Systematic review

# Diagnostic accuracy of self-report and subjective history in the diagnosis of low back pain with non-specific lower extremity symptoms: A systematic review

Susan Shultz, Kristina Averell, Angela Eickelman, Holly Sanker, Megan Burrowbridge Donaldson*

*Walsh University, 2020 East Maple Street, North Canton, OH 44720, USA*

## ARTICLE INFO

## ABSTRACT

Subjective history questions/self-report items are commonly used to triage the patient with low back pain and related leg symptoms. However the value of the history taking process for decision-making to identify common classifications/diagnosis for patients presenting with low back related leg pain (LBRLP) have not been considered. The purpose of this study was to investigate the diagnostic accuracy of self-report items/history-taking questions used to identify patients with LBRLP.

Eligible studies included: 1)subjects with low back pain AND related lower extremity pain, 2)details of subjective examination/self-report items, 3)cohort, prospective/longitudinal studies, and randomized control trials, 4)use of statistical reporting, 5)an acceptable reference standard. Quality was evaluated using the *Quality Assessment of Diagnostic Accuracy Studies 2.* A synthesis of history items that met the threshold for at least a small shift in the likelihood of the condition with a $+LR \geq 2$ or $-LR \leq 0.5$ were reported.

Conditions commonly reported in the literature: lumbar spinal stenosis, lumbosacral nerve root compression/radiculopathy, disc herniation and neurophysiological low back pain $\pm$ leg pain. Eleven studies met the inclusion criteria.

This is the first systematic review of diagnostic accuracy studies that examined only the history-taking items for their ability to identify LBRLP conditions. Clustering key items may provide a more precise clinical picture necessary to detect and treat a patient's presentation. History questions formed within the interview and their contributing value for decision-making remain understudied. There is a need for better designs to determine a more accurate diagnostic power to identify conditions with LBRLP.

© 2014 Elsevier Ltd. All rights reserved.

## 1. Background

Low-back and low-back related leg pain complaints are frequently addressed by healthcare practitioners. These conditions can be seriously debilitating to patients and impose a significant social and economic burden on the community (Delitto et al., 2012). Accompanying leg pain is present in approximately 25–57% of all LBP cases (Shäfer et al., 2009). Low-back related leg pain with or without nerve root involvement is also associated with a poor prognosis compared to low back pain (LBP) alone (Konstantinou et al., 2012). In order to treat these conditions effectively, healthcare practitioners ask key questions related to symptom presentation that help generate or eliminate probable diagnoses (Deyo et al., 1992).

Initially, the patient self-report questionnaire items and the subjective examination/history-taking is often used early to help clinicians generate probable hypotheses which may help differentiate those patients with pain of musculoskeletal origin from those with non-spinal or serious spinal pathology (Rubinsten, & van Tulder, 2008). Initial diagnosis may present significant challenges for healthcare practitioners due a level of uncertainty with history questions provided by a patient (Hill et al., 2008). Based on the literature clinicians cannot distinguish with reliable accuracy between those patients with benign conditions and those with radicular pain or serious spinal pathology (Cook et al., 2011).

* Corresponding author.
   *E-mail addresses:* mismeg77@aol.com, mdonaldson@walsh.edu (M.B. Donaldson).

Often, the pain associated with low back-related leg pain is unclear.

In most primary care settings, clinicians and or physicians may do a history, physical exam and possibly imaging. An accurate diagnostic strategy should include both components (Deyo et al., 1992). However the reliance on advanced imaging findings to confirm the presence of neurologic compression to triage for potential surgical referrals is expensive and has an unacceptably high rate of false-positive results (Simon et al., 2009). Furthermore, there are several studies that have examined the accuracy of individual clinical index tests used to identify spinal conditions, and have found them not be helpful in clarifying the cause of low back related leg pain (van der Windt et al., 2010; Iverson et al., 2013). When examining the evidence associated with the diagnostic accuracy of history taking questions or patient self-report questions from patients with low back related leg pain, there is a lack of synthesized literature to identify the accuracy of the history taking to identify underlying conditions.

Currently, there is no systematic review that has evaluated the diagnostic accuracy of history taking (questions) for patients with low back related leg pain to identify underlying spinal conditions. Therefore, the purpose of this study is to examine the diagnostic accuracy of the history taking or questions to identify underlying spinal conditions that commonly cause low back-related leg pain.

