

## ORIGINAL ARTICLE

# Diagnostic Value of Assessment Tools for Sciatica in Clinical Practice: A Systematic Review and Narrative Synthesis

Nikolaos Kontakiotis, PT, MSc<sup>1</sup>, Evangelia Bratsi, PT<sup>2</sup>, Ioanna Findrili, PT<sup>2</sup>, Evdokia Billis, PT, PhD, MSc<sup>3</sup>, George Gioftsos, PT, PhD, MSc<sup>1</sup>

<sup>1</sup>Department of Physiotherapy, Laboratory of Advanced Physiotherapy, Faculty of Health and Caring Sciences, University of West Attica, Athens, Greece

<sup>2</sup>Department of Physiotherapy, Faculty of Health and Caring Sciences, University of West Attica, Athens, Greece

<sup>3</sup>Department of Physiotherapy, School of Health Rehabilitation Sciences, University of Patras, Aigion, Greece

**KEY WORDS:** *assessment, tools, diagnosis, sciatica*

#### ABBREVIATION LIST:

PROMs: patient reported outcome measures  
 PBOs: performance-based measures  
 S-LANSS: self-completed leeds assessment of neuropathic symptoms and signs  
 PDQ: painDETECT questionnaire  
 S-DN4: self completed douleur neuropathique  
 SQST: Standardized Qualitative Sensory Testing  
 StEP: standardized evaluation of pain  
 SLR: straight leg raise  
 LBP: lumbar spine pain

#### Correspondence to:

Nikolaos Kontakiotis,  
 Laboratory of Advanced  
 Physiotherapy, Physical Therapy  
 Department, Faculty of Health  
 & Caring Sciences, University  
 of West Attica,  
 28 Agiou Spiridonos street, Egaleo,  
 122 43, Athens, Greece  
 E-mail: nkontakiotis@uniwa.gr;  
 kontak766@hotmail.com

Manuscript received August 2, 2021;

Accepted November 22, 2021

#### ABSTRACT

**BACKGROUND:** Sciatica is one of the most common reasons for seeking healthcare for musculoskeletal pain and can be a challenge to healthcare providers to diagnose and treat. In view of the variability of sciatica symptoms, a great range of patient reported outcome measures (PROMs) and performance-based measures (PBOs) have been developed for its assessment and management, with however, often poor or controversial results in their reliability and discriminative ability. Accurate diagnosis of sciatica is crucial to ensure appropriate intervention is given. However, to date there is no gold standard to diagnose sciatica. There has been no systematic review conducted to compare the diagnostic validity of assessment tools of sciatica.

**OBJECTIVE:** To evaluate the diagnostic value of tools (PROMs and PBOs) used to assess patients presenting with sciatica.

**METHODS:** This review informed and reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis-Protocols. PubMed, Science Direct, Cochrane Library, CINAHL, MEDLINE, EMBASE, key journals and grey literature searched rigorously to find diagnostic accuracy studies investigating patient with sciatica. Two independent reviewers conducted the search, extracted the data and assessed risk of bias for included studies using the Quality Assessment of Diagnostic Accuracy Studies 2 tool. The overall quality of included studies evaluated using Grading of Recommendations, Assessment, Development and Evaluation guidelines.

**RESULTS:** From 8347 studies, 11 studies were included. Nine studies out of the 11 were at risk of bias. Very low level evidence supports the use of dermatomal patterns and low level evidence supports the use of 7 tools (neurological examination, Bragard test, S-LANSS, ID Pain, PDQ, S-DN4, SQST) for diagnosing sciatica. Moderate level evidence supports a cluster of eight signs (age, duration of disease, paroxysmal pain, pain worse in leg than back, typical dermatomal distribution, worse on coughing/sneezing/straining, finger to floor distance and paresis), twenty items of patient history, self-reported items (pain below knee, which pain worst, numbness pins and needles), question «worsening of pain during sneezing coughing and straining» and Slump test for diagnosing sciatica. Also, moderate level evidence supports the use of the StEP tool for diagnosing lumbar radicular pain, demonstrating high sensitivity (92%) and

**Conflict of Interest:** none declared

**PROSPERO registration number:** CRD42020168467

specificity (97%) values. SLR showed moderate level evidence in one study and high level evidence in another study in diagnosing sciatica with sensitivity 63.46% and specificity 45.88%.

