaired tendon reflexes, a positive straight leg raise test (at 60° or less), or a positive prone knee bend test combined with pain to the anterior thigh.

The study was conducted according to the Declaration of Helsinki and, before inclusion, each patient gave written informed consent for research use and publication of their de-identified data. The Regional Scientific Ethics Committee for Southern Denmark determined that, under the Danish legal framework, this study did not require formal ethics approval (reference number S-20102000-58).

**Clinical characteristics**

Data on patients’ demographic and clinical characteristics were collected using internet-based patient self-reported questionnaires as part of the Spine Centre’s standard procedure.

**Clinical investigation and SI joint tests**

As a part of the standard procedure at the Spine Centre, information from the clinical examination was collected using an internet-based evaluation form. Information collected from the physical examination included neurological examination and three SI joint pain provocation tests (positive or negative): Gaenslen’s test (15), the thigh trust test (16, 17), and the long dorsal sacroiliac ligament test (18). The SI joint tests were considered positive when reproducing pain in the SI joint (Table 1). All 16 members of the multidisciplinary project team participated in the collection of the SI joint test data.

**MRI**

Details of the scan protocol have been published previously (19). In brief, MRI of the whole spine and SI joints was performed with a 1.5-T MRI System (Philips Achieva, Best, The Netherlands). For the SI joint, the following sequences were used: semi-coronal T1-weighted turbo spin–echo (TSE), semi-coronal T1-weighted acquisition with spectral pre-saturation inversion recovery (SPIR), and semi-axial T2-weighted short-tau inversion recovery (STIR).

Three spondyloarthritis expert radiologists, blinded to all clinical information except the patients’ age and gender, participated in the reading of the MRI scans. Each MRI scan was evaluated by one reader and, in the case of any uncertainty, consensus was reached by

---

**Table 1. Sacroiliac (SI) joint tests evaluated in the study.**

| Test name | Technique |
|-----------|-----------|
| Gaenslen’s test (pelvic torsion test) (15) | Supine, affected side close to the edge and partially off the examination table |
| Thigh trust test (posterior pelvic pain provocation test) (16, 17) | Supine, patient’s knee is flexed to the chest. The patient is then shifted to the side of the examination table so that the opposite leg extends over the edge while the other leg remains flexed |
| Long dorsal sacroiliac ligament test (18) | Prone, the examiner tests for tenderness on palpation of the long dorsal sacroiliac ligament directly under the caudal part of the posterior superior iliac spine |

In the current study, all tests were considered positive if pain in the SI joint was reproduced.
consulting another reader (6.1% of the evaluations). In the included patients, the median time between the clinical examination and the MRI scan was 15 days [inter-quartile range (IQR) 13–21 days].

Each joint was subdivided into four osseous locations: the cartilaginous and ligamentous portion of the iliac and the sacral bones, respectively (20). Sacroiliitis was defined as a minimum of two SI joint regions with bone marrow oedema lesions or a minimum of one SI joint region with bone marrow oedema at > 25% of the periarticular area, and fulfilling the minimum definition of bone marrow oedema used in the Assessment of SpondyloArthritis international Society’s (ASAS) criteria for axial spondyloarthritis (21). The agreement of the evaluation of bone marrow oedema at the SI joint has previously been tested with kappa values > 0.8 for inter- and intra-observer agreement (19). An imaging example of sacroiliitis is shown in Figure 2.

Data analysis

Information from the physical examination, the self-reported questionnaire, and the MRI evaluation were entered directly into an electronic database [the SpineData database (22)] and were analysed using STATA version 13.1 (StataCorp, College Station, TX, USA).

