Are chiropractic tests for the lumbo-pelvic spine reliable and valid: a systematic critical literature review

Hestboek L, Leboeuf-Yde C

Authors' objectives
To investigate the evidence for intra-examiner reliability, inter-examiner reliability, and the validity of different chiropractic tests used to diagnose the presence of manipulative lesions in the lumbo-pelvic spine area.

Searching
CHIROLARS was searched from 1976 to 1995 using the index terms 'chiropractic tests', 'chiropractic adjusting technique', 'motion palpation', 'movement palpation', 'leg length', 'applied kinesiology' and 'sacrooccipital technique'. In addition, a manual search was performed at the libraries of the Nordic Institute of Chiropractic and Clinical Biomechanics (Odense, Denmark) and the Anglo-European College of Chiropractic, (Bournemouth, UK). Before completion, a superficial update was conducted through manual searches of Spine and the Journal of Manipulative and Physiological Therapeutics from October 1996 to October 1998. Only studies that were published as a journal article and written in English were included. The studies also had to have been published between January 1, 1976 and December 31, 1995 and be available to the chiropractic profession.

Study selection
Study designs of evaluations included in the review
No a priori inclusion criteria were reported in relation to the study design. Studies that were anecdotal, speculative, or editorial in nature were excluded. Studies that contained only information that was published elsewhere and already included in the review were also excluded.

Specific interventions included in the review
Tests or examination methods used by chiropractors, which were aimed at detecting a lesion and were suitable for manipulation, were eligible for inclusion. The tests investigated were required to be manual or visual only, without the use of instrumentation. Studies that included a whole regimen of tests or methods without separate data for each test were excluded. Studies that explored relationships between tests, and did not relate them to the patient's condition or indication for manipulation, were also excluded.

The types of chiropractic test used in the studies included: motion palpation of the lumbar spine; motion palpation of the sacroiliac joints; leg-length inequality; sacrooccipital technique; palpation for tenderness; and a combination of several tests. Some studies included several tests (parallel testing) performed by the same examiner(s). The professionals undertaking the examination varied considerably, comprising experienced or student chiropractors, sacrooccipital practitioners, physiotherapists, and instructors of the activator technique. The number of examiners ranged from 2 to 45 in different intra-examiner studies.

Reference standard test against which the new test was compared
No a priori inclusion criteria were specified with regard to the reference standard. The reference standards reported by validity studies included mechanical models to study motion palpitation, radiography and patient-reported low-back pain (LBP).

Participants included in the review
No a priori inclusion criteria were specified with regard to the participants, other than the included studies had to pertain to the human lumbar spine. Within the included studies, the participants varied considerably from patients with chronic LBP to asymptomatic students.

Outcomes assessed in the review
No a priori inclusion criteria were specified with regard to the outcome measures. Only studies that evaluated intra-examiner reliability, inter-examiner reliability, and/or validity of the chiropractic tests were included. The measures of...
validity reported in the included studies were sensitivity, specificity, likelihood ratios and the diagnostic odds ratio. The authors also looked at whether the results varied in relation to the experience of the examiner (i.e. students versus qualified chiropractors or experts), and in relation to whether the test was initiated by the examiner (i.e. observation) or by the patient (i.e. pain report).

How were decisions on the relevance of primary studies made?
One reviewer screened the retrieved articles to determine their eligibility according to predefined criteria.

Assessment of study quality
Specific checklists were constructed to quantify the quality of each study for intra-examiner reliability, inter-examiner reliability and validity. All of the articles were evaluated once according to these checklists, and then the checklists were adjusted. Each article was subsequently given a quality score, which consisted of the proportion of acceptable items in relation to the total number of relevant items on the checklist (maximum score 8). The scoring system was described in the text. The authors also investigated whether the quality of the studies was related to the year of publication, the technique under scrutiny, or the study design used. Two reviewers independently assessed the quality of the data; the results were then compared. In cases of disagreement, the matter was discussed until consensus was reached. The conclusions of the review were drawn from studies with a quality score of greater than or equal to 80%.

Data extraction
Data on the following categories were extracted: reference details; test(s) studied; study participants; type of examiner; description of the statistics used; description of any problems; whether the study met any of the additional quality scoring items (such as blinding); and the overall quality score. The authors did not state how many reviewers performed the data extraction.