**CONCLUSIONS:** Overall low-moderate level evidence supports the diagnostic utility of the tools examined in this review in diagnosing sciatica. The weak evidence base is largely due to methodological flaws and indirectness regarding applicability of the included studies. The most promising diagnostic tools include a cluster of 8 patient history/clinical examination signs, the StEP tool and the SLR test. From these results it is easily understood that history taking has a major role as assessment tool of sciatica in clinical practice. Low risk of bias and high level of evidence diagnostic utility studies are needed, in order for stronger recommendations to be made.

---

## INTRODUCTION

---

A large proportion of patients with lumbar spine pain (LBP) in primary care, approximately 60%, also report lower leg pain<sup>1</sup>. This pain is commonly referred to as sciatica and is often described as pain radiating to the buttocks, thighs and below the knee, foot and/or toes. It may be accompanied by physical findings of nerve root entrapment, such as decreased sensitivity, alteration in reflex and/or muscle weakness in entrapped nerve root distribution<sup>2</sup>. Sciatica is a relatively common<sup>3</sup> and often a persistent<sup>4</sup> nuisance that leads to the use of health services<sup>5</sup> and prolonged sick leave<sup>6</sup>. The prevalence of sciatica reported in the literature varies considerably from 1.6% in the general population to 43% in specific working population<sup>3</sup>. Although prognosis is good for most patients, a significant proportion (up to 30%) still have pain for a year or more<sup>7</sup>.

Sciatica has a significant impact on patients, health care and social costs<sup>8</sup>. The annual impact on the UK economy is estimated as GBP 268 million and GBP 1.9 billion for indirect expenditure; based on a Dutch study showing that the social costs of sciatica accounts for 13% of the cost regarding the LBP<sup>9</sup>.

Various environmental and hereditary factors affecting the development of sciatica have been studied, including gender, body weight, number of births, age, genetic factors, occupation and environmental factor. A study of 2946 women and 2727 men showed that neither gender nor body mass influenced the development of sciatica, although body mass may be related to LBP<sup>10</sup>. Height can be a risk factor for sciatica, although this seems to be important only in men aged 50-64<sup>10</sup>. The number of births of up to six children has also been recognized to be unrelated to sciatica<sup>11</sup>. The frequency of sciatica is related to age. It is rarely seen before the age of 20 and is more common in the fifth decade<sup>1</sup>. The odds ratio (OR) for an episode of sciatica is increased by 1.4 for every additional 10 years until age 64<sup>1</sup>. Relationships between psychosocial factors and sciatica have been poorly studied and results are conflicting<sup>12</sup>. However, the association of sciatica with psychological stress has been reported<sup>4</sup>.

Differentiating between the various causes of sciatica is

important for diagnosis and determination of the underlying pathology. The primary pathology that causes radiation pain is often difficult to differentiate as many structures are capable of causing similar pain symptoms<sup>13</sup>. Diagnosis is therefore essential to effective clinical practice as it determines the pathology of the problem and the goals of treatment. There are many clinical assessment tools and tests for patients with sciatica, used in clinical practice. A variety of patient reported outcome measures (PROMs) or performance-based measures (PBOs) have been developed for the assessment of sciatica. Many of these have not been fully evaluated and their diagnostic value has not been substantiated. There is a need to investigate the diagnostic value of these evaluation tools<sup>14</sup>.

## OBJECTIVE

To evaluate the diagnostic value of tools (PROMs and PBOs) used to assess patients presenting with sciatica.

---

## METHODS

---

This systematic review follows the Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines<sup>15</sup>. Methods and inclusion criteria were specified in advance and documented in a protocol published in PROSPERO (CRD42020168467). This systematic review and narrative synthesis conducted based in the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy<sup>16</sup>.

## ELIGIBILITY CRITERIA

Eligibility criteria were informed using the Sample, Phenomenon of Interest, Design, Evaluation, Research Type (SPIDER) guidelines<sup>17</sup>.