Descriptive data were tabulated. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were estimated and are presented as percentages with 95% confidence intervals (CIs) for each SI joint test and for ≥ 1 positive out of three tests, ≥ 2 positive out of three tests, and all three tests positive. The area under the receiver operating characteristic (ROC) curve (AUC) was used to assess the diagnostic value and is reported with 95% CI. For binary predictions, the AUC equals balanced accuracy, that is (sensitivity + specificity)/2 (23). The AUC can range from 0.5 (useless model) to 1.0 (perfect discrimination). The following general interpretation of AUC was used: < 0.6 means no clinical value, 0.6–0.7 limited value, 0.7–0.8 modest value, and > 0.8 discrimination adequate for genuine clinical utility (24).

To test for statistically significant difference between groups, a two-sample Wilcoxon rank-sum test was used for continuous and ordinal categorical variables, and Pearson’s χ² test was used for binary variables. The significance level was set at 5%.

Results

Of the 1037 patients included in the SSD cohort, 106 were excluded from the current analyses because of clinical signs of nerve root compression and 477 because of missing data on the SI joint pain provocation tests. The remaining 454 were included in the analysis (Figure 1). The patients excluded due to missing values were generally similar to the included patients regarding the prevalence of sacroiliitis and the descriptive and clinical data (see Supplementary Table S1 for details).

The median age of the 454 included patients was 33 (IQR 27–37) years and 241 (53%) were women. A total of 36 patients had changes observed on MRI classified as ‘sacroiliitis’ in one or two SI joints. The remaining 418 patients were classified as ‘non-inflammatory LBP’. Descriptive and clinical data for the two groups are presented in Table 2. Patients with MRI-identified sacroiliitis were more often on sick leave compared with patients classified as non-inflammatory LBP. No differences between the groups were found on the other parameters shown in Table 2.

Women were twice as likely to have a positive test response compared with men (Table 3). Of the 908 included joints, 45 (5%) were classified with sacroiliitis. There was no statistically significant gender difference in the prevalence of sacroiliitis (4%) in men vs. (6%) in women, p = 0.12).

Diagnostic value of the assessed SI joint tests

The sensitivity, specificity, PPV, NPV, and AUC for each of the SI joint pain provocation tests calculated for the whole group are shown in Table 4. Only the thigh trust test was significantly associated with sacroiliitis in the whole group.

The diagnostic parameters of the SI joint pain provocation tests for men and women, respectively, are shown in Table 5. In men, sacroiliitis was significantly associated with all the SI joint tests and multi-test regimens, with the greatest AUC found for at least one positive test out of three tests. In women, no significant associations were found between the SI joint tests assessed and sacroiliitis.
Table 2. Demographic and clinical characteristics in ’sacroiliitis’ and ‘non-inflammatory LBP’ patients.

| Characteristics                                      | Sacroiliitis n = 36 | Non-inflammatory LBP n = 418 |
|------------------------------------------------------|----------------------|-----------------------------|
| Age (years), median (IQR)                            | 31 (26–35)           | 33 (27–37)                  |
| Women, % (95% CI)                                    | 64 (47–80)           | 52 (47–57)                  |
| Employed, % (95% CI)                                 | 69 (52–85)           | 73 (69–77)                  |
| Sick leave due to LBP in the previous 3 months, % (95% CI) | 68 (51–84)           | 46 (41–51)*                 |
| LBP duration in months, median (IQR)                 | 20 (5–40)            | 12 (4–41)                   |
| Previous LBP episode(s), % (95% CI)                  | 79 (65–94)           | 73 (69–77)                  |
| LBP intensity (0–10), median (IQR)                   | 6 (5–7)              | 6 (5–7)                     |
| Leg pain, % (95% CI)                                 | 83 (70–96)           | 77 (73–81)                  |
| Leg pain intensity (0–10), median (IQR)              | 4 (3–6)              | 4 (2–6)                     |
| Activity limitation (RMDQ 0–100), median (IQR)       | 61 (43–78)           | 52 (39–74)                  |

IQR, Interquartile range; CI, confidence interval; LBP, low back pain [LBP intensity is averaged on 0–10 numerical rating scales on present LBP, worst LBP last 14 days, and typical LBP last 14 days (25); Leg pain intensity measured in the same way as for LBP; RMDQ, Roland Morris Disability Questionnaire (26) [calculated as a proportional score (0% = no activity limitation; 100% = maximum activity limitation) (27)].

n varies for each item due to missing values.* p < 0.05. † For patients with leg pain > 0.

