The diagnostic validity of clinical tests in temporomandibular internal derangement: a systematic review and meta-analysis
Chaput E, Gross A, Stewart R, Nadeau G, Goldsmith CH

CRD summary
This review concluded that deflection alone, crepitation alone or crepitation, deflection, pain and limited mouth opening, were the most valuable for identifying internal derangement without reduction. The quality and quantity of the evidence on which the conclusions were based was very low and the conclusions seem too firm.

Authors' objectives
To assess the diagnostic accuracy of clinical tests for temporomandibular internal derangement compared with magnetic resonance imaging (MRI).

Searching
MEDLINE and EMBASE were searched for English-language articles from 1994 to October 2009; the search strategy was reported.

Study selection
Prospective cohort and case-control studies recruiting patients with temporomandibular joint disorders or internal derangement with or without reduction were eligible for inclusion. Patients had to be at least 14 years old. Studies had to use tests that were easily performed with minimal equipment, either alone or in combination. MRI had to include the sagittal and coronal planes.

Where reported in the included studies, the mean age ranged from 26.4 to 55.3 years, the vast majority of participants were female, and most studies were conducted in tertiary care settings. Most studies combined the sagittal and coronal views from the MRI. Some studies assessed single index tests and a similar number assessed combinations of tests.

Three reviewers independently selected studies for the review.

Assessment of study quality
Study quality was assessed by two independent reviewers using the 14-point Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool. Studies scoring less than nine were excluded. The Grades of Recommendation Assessment, Development and Evaluation (GRADE) system was used to assess the generalisability of the studies; the quality of the evidence was considered to be high, moderate, low or very low.

Data extraction
Three reviewers independently extracted data to construct 2x2 tables of test performance, from which the sensitivity, specificity, and positive and negative likelihood ratios, with 95% confidence intervals, were calculated.

Methods of synthesis
The sets of data were categorised based on the MRI view, type of index test, and pathological condition. Pooled estimates of diagnostic accuracy were derived using a DerSimonian and Laird random-effects model. Heterogeneity was assessed using $\chi^2$ and $I^2$. Subgroup analyses were planned to investigate the impact of study quality and pre-test probability. Publication bias was not investigated as there were too few studies.

Results of the review
Eight studies met the inclusion criteria, with 952 patients (range 40 to 242). On QUADAS, three studies scored the full 14 points, two scored 12, one scored 11, one scored 10, and one scored nine; progression bias was the most common potential bias, with only three studies avoiding it.

The presence of a click on palpation was rarely important for identifying internal derangement (four studies), with reduction (one study) or without it (three studies), as was the presence of pain on palpation, functioning and opening
(five studies). Deflection alone (LR+ 6.37, 95% CI 2.13 to 19.03), crepitation alone (LR+ 5.88, 95% CI 1.95 to 17.76), and crepitation, deflection, pain and limited mouth opening altogether (LR+ 6.37, 95% CI 2.13 to 19.03), were of significant value in identifying internal derangement without reduction. Click, deviation, and pain together ruled out internal derangement with reduction (LR- 0.09 95% CI 0.01 to 0.72). The quality of the evidence was very low or low for these estimates.

**Authors’ conclusions**
Limited evidence suggests deflection alone, crepitation alone or the cluster of crepitation, deflection, pain and limited mouth opening, were the most valuable for identifying internal derangement without reduction. The test cluster click, deviation, and pain ruled out internal derangement with reduction. No single test or cluster of tests was conclusive and of significant value for ruling in internal derangement with reduction.

**CRD commentary**
The authors addressed a clear review question with reproducible inclusion criteria. The search was limited, and there were no attempts to reduce potential language or publication bias. Diagnostic filters were used, and studies could have been missed. Each stage of the review process was conducted in duplicate, reducing the risks of error and bias. Appropriate criteria were used to assess study quality, a summary score was produced, and the results were reported in full. There was substantial clinical variation across the studies. Pooled estimates of sensitivity and specificity were derived from analyses that did not maintain the within-study relationship between these measures, and that used a small subset of studies. A random-effects model was used where too few studies were available to assess the distribution of accuracy across the studies.

The quality and quantity of the evidence on which the conclusions were based was low, and the suggestions for the value of single or combinations of tests seem too strong.

**Implications of the review for practice and research**
- **Practice**: The authors did not state any implications for practice.
- **Research**: The authors stated that future research should focus on improving quality, should use both sagittal and coronal views on the MRI, and should recruit larger samples of patients.

**Funding**
Not stated.

**Bibliographic details**

**Original Paper URL**
http://utpjournals.metapress.com/content/c4g25415678x3527/

**Indexing Status**
Subject indexing assigned by CRD

**MeSH**
Temporomandibular Joint Disorders; Humans; Predictive Value of Tests; Diagnosis; Magnetic Resonance Imaging

**Accession Number**
12012026357

**Date bibliographic record published**
12/10/2012

**Date abstract record published**
13/11/2012

**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.