## INCLUSION CRITERIA

- Cross sectional studies, observational studies and clinical trials in English language published in the last 20 years (2000-2020).
- Humans with diagnosis of back pain with referred leg pain (sciatica) aged 18-75 years old.
- Studies that evaluate the diagnostic ability of sciatica assessment tools without intervention.

- Patient Reported Outcome Measures (PROMs) or Performance Based Measures (PBOs).

#### EXCLUSION CRITERIA

- Single case studies, free conference announcements, commentary articles, abstracts, letters.
- Studies with patients generally referred to back pain without defining the presence of reported pain in the lower extremities. Patients with serious pathology and red flags (cauda equina syndrome, inflammatory arthritis, malignancy etc.), severe trauma, psychiatric illness or personality disorder, pregnant women and animals.
- Studies evaluating the effectiveness of diagnostic mechanisms for sciatica. Studies investigating economic parameters and epidemiological evidence of sciatica.

#### INFORMATION SOURCES

A comprehensive search was conducted without the restriction in the period of the twenty last years (2000 - 2020) to identify relevant studies in various electronic databases including, Science Direct, Pubmed, Cochrane library, CINAHL, MEDLINE, EMBASE. To extend the scope of the review, was conducted complementary hand searching of field and topic relevant journals including reference lists of potentially relevant articles.

#### SEARCH STRATEGY

Natural language expressions and medical subject heading (MeSH) terms were used for the search in the electronic databases to be sure that any term related to sciatica have been missed.

#### STUDY SELECTION

Due to the large number of studies identified from the preliminary search the study selection was conducted in 3 stages. At the first stage 20 years were set as an inclusion criterion (2000-2020) with the cutoff date being 31st of December 2020. Then eligibility criteria were applied to the title and abstract of the search results by one reviewer. During the third stage, the same eligibility criteria were applied to the full-text articles of the remaining studies by two independent reviewers. The agreement between the two reviewers regarding study selection was analyzed using the Kappa-Cohen statistic<sup>18</sup>. Disagreements between the two reviewers were resolved by consensus. A third reviewer was used if disagreement could not be resolved in discussion.

#### DATA COLLECTION PROCESS AND DATA ITEMS

Based on the guidelines in chapter 7 of the Cochrane Handbook<sup>16</sup> a data extraction sheet was developed for data collection and process. There was a pilot tested on three randomly selected included studies from the two reviewers. One reviewer extracted the data from included studies and

the second reviewer checked for correctness. Information was extracted from each included study on: (1) authors (2) type of study (3) characteristics of trial participants (sample size, age, sex) (4) region of study (5) reference standard and (6) index test. In studies with multiple comparisons, multiple measurement points or multiple outcome measures, only those relevant to the aim of the systematic review were extracted. All characteristics were extracted using a standardized form and are shown in Table 1.

#### RISK OF BIAS IN INDIVIDUAL STUDIES

The risk of bias and concerns regarding applicability of included articles was assessed by two independent reviewers (IF and EB) using the Quality Assessment for Diagnostic Accuracy Studies (QUADAS-2) tool<sup>19</sup>. The four domains of the QUADAS-2 tool (patient selection, index test, reference standard and flow and timing) were independently assessed and judged by each reviewer (IF, EB) as 'high', 'low' or 'unclear risk'. The reviewers' agreement on risk of bias (RoB) items before discussion was calculated using the Cohen's correlation coefficient. In the case of disagreement that could not be resolved in a discussion, a third independent reviewer (NK) was used.

#### DATA SYNTHESIS

All studies identified as eligible at full text level were included in the qualitative synthesis. Due to heterogeneity across studies regarding populations, studies using different reference standards, different interventions and especially types of outcome measures used, it was not possible to pool data in a meta-analysis. Also, no meta-analysis could be performed because no test results with both sensitivity and specificity values were reported in more than one study.

#### CONFIDENCE IN CUMULATIVE EVIDENCE

Evidence quality was graded with the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach for diagnostic tests<sup>20</sup>. It involves consideration of within-study risk of bias (methodological quality), directness of evidence, heterogeneity and precision of effect estimates. Evidence profile was created that indicated the quality of the evidence assessment according to GRADE approach in each diagnostic tool.