Table 3. Prevalence of sacroiliac joints with positive tests in men and women.

|                              | Men n = 426 joints | Women n = 482 joints | All n = 908 joints |
|------------------------------|--------------------|----------------------|-------------------|
| Single tests                 |                    |                      |                   |
| Gaenslen’s test              | 7 (5–10)           | 14 (11–17)*          | 11 (9–13)         |
| Thigh trust test             | 10 (7–13)          | 21 (18–25)*          | 16 (14–18)        |
| Long dorsal sacroiliac ligament test | 17 (14–21)        | 38 (34–43)*          | 28 (25–31)        |
| Multi-test regimens          |                    |                      |                   |
| ≥ 1 positive out of three tests | 21 (17–25)         | 41 (36–45)*          | 31 (28–34)        |
| ≥ 2 positive out of three tests | 9 (6–11)           | 20 (17–24)*          | 15 (13–17)        |
| All three tests positive     | 5 (3–8)            | 12 (9–15)*           | 9 (7–11)          |

Values are percentages with 95% confidence intervals in parentheses.
Difference in prevalence between men and women: * p < 0.01.

Table 4. Diagnostic value of sacroiliac joint tests for sacroiliitis.

|                              | Sensitivity | Specificity | PPV | NPV  | AUC      | p-value |
|------------------------------|-------------|-------------|-----|------|----------|---------|
| Single tests                 |             |             |     |      |          |         |
| Gaenslen’s test              | 13 (5–27)   | 89 (87–91)  | 6 (2–13) | 95 (94–97) | 0.51 (0.46–0.57) | 0.573 |
| Thigh trust test             | 31 (18–47)  | 85 (82–87)  | 10 (5–16) | 96 (94–97) | 0.58 (0.51–0.65) | 0.005 |
| Long dorsal sacroiliac ligament test | 36 (22–51) | 72 (69–75) | 6 (4–10) | 96 (94–97) | 0.54 (0.47–0.61) | 0.276 |
| Multi-test regimens          |             |             |     |      |          |         |
| ≥ 1 positive out of three tests | 42 (28–58) | 69 (66–72)  | 7 (4–10) | 96 (94–97) | 0.56 (0.48–0.63) | 0.105 |
| ≥ 2 positive out of three tests | 24 (13–40) | 86 (83–88)  | 8 (4–14) | 96 (94–97) | 0.55 (0.49–0.62) | 0.064 |
| All 3 tests positive         | 13 (5–27)   | 91 (89–93)  | 7 (3–15) | 95 (94–97) | 0.52 (0.47–0.57) | 0.317 |

PPV, Positive predictive value; NPV, negative predictive value; AUC, area under the receiver operating characteristic (ROC) curve [equals (sensitivity + specificity)/2].
Values given are percentages with 95% confidence intervals in parentheses.
Pre-test probability for sacroiliitis in a joint: 5%.
Table 5. Diagnostic value of sacroiliac joint tests for sacroiliitis among men and women.

