Specificity, sensitivity, and predictive values of clinical tests of the sacroiliac joint: a systematic review of the literature

Stuber K J

CRD summary
This review found a lack of evidence on clinical tests of the sacroiliac joint. However, the review suffered from a number of methodological limitations, especially in relation to the assessment of study quality and the failure to present appropriate study details. The findings should therefore be interpreted with caution.

Authors' objectives
To evaluate the accuracy of physical examination tests for the presence of sacroiliac (SI) joint injuries and/or dysfunction.

Searching
MEDLINE, EMBASE, CINAHL, AMED and ICL were searched from inception to 2005. The search terms, which were reported, included a diagnostic filter. Six relevant journals were handsearched, references of retrieved articles were screened and experts were contacted to identify unpublished studies and conference abstracts. No language restrictions were applied.

Study selection
Studies that evaluated any physical examination test of the SI joint, used SI joint block or injection as the reference standard, included patients with mechanical low back pain of likely SI joint origin, and reported data on sensitivity, specificity and/or predictive values, were eligible for inclusion. Studies of pregnant women, patients with degenerative joint disease, inflammatory arthropathies, malignancy or other systemic diseases were excluded.

The included studies evaluated a variety of different clinical tests.

The author did not state how the papers were selected for the review, or how many reviewers performed the selection.

Assessment of study quality
Studies were assessed for methodological quality using a 10-item form developed by the author. The tool considered description of: study population, inclusion and exclusion criteria, number of participants, clinical test, examiner, outcome measures of test, reference standard and examiner conducting the reference standard. The tool also considered blinding to group allocation and the person interpreting the reference standard, randomisation to active gold standard group or placebo group, use of confirmatory diagnostic blocks and the reporting of appropriate accuracy data. Studies were assigned points for each item fulfilled to give a maximum possible quality score of 100.

The author did not state how many reviewers performed the validity assessment.

Data extraction
Data were extracted on the sensitivity, specificity and predictive values for tests used singly or in combination. A pre-formulated data extraction tool, which enabled the extraction or calculation of sensitivity, specificity and predictive values, was used.

The author did not state how many reviewers performed the data extraction.

Methods of synthesis
The data were combined in a narrative.
Results of the review
Five studies reported in six publications were included in the review (number of participants unclear). An additional relevant study could not be obtained.

Study quality scores ranged from 52.5 to 90 out of 100.

The included studies evaluated a large number of different tests. The sensitivity of these tests ranged from 36 to 95% and the specificity ranged from 9 to 100%. Even when studies evaluated the same test there was considerable heterogeneity in estimates of accuracy. Using composite tests based on one or more positive clinical tests appeared to offer improved sensitivity and specificity combinations.

Authors' conclusions
There is a lack of evidence on most clinical SI joint tests. None of the tests evaluated was overwhelmingly superior to the others.

CRD commentary
This review addressed a focused question that was supported by clearly defined inclusion criteria. Although a broad range of databases were searched, the search included a diagnostic filter and it is therefore likely that some relevant studies were missed. Attempts were made to limit language and publication bias. Details of the review process were not reported, but given that the review has only one author it is likely that appropriate steps were not taken to minimise bias and errors. A detailed quality assessment was undertaken but this focused largely on reporting and included items more appropriate for randomised controlled trials, rather than diagnostic accuracy studies. The results of the quality assessment were presented in detail but failure to consider important quality indicators for diagnostic studies, such as the inclusion of an appropriate patient spectrum and avoidance of verification bias, means that the validity of the included studies remains unclear. Very few details of the included studies were provided, so the generalisability of the review findings is unclear. A narrative synthesis was appropriate given the differences between the studies, but considering sensitivity and specificity separately in the results and tables, and stratifying by study rather than test, make the findings difficult to follow. The author's conclusions should be interpreted with some degree of caution given the possibility that studies might have been missed, study quality was not appropriately assessed and individual study details were lacking.

Implications of the review for practice and research
Practice: The author stated that the only tests which have a sensitivity greater than 60% in at least one study, and which practitioners may consider using, are the distraction test, compression test, thigh thrust/posterior shear, sacral thrust and resisted hip abduction.

Research: The author stated that further research using improved methodology is required to determine optimum test combinations to identify SI joint interventions.

Funding
University of Sheffield School of Health and Related Research Training in Health bursary.

Bibliographic details

PubMedID
17657289

Original Paper URL

Indexing Status
Subject indexing assigned by CRD

MeSH
Diagnostic Tests, Routine; Sacroiliac Joint; Sensitivity and Specificity

AccessionNumber
12007009201

Date bibliographic record published
09/08/2008

Date abstract record published
01/12/2008

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.