Authors' objectives
To assess the diagnostic performance of the straight-leg-raising (SLR) test (also known as Lasegue) and the cross SLR test to detect herniated discs in patients with low-back pain.

Searching
MEDLINE and EMBASE were searched (from 1992 to 1997) to update a search from an earlier review; the search terms were listed. The reference lists from identified publications were checked.

Study selection

Study designs of evaluations included in the review
The authors did not explicitly state which study designs were eligible for inclusion; all of the studies included in the review were surgical case series. Studies involving fewer than 10 patients were excluded, as were reviews.

Specific interventions included in the review
Studies evaluating the SLR test or the cross SLR test were eligible for inclusion.

Reference standard test against which the new test was compared
Studies that used surgery as the reference standard were eligible for inclusion.

Participants included in the review
Studies involving patients presenting with low-back pain were included in the review, whereas studies of Cauda Equina syndrome were excluded.

Outcomes assessed in the review
Studies providing sensitivity data, or sensitivity and specificity data on the SLR or cross SLR test were eligible for inclusion.

How were decisions on the relevance of primary studies made?
One reviewer selected the studies for inclusion.

Assessment of study quality
The authors used the Cochrane checklist on 'Meta-analysis of diagnostic and screening tests' to assess the quality of the included studies (see Other Publications of Related Interest no.1). Two further quality checkpoints, taken from Mulrow et al., supplemented these criteria; these referred to the cut-off point of the reference test and the purpose of the test (see Other Publications of Related Interest no.2). Two reviewers independently assessed the papers for quality. Any disagreements were resolved through consensus and agreement between the reviewers was measured using Cohen's kappa.

Data extraction
The authors did not state how the data were extracted for the review, or how many reviewers performed the data extraction.

The extracted data included: year of study, disease prevalence at the setting, sample size and size of the smallest group, prior surgery, level of hernia, and method of participant recruitment (prospective or retrospective). The sensitivity and specificity of the diagnostic odds ratio (DOR) of the SLR and cross SLR tests were calculated from the absolute numbers provided in the studies, unless these were unavailable.
Methods of synthesis
How were the studies combined?
Studies with sensitivity and specificity data were selected for pooling. The average predictive values were based on the pooled sensitivity and specificity estimates at the mean prevalence of the pooled studies. The DOR was calculated for each study. Eleven of the studies evaluating the SLR test were combined in a meta-analysis; a second meta-analysis included 6 studies evaluating the cross SLR test. A summary receiver operating characteristic (ROC) curve was presented.

How were differences between studies investigated?
Statistical heterogeneity of sensitivity and specificity across the studies was investigated using the chi-squared test. Statistical heterogeneity of the DORs was assessed using the method of DerSimonian and Laird. A fixed-effect model was used if the studies were not statistically significantly heterogeneous. A univariate regression was used to assess the effect of the independent variables: cut-off point of the index test, validity score, numerical information about study characteristics, and study population.

Results of the review
Fifteen studies were included in the review. Eleven studies evaluating the SLR test and 6 studies assessing the cross SLR test were pooled in separate meta-analyses.

The observer agreement for classifying quality criteria was 76.5% (range: 32.9 to 100); the overall agreement was moderate (kappa 0.56) for all criteria. The median score was 3 out of 6 (50%) for internal validity and 5 out 11 (45%) for external validity. Four studies had a score of 4 out of 6 for internal validity. The median total validity was 47%; 6 of the 15 studies scored 50% or more for total validity.

Results of the meta-analysis for the SLR test.
Statistical heterogeneity was significant for sensitivity and specificity. Using a random-effects model, the pooled sensitivity (11 studies) was 0.91 (95% confidence interval, CI: 0.82, 0.94) and the pooled specificity was 0.26 (95% CI: 0.16, 0.38). Significant heterogeneity was found for the DORs. The unweighted pooled DOR was 3.74 (95% CI: 1.23, 11.4). An outlying study was excluded from further analysis, as the studies were homogeneous in its absence. The pooled DOR after exclusion of the outlying study was 2.96 (95% CI: 2.05, 4.28).

The mean predictive value of a positive result at the average prevalence of 0.86 was 0.89; the negative predictive value was 0.33.

The regression found a strong negative correlation between natural log (ln) of the DOR and year of publication, where the DOR was lower in recently published studies. A significant positive correlation was found between presence of verification bias and ln(DOR).

Results of the meta-analysis for the cross SLR test.
Statistical heterogeneity was significant for sensitivity and specificity. Using a random-effects model, the pooled sensitivity (6 studies) was 0.29 (95% CI: 0.24, 0.34) and the pooled specificity was 0.88 (95% CI: 0.86, 0.90). The DORs were homogeneous. The unweighted pooled DOR was 4.39 (95% CI: 0.74, 25.9).

The predictive value of a positive test was 0.92 at a prevalence of 0.82; the negative predictive value was 0.22.

Authors' conclusions
The diagnostic accuracy of the SLR test was limited by its low specificity. More valid designs, more homogeneous case-mixes and year of publication showed a decrease in discriminatory power. The studies did not enable a valid evaluation of the diagnostic accuracy of the SLR test or cross SLR test.
CRD commentary
The review question and inclusion criteria were relatively clear, although it was not possible to decipher whether the population was restricted to adults. It was also unclear whether the authors included studies published in all languages, and publication bias was not discussed or measured. The authors did not present all the study details that they claimed to have extracted, although the impact of population differences were discussed. A thorough quality assessment was performed and was used in the investigation of heterogeneity. The authors found significant heterogeneity for sensitivity and specificity, therefore, the decision to statistically combine the studies might not have been appropriate. Differences between the studies were suitably investigated. The authors acknowledged a number of other limitations, including poor study quality, inappropriate design and generalisability, which are reflected in their cautious conclusions.

Implications of the review for practice and research
Practice: The authors did not state any implications for practice.

Research: The authors stated that diagnostic research needs to evaluate the validity of the complete diagnostic process and investigate the evidence on the added value of different tests.

Bibliographic details

PubMedID
10788860

Other publications of related interest

Indexing Status
Subject indexing assigned by NLM

MeSH
Adult; Confidence Intervals; Female; Humans; Intervertebral Disc Displacement /diagnosis /surgery; Leg; Male; Middle Aged; Odds Ratio

AccessionNumber
12000001059

Date bibliographic record published
31/03/2005

Date abstract record published
31/03/2005

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.