|                  | Men                          | Women                       |
|------------------|------------------------------|------------------------------|
|                  | Sensitivity | Specificity | PPV     | NPV     | AUC   | p-value | Sensitivity | Specificity | PPV     | NPV     | AUC   | p-value |
| Single tests     |             |             |         |         |       |         |             |             |         |         |       |         |
| Gaenslen’s test  | 22 (6–48)   | 93 (91–96)  | 13 (4–30) | 97 (94–98) | 0.58 | 0.48–0.68 | 0.013 | 7 (1–24) | 86 (82–89) | 3 (0–10) | 94 (91–96) | 0.47 | 0.41–0.52 | 0.316 |
| Thigh trust test | 39 (17–64)  | 91 (88–94)  | 16 (7–30) | 97 (95–99) | 0.65 | 0.53–0.77 | 0.000 | 26 (11–46) | 79 (75–83) | 7 (3–14) | 95 (92–97) | 0.53 | 0.44–0.61 | 0.533 |
| Long dorsal      | 44 (22–69)  | 84 (80–88)  | 11 (5–21) | 97 (95–99) | 0.64 | 0.52–0.76 | 0.002 | 30 (14–50) | 61 (56–66) | 4 (2–8)  | 94 (90–96) | 0.45 | 0.36–0.54 | 0.336 |
| sacroiliac test  |             |             |         |         |       |         |             |             |         |         |       |         |
| Multi-test       |             |             |         |         |       |         |             |             |         |         |       |         |
| ≥ 1 positive     | 56 (31–79)  | 81 (77–85)  | 11 (6–20) | 98 (95–99) | 0.68 | 0.56–0.80 | 0.000 | 33 (17–54) | 59 (54–64) | 5 (2–9)  | 94 (90–96) | 0.46 | 0.37–0.56 | 0.425 |
| out of three tests |             |             |         |         |       |         |             |             |         |         |       |         |
| ≥ 2 positive     | 28 (10–54)  | 92 (89–95)  | 14 (5–29) | 97 (94–98) | 0.60 | 0.49–0.71 | 0.003 | 22 (9–42) | 80 (76–83) | 6 (2–13) | 95 (92–97) | 0.51 | 0.43–0.59 | 0.802 |
| All three tests  | 22 (6–48)   | 95 (93–97)  | 17 (5–39) | 97 (94–98) | 0.59 | 0.49–0.69 | 0.001 | 7 (1–24) | 87 (84–90) | 3 (0–12) | 94 (91–96) | 0.47 | 0.42–0.53 | 0.415 |

PPV, Positive predictive value; NPV, negative predictive value; AUC, area under the receiver operating characteristic (ROC) curve [equals (sensitivity + specificity)/2].

Values given are percentages with 95% confidence intervals in parentheses.

Pre-test probability for sacroiliitis in a joint: 4% for men and 6% for women.

Discussion

In this study, we investigated the diagnostic value of SI joint pain provocation tests for identifying early onset of persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP. In the whole study sample, only one of the SI joint pain provocation tests was found to be useful in identifying patients with sacroiliitis in a large sample of young patients with persistent LBP.
Overall, there appears to be a need for further investigation of the different causes of SI joint pain among men and women. This may also be beneficial for the evaluation of SI joint tests for SI joint dysfunction in general. The majority of previous studies on SI joint tests have evaluated men and women at the same time, but women seem to be over-represented in the study populations. In a systematic review on the diagnostic validity of SI joint tests regarding patients with SI joint pain, the overall prevalence of women in the included studies was 60% (4). However, the results from the current study indicate that the diagnostic value of SI joint pain provocation tests could be optimized by conducting gender-separate analyses.

The methodological strengths of the current study are, first, the large number of patients included in the analyses, which strengthened the precision of the estimates of diagnostic value and also made gender-separate analysis possible. Second, the independent interpretation of information from the MRI evaluation and the physical examination reduced the risk of bias. Third, the time period between the clinical examination and the MRI evaluation was relatively short, reducing the risk of changes regarding onset of sacroiliitis in the time span after the examination.

One of the limitations of the current study was that the primary aim with the data collection was not to evaluate SI joint tests. Consequently, the performance of the pain provocation tests was not standardized and the reproducibility of the test procedure was not evaluated in the study. Other studies have found an acceptable reliability for two of the three included SI joint tests (Gaenslen’s test and the thigh thrust test), although there is a general need for further investigation of the reliability of SI joint tests (31). Moreover, because the SI joint test data were collected and recorded as part of the daily clinical examination, there was a relatively high number of missing values in the current study. Thus, it is unknown whether the relatively low diagnostic value of the tests observed in the current study could partly be due to clinicians inconsistently interpreting the tests prior to recording. However, we expect the results to be more generalizable to daily clinical practice, as the data were collected as a part of the everyday routine in the Department.