## 2. Methods

### 2.1. Study design

A systematic review conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Guidelines for Reporting Systematic Reviews (Liberati et al., 2009). PRISMA is intended to improve the "transparency and scientific merit" of systematic reviews and meta-analyses using a 27-item checklist (Liberati et al., 2009). The PRISMA guidelines were followed and the 27-item checklist was completed in order to improve the reporting of this systematic review. Additionally, the *Cochrane Handbook for Systematic Reviews of Interventions* was used as a protocol for entire review process (O'Connor et al., 2008).

### 2.2. Search strategy

A systematic search of electronic databases (MEDLINE, PubMed, CINAHL, and SCOPUS) was initiated in August 5, 2013 to retrieve studies that had investigated the accuracy of subjective history questions (e.g. interview, questionnaire, and/or self-report) to diagnose low back-related leg pain. The search strategy included MeSH terms (Table 1) and keyword searches, as well as a combination of both for a sensitive and specific search strategy (Appendices 1—3). References of each full-text were searched in

**Table 1**
MeSH terms.

| |
|---|
| Medical history taking |
| Low back pain |
| Radiculopathy |
| Referred pain |
| Sciatica |
| Intermittent claudication |
| Sciatic neuropathy |
| Nociceptive pain |
| Peripheral neuropathy |
| Neuralgia |
| Piriformis muscle syndrome |
| Humans |

addition to an extensive hand search of this literature. Subsequent hand searches were completed weekly to exhaust the search of any available grey literature throughout the stated databases and terminated on November 17, 2013.

### 2.3. Eligibility criteria

In determining study eligibility we followed the recommendations of the *Cochrane Handbook for Systematic Reviews of Interventions* (O'Connor et al., 2008). A study was included when it met all of the following criteria: 1) include patients with low back pain (LBP) AND related lower extremity pain condition 2) include specific details of subjective examination or patient self-report items, 3) study designs appropriate for diagnostic accuracy including cohort, and prospective/longitudinal studies, epidemiological studies, 4) used of statistical reporting within the study, 5) written in the authors primary English Language for interpretation, 6) used an acceptable reference standard for diagnosis (i.e. MRI/surgery for lumbar spine stenosis, lumbar nerve root compression/radiculopathy, and disc herniation and clinical impression for nociceptive/central sensitization). Articles were excluded if they did not meet one or more of the inclusion criteria, if they solely assessed diagnostic value of imaging or blood work; examined patients with low back pain only (as defined from the gluteal fold up to the 12th rib); and/or did not use an acceptable reference standard to diagnosis specific condition.

### 2.4. Study selection and data collection

Articles retrieved from each database were initially independently reviewed by two reviewers at each phase, title reviewers (HS and AE), abstract reviewers (HS and SS), and full text reviewers (AE and SS). Two reviewers (AE and HS) independently reviewed the full text articles for quality standards and came to agreement with any discrepancies. Any disagreements were mediated by a third reviewer (KA) who was not involved in the specific search strategy. Relevant articles were included in the review based on the aforementioned criteria. Items included within the data collection included author, year, subject characteristics, index standard, reference standard, diagnostic test item. Data on sensitivity, specificity, likelihood ratios and diagnostic odds ratios were also extracted.

### 2.5. Statistical analysis and quality assessment

Cohen's kappa of agreement is a statistical analysis utilized to measure the inter-rater agreement for qualitative items including review of the titles, abstracts and full-text articles (Van Tulder et al., 1997). The kappa inter-rater agreement was performed independently between two raters who identified articles as „yes" or „no" for acceptance.

Results were compared using the kappa formula: Kappa = Pr($a$)-Pre($e$)/1 − Pre($e$), where Pr($a$) was the relative observed agreement among raters and Pre($e$) equaled the hypothetical probability of chance agreement (Stribos et al., 2006).

The revised Quality Assessment of Diagnostic Accuracy Studies tool (QUADAS-2) builds on the original assessment tool for systematic reviews of diagnostic accuracy studies by identifying four key domains (patient selections, index test(s), reference standard and flow and timing) (Whiting et al., 2011). All domains are assessed for risk of bias. For the signaling questions a "yes" indicated a low risk of bias, a "no" answer indicated potential for bias and "unclear" indicated that there was insufficient information was provided in order to make a proper judgment (Whiting et al., 2011).