---

## RESULTS

---

#### STUDY SELECTION

The selection of studies is outlined in the PRISMA flow diagram (fig. 1). A total of 8347 articles were identified during the electronic and hand-searching processes, of which 362 were duplicates. After putting as an exclusion criterion the date (2000-2020) and language (english) 7565 articles

ASSESSMENT TOOLS FOR SCIATICA IN CLINICAL PRACTICE

TABLE 1. Characteristics of included studies

Authors/Year	Type of study	Characteristics	Region	Reference Standard	Index Test
Gudala et al (2017)	Cross-sectional observational study	n = 215 (Female: n = 104, Male: n = 111) Mean (SD) age: 46.6 (13.9)	India	Physician opinion	S-DN4, ID Pain, pain DETECT questionnaire, S-LANSS
Verwoerd et al (2014)	Cross-sectional Observational study	n=395 (Female: n = 147, Male: n = 248) Mean (SD) age: 42.8 (10)	Netherlands	MRI	History taking
Konstantinou et al (2012)	Randomized Controlled Trial (RCT)	n=511 Patients with LBP, with or without LP, age ≥18 years old	Primary Care Trusts in Staffordshire, England	Clinical examination	Self-reported items
Lin et al (2017)	Cross-sectional observational study	n = 60 (Female: n = 38, Male: n = 22) Mean (SD) age: 61.37 (Nil reported)	Taiwan	MRI: grade 3 lateral stenosis	SQST
Hancock et al (2011)	Randomized Controlled Trial (RCT)	n=283 Patients with LB radicular syndrome, age 18-65 years old	9 hospitals of Netherlands	MRI	Neurologic examination
Verwoerd et al (2016)	Cross-sectional study	n=395 Patients with severe low back pain, age 18-65 years old	9 clinics of West Netherlands	MRI	«worsening of pain during coughing, sneezing and straining»
Taylor et al (2013)	Observational study	n=181 Patients with nerve route compression	Spinal Assessment Clinic	MRI and surgery	Patient report of dermatomal patterns
Homayouni et al (2018)	Cross-sectional study	n=506 Patients with unilateral radiating LBP, age 20-80 years old	Iran	Electrodiagnosis	Modified Bragard Test, Straight Leg Raise
Walsh J & Hall T (2009)	Observational Cross-sectional study	n=45 Age 18-70 years old	Clinic in Dublin	Diagnostic Imaging	Straight Leg Raise & Slump Test
Vroomen et al (2002)	Cross-sectional observational study	n=274 (Female: n = 139, Male: n = 135) Mean (SD) age: 46	Netherlands	MRI	History and physical examination
Scholz et al (2009)	Cross-sectional observational study	n = 138 (Female: n = 78, Male: n = 60) Mean (SD) age: 45	Usa	Independent physician clinical diagnosis	StEP tool