Finally, the current study investigated only three tests whereas a multi-test regimen with four or more tests is recommended based on results from previous studies on the validity of tests for SI joint dysfunction (4). Moreover, it is possible that pain provocation tests, which correlate with SI joint dysfunction, may differ from those correlating with sacroiliitis, which investigating a greater variety of tests might reveal. Thus, the observed diagnostic value in the current study may be improved by investigating combinations of more tests in future studies.

In summary, the three SI joint pain provocation tests that were assessed were all positively associated with sacroiliitis among men, but with relatively low diagnostic value, while no associations were found between the SI joint tests and sacroiliitis in women. These results reveal that the diagnostic value of pain provocation tests for sacroiliitis may be optimized by gender-separate analyses. Moreover, the results indicate a need for further investigation of the different causes of SI joint pain among men and women, respectively, which also could add to knowledge regarding tests for SI joint dysfunction in general.

Acknowledgements

We thank the patients and staff of the Spine Centre of Southern Denmark who participated in the data collection, and Suzanne Capell, academic English language editor, for proofreading this manuscript. The study was supported financially by the Foundation for Chiropractic Research and Postgraduate Education and the Knud and Edith Eriksen Foundation.

References

1. Akkoc N. Are spondyloarthropathies as common as rheumatoid arthritis worldwide? A review. Curr Rheumatol Rep 2008;10:371–8.
2. Feldtkeller E, Khan MA, van der Heijde D, van der Linden S, Braun J. Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. Rheumatol Int 2003;23:61–6.
3. Sieper J, Rudwaleit M. Early referral recommendations for ankylosing spondylitis (including pre-radiographic and radiographic forms) in primary care. Ann Rheum Dis 2005;64:659–63
4. Szadek KM, Van Der Wurff P, Van Tulder MW, Zuurmond WW, Perez RS. Diagnostic validity of criteria for sacroiliac joint pain: a systematic review. J Pain 2009;10:354–68.
5. Vleeming A, Albert HB, Ostgaard HC, Sturesson B, Stuge B. European guidelines for the diagnosis and treatment of pelvic girdle pain. Eur Spine J 2008;17:794–819.
6. Ozgocmen S, Bozgeyik Z, Kalcik M, Yildirim A. The value of sacroiliac joint pain provocation tests in early active sacroiliitis. Clin Rheumatol 2008;27:1275–82.
7. Rudwaleit M, Van Der Heijde D, Khan MA, Braun J, Sieper J. How to diagnose axial spondyloarthritis early. Ann Rheum Dis 2004;63:535–43.
8. Brandt HC, Spiller I, Song IH, Vahlidiek JL, Rudwaleit M, Sieper J. Performance of referral recommendations in patients with chronic back pain and suspected axial spondyloarthritis. Ann Rheum Dis 2007;66:1479–84.
9. Rudwaleit M, van der Heijde D, Landewe R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68:777–83.
10. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361–8.
11. Battafarano DF, West SG, Rak KM, Fortenberry EJ, Chantelois AE. Comparison of bone scan, computed tomography, and magnetic resonance imaging in the diagnosis of active sacroiliitis. Semin Arthritis Rheum 1993;23:161–76.
12. Blum U, Buitrago-Tellez C, Mundinger A, Krause T, Laubenger J, Vaith P, et al. Magnetic resonance imaging (MRI) for detection of active sacroiliitis—a prospective study comparing conventional radiography, scintigraphy, and contrast enhanced MRI. J Rheumatol 1996;23:2107–15.
13. Yu W, Feng F, Dion E, Yang H, Jiang M, Genant HK. Comparison of radiography, computed tomography and magnetic resonance imaging in the detection of sacroiliitis accompanying ankylosing spondylitis. Skeletal Radiol 1998;27:311–20.
14. Dietrichs E. Anatomy of the pelvic joints—a review. Scand J Rheumatol Suppl 1991;19(Suppl 88):4–6.