Methods of synthesis
How were the studies combined?
A narrative synthesis of the studies was undertaken.

How were differences between studies investigated?
Differences between the studies were discussed in the text.

Results of the review
Thirty original articles containing 61 studies were included. Fifteen studies looked at the intra-examiner reliability (n=5), inter-examiner reliability (n=7) and validity (n=3) of motion palpation of the lumbar spine. Six studies looked at the intra-examiner reliability (n=2), inter-examiner reliability (n=3) and validity (n=2) of motion palpation of the sacroiliac joints. Fourteen studies looked at the intra-examiner reliability (n=3), inter-examiner reliability (n=5) and validity (n=6) of leg-length inequality. Eight studies looked at the intra-examiner reliability (n=2), inter-examiner reliability (n=4) and validity (n=2) of sacrooccipital technique. Nine studies looked at the intra-examiner reliability (n=2), inter-examiner reliability (n=4) and validity (n=3) of palpation for pain. Two studies looked at the intra-examiner reliability (n=2) of palpation for muscle tension. One study looked at the intra-examiner reliability (n=1) of palpation for misalignment. Four studies looked at the intra-examiner reliability (n=1) and inter-examiner reliability (n=3) of visual inspection.

None of the test studies had been sufficiently evaluated in relation to reliability and validity. Only tests for palpation for pain had consistently acceptable results. Motion palpation of the spine might be valid, but it showed poor reliability, whereas motion palpation of the sacroiliac joints seemed to be slightly reliable, but was not shown to be valid. Measures of leg-length inequality seemed to correlate with radiographic measurements, but consensus on method and interpretation was lacking. For sacrooccipital techniques, some evidence favoured the validity of the arm-fossa test, but the rest of the test regimen remains poorly documented. Documentation of applied kinesiology was not available. Palpation for muscle tension, palpitation for misalignment, and visual inspection were either undocumented, unreliable, or not valid.
The quality scores ranged from 22 to 100% (average 72%). Thirty studies achieved a minimum score of 80% (9 of the 13 intra-examiner studies, 14 of the 23 inter-examiner studies, and 7 of the 15 validity studies). The statistical analyses were often unclear. The quality of pre-1985 studies was generally poor. There was no correlation between year of publication and quality after 1984, or between quality and the test studied.

**Authors' conclusions**
The detection of the manipulative lesion in the lumbo-pelvic spine depends on valid and reliable tests. Such tests have not been established, therefore the presence of the manipulative lesion remains hypothetical. Great effort is needed to develop, establish, and enhance valid and reliable test procedures.

**CRD commentary**
The review included a clear objective, as well as predefined inclusion and exclusion criteria, although the a priori inclusion criteria in relation to type of participants, study design and outcome measures were not stated. One electronic database was searched and only studies published between January 1, 1976 and December 31, 1995 were included. While some additional handsearching was undertaken, some important information may still have been missed. Only studies published in English were included, thus publication bias cannot be ruled out. Details of the review process were clearly presented and the validity of the included studies was assessed.

Relevant details of the primary studies were presented both in tabular format and as a narrative, along with comments on the validity of each study. In view of the great variation between the included studies, the narrative synthesis of the results was appropriate. The authors' conclusions follow from the results.

**Implications of the review for practice and research**
Practice: The authors noted that, to date, no manual or visual test has been identified that fulfils the minimal criteria of consistent reliability and validity. Until such tests have been established, the presence of the manipulative lesion remains hypothetical.

Research: The authors suggested that an expert panel designs a series of acceptable study protocols for the different types of study designs, and that the various chiropractic institutions divide the tasks of systematic testing of the most commonly used chiropractic test procedures among themselves.

**Bibliographic details**

**PubMedID**
10820299

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Chiropractic /methods; Evaluation Studies as Topic; Female; Humans; Lumbar Vertebrae; Lumbosacral Region; Male; Manipulation, Spinal /methods; Observer Variation; Reproducibility of Results; Sensitivity and Specificity; Spinal Diseases /diagnosis /therapy

**AccessionNumber**
12000001185

**Date bibliographic record published**
30/11/2004
Date abstract record published
30/11/2004

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.