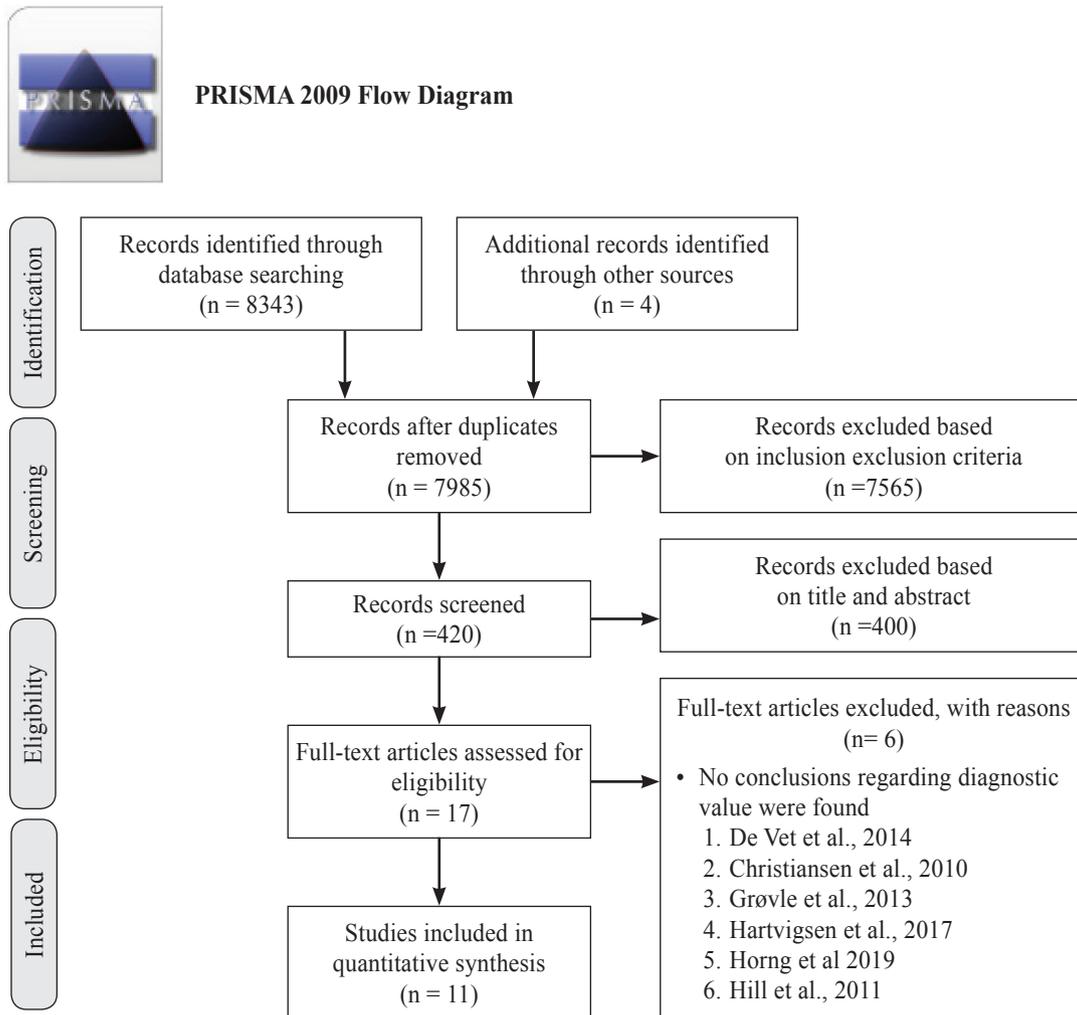


FIGURE 1. Flow diagram of selected studies

were excluded. 420 articles were eligible for title and abstract screening. After title and abstract screening, the full texts of 17 articles were considered for full text review. Of these, 6 were rejected. The first<sup>21</sup> was rejected because a specific scale was used to refer to hallucinations, weakness and lower extremity pain (Sciatica Bothersomeness Index), without assessing its diagnostic value. The second study<sup>22</sup> was rejected for a similar reason. It did not assign the diagnostic value of the questionnaire, but the Minimal Important Change (MIC) for two questionnaires. The third study<sup>23</sup> was rejected because it focuses on McKenzie's treatment rather than its diagnostic value. Three more studies<sup>1,24,25</sup> were rejected because they did not evaluate sciatica. Finally, a total of 11 studies were included for the review.

**STUDY CHARACTERISTICS**

Table 1 shows the study characteristics of all included stud-

ies. Each index test was used in only one study and compared with others tests. Only straight leg raise (SLR) was used in two studies but was compared with different assessment test. Analytically the 11 studies that were selected to extract the results in this review were: Verwoerd et al.<sup>26</sup>, Konstantinou et al.<sup>27</sup>, Hancock et al.<sup>28</sup>, Gudala et al.<sup>29</sup>, Lin et al.<sup>30</sup>, Scholz et al.<sup>31</sup>, Verwoerd et al.<sup>32</sup>, Vroomen et al.<sup>33</sup>, Taylor et al.<sup>34</sup>, Homayouni et al.<sup>35</sup>, Walsh and Hall<sup>36</sup>.