www.scandjrheumatol.dk
15. Gaenslen F. Sacroiliac arthrodesis: indications, author's technic and end-results. J Am Med Assoc 1927;89:2031–5.
16. Ostgaard HC, Zetherstrom G, Roos-Hansson E. The posterior pelvic pain provocation test in pregnant women. Eur Spine J 1994;3:258–60.
17. Laslett M, Williams M. The reliability of selected pain provocation tests for sacroiliac joint pathology. Spine (Phila Pa 1976) 1994;19:1243–9.
18. Vleeming A, de Vries HJ, Mens JM, van Wingerden JP. Possible role of the long dorsal sacroiliac ligament in women with peripartum pelvic pain. Acta Obstet Gynecol Scand 2002;81:430–6.
19. Ambak B, Jensen TS, Manniche C, Zejden A, Egund N, Jurik AG. Spondyloarthritis-related and degenerative MRI changes in the axial skeleton - an inter- and intra-observer agreement study. BMC Musculoskelet Disord 2013;14:274.
20. Madsen KB, Jurik AG. Magnetic resonance imaging grading system for active and chronic spondylarthritides in the sacroiliac joint. Arthritis Care Res (Hoboken) 2010;62:11–18.
21. Rudwaleit M, Jurik AG, Hermann KG, Landewe R, Van Der Heijde D, Baraliakos X, et al. Defining active spondylitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. Ann Rheum Dis 2009;68:1520–7.
22. Kent P, Kongsted A, Jensen TS, Albert HB, Schiottz-Christensen B, Manniche C. SpineData - a Danish clinical registry of people with chronic back pain. Clin Epidemiol 2015;7:369–80.
23. Cantor SB, Kattan MW. Determining the area under the ROC curve for a binary diagnostic test. Med Decis Making 2000;20:468–70.
24. Ohman EM, Granger CB, Harrington RA, Lee KL. Risk stratification and therapeutic decision making in acute coronary syndromes. JAMA 2000;284:876–8.
25. Manniche C, Asmussen K, Lauritsen B, Vinterberg H, Kreiner S, Jordan A. Low back pain rating scale: validation of a tool for assessment of low back pain. Pain 1994;57:317–26.
26. 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 (Phila Pa 1976) 1983:8:141–4.
27. Kent P, Lauridsen HH. Managing missing scores on the Roland Morris disability questionnaire. Spine (Phila Pa 1976) 2011;36:1878–84.
28. Gran JT, Huby G, Hordvik M. Prevalence of ankylosing spondylitis in males and females in a young middle-aged population of Tromso, northern Norway. Ann Rheum Dis 1985;44:359–67.
29. Calin A, Fries JF. Striking prevalence of ankylosing spondylitis “healthy” w27 positive males and females. N Engl J Med 1975;293:835–9.
30. Albert HB, Godskesen M, Westergaard JG. Incidence of four syndromes of pregnancy-related pelvic joint pain. Spine (Phila Pa 1976) 2002;27:2831–4.
31. Van Der Wurff P, Hagmeijer RH, Meyne W. Clinical tests of the sacroiliac joint. A systematic methodological review. Part 1: reliability. Man Ther 2000;5:30–6.

Supporting Information
Additional Supporting Information may be found in the online version of this article.

Supplementary Table S1. Demographic and clinical characteristics in patients excluded due to missing values and patients included in the study.
Please note that the editors are not responsible for the content or functionality of any supplementary material supplied by the authors. Any queries should be directed to the corresponding author.