**RISK OF BIAS WITHIN STUDIES**

Judgements of risks of bias and concerns regarding applicability were made using the QUADAS-2 tool, as shown in Table 2. Complete agreement was achieved between the two reviewers for assessment of RoB, and thus the third reviewer was not required. Two studies were assessed as low RoB<sup>31,33</sup> and six studies showed high RoB in at least two domains<sup>26,28,29,32,34,35</sup>. Eight studies showed low concerns regarding applicabil-

TABLE 2. QUADAS 2

Study	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Vroomen et al (2002)	☺	☺	☺	☺	☺	☺	?
Gudala et al (2017)	☺	?	?	☺	☹	☺	?
Verwoerd et al (2014)	?	?	☺	☺	☹	☺	?
Konstantinou et al (2012)	☺	?	☺	☺	☺	☺	?
Scholz et al (2009)	☺	☺	☺	☺	☺	☺	?
Lin et al (2017)	☺	☺	☺	☺	☹	☺	?
Hancock et al (2011)	☺	?	☺	?	☺	☺	?
Verwoerd et al(2016)	☺	?	☺	?	☺	☺	?
Taylor et al (2013)	☺	☹	☹	?	☺	☺	?
Homayouni et al (2018)	☺	☹	?	?	☺	☺	?
Walsh & Hall (2009)	☺	☺	☹	☺	☺	☺	?

☺ Low Risk    ☹ High Risk    ? Unclear Risk

ity<sup>26-28,31,33-36</sup> and all studies were no clear in the applicability as far as concerns the reference standard. The primary concerns in relation to the at RoB studies were blinding of index and reference tests/insufficient description of procedures involved in index and reference test. This is because in the absence of a gold standard or clear recommendations/ guidelines to diagnose sciatica it is unclear whether the reference standards used in the studies correctly classify the target condition.

## SYNTHESIS OF RESULTS

### Physical examination

In the study of Vroomen et al.<sup>33</sup> identified 8 signs (including patient history and clinical examination signs) which were predictive of lumbosacral nerve root compression demonstrating moderate sensitivity (72%) and moderate/high specificity (80%). This study described one more item the pain referred in a dermatomal distribution. Vroomen et al.<sup>33</sup> used MRI as a reference standard, which has been questioned for its diagnostic validity<sup>37</sup>, furthermore this study was investigating lumbosacral nerve root compression which does not necessarily infer neuropathic pain. Using GRADE, there is a moderate level of evidence to support Vroomen et al's<sup>33</sup> eight signs in diagnosing lumbosacral nerve root compression.

### Patient history

The study of Verwoerd et al.<sup>32</sup>, investigated patient history data in relation to diagnosing nerve root compression or herniated disc in patients with LBLP. This study investigated 20 separate patient history items. Of the 20 items, moderate/high and high sensitivity values in both herniated disc and nerve compression groups were observed for health-related absenteeism (81 and 80% respectively) and in subjective sensory loss (89 and 90% respectively). Having had pain in the same leg previously demonstrated the highest specificity, in both herniated disc and nerve compression groups (90 and 91% respectively). Indirectness of evidence was highlighted as a highly selective population of patients were recruited. Using GRADE, there is low quality of evidence to support the use of Verwoerd et al's<sup>32</sup> patient history indicators in diagnosing nerve root compression or herniated discs.

### Questionnaire Self-reported items

This questionnaire consists of self-reported items, which was answered by patients with low back pain with pain radiating to the leg(s) in the study of Konstantinou et al.<sup>27</sup>. The items "pain below the knee", "which pain worst", "numbness, pins and needles" had 0.6, as designated by the AUC value. When

they made a cluster of these three items, they observed that they had an improvement in an acceptable level, as the AUC value was 0.72-0.74. Sensitivity was above 50% for the “pain below knee”, “numbness, pins and needles”, “which pain worst” and two more items. Specificity was over 50% for all items. Only these three items were significantly independent, at the level of  $P < 0.05$ . They also found that the NPV was higher than the PPV, 0.76 and 0.59 respectively. So, it was more possible to predict the absence of the nerve root involvement than the presence. Using GRADE, there is a moderate level of evidence to support the use of these self-reported items in diagnosing sciatica.

#### *Question «worsening of pain during coughing, sneezing and straining»*

Worsening of pain during coughing, sneezing and straining is a single question, which was used at this study of Verwoerd et al.<sup>26</sup>, to identify the ability to estimate if the localization of pain matters at a lumbosacral nerve root compression or a disc herniation. The patients had to answer this question, choosing one of the four options: a) no worsening of pain, b) worsening of back pain, c) worsening of leg pain, d) worsening of back and leg pain. Diagnostic odd ratios (DORs), sensitivity, specificity, and 95% CI were calculated. In both nerve root compression and herniated disc, the sensitivity of worsening of leg and back pain was 0.71 (0.66-0.76) and specificity was 0.31-0.32 (0.22-0.43). The odd ratios were 1.15 and 1.10 respectively. Similarly, the sensitivity of worsening of leg pain was 0.56 (0.50-0.61) and 0.54 (0.49-0.60) and the specificity 0.61 (0.49-0.72) and 0.59 (0.45-0.71). The odd ratios were 1.94 (1.17-3.21),  $p=0.01$  and 1.67 (0.95-2.94). Finally, sensitivity of sole worsening of leg pain was 0.40 (0.35-0.46) and 0.39 (0.34-0.45) and specificity was 0.77 (0.66-0.86) and 0.79 (0.67-0.89). The odd ratios were 2.28 (1.28-4.04),  $p < 0.01$  and 2.50 (1.27-4.90),  $p < 0.01$ , which was the highest DOR. Using GRADE, there is a moderate level of evidence to support the use of these questions in diagnosing sciatica.

#### *Dermatomal Pattern*

Dermatomal pattern charts show which nerve innervate each part of the body. In the study of Taylor et al.<sup>34</sup>, participated 181 patients which had surgically proven S1 (83) or L5 (98) nerve root compression. They all completed a computer-based questionnaire that allow them to indicate areas where experience pain and/or pins and needle. The results showed a substantial overlap of the dermatomes in the most patients and the distribution of pain and pins and needles did not correspond well with the dermatomal pattern. Using GRADE, there is a very low level of evidence to support the use of dermatomes in assessing sciatica.

#### *Standardized Evaluation of Pain (StEP tool)*

The study of Scholz et al.<sup>31</sup> was deemed at RoB as patient

applicability was compromised, this was partly due to the recruitment of patients with LBP with or without leg pain which is not consistent with the target population for this review. Furthermore, the reference standard, clinical judgement, was not adequately described and thus subject to bias. Additionally, this study was completed in a different language and cross-cultural validation cut of points used are yet to be validated. The StEP tool was found to have a high sensitivity (92%) and specificity (97%) when diagnosing lumbar radicular pain, this evidence was found to be of low RoB. Using GRADE, there is a moderate level of evidence to support the use StEP tool for diagnosing lumbar radicular pain.

#### *Self-completed douleur neuropathique (S-DN4), ID Pain, PainDETECT questionnaire (PDQ), S-LANSS*

The study of Gudala et al.<sup>29</sup> investigated four screening tools; S-DN4 (Self completed douleur neuropathique, ID Pain, PDQ (painDETECT questionnaire) and S-LANSS (Self-completed Leeds Assessment of Neuropathic symptoms and Signs) to identify NP in LBP. Three of the screening tools were identified as having a range of low/moderate to high sensitivity and specificity values; In the above study found that SDN4 had sensitivity and specificity values 58.5% & 98%. ID pain had sensitivity and specificity values 70.7% & 84.3% and PDQ had sensitivity and specificity values 76.8% & 78.4%. However, in this study, the S-LANSS was identified as having a low sensitivity of 13% but had excellent specificity value of 100%. Using GRADE, there is a low level of evidence to support the use of all of these tools for diagnosing sciatica.

#### *Straight Leg Raise (SLR)*

The SLR is well known clinical test used in the examination of low back related leg pain. In this observational cross-sectional study of Walsh and Hall<sup>36</sup>, SLR and SLUMP test were compared to prove agreement and correlation between them. Positive and negative finding of the SLR were cross tabulated by those of Slump and  $\kappa$  was used to determine agreement between them. The  $\kappa$  score between the results of the 2 examiners was 0.8 and the  $\kappa$  score of the agreement for both tests was 0.71. A strong correlation of the ROM of Slump and SLR was found on the symptomatic side with  $r=0.64$   $p<0.01$  but a weak one of the asymptomatic  $r=0.3$   $p=0.05$ . These results indicate there is reliability between the 2 raters and the procedure and there is a substantial agreement between the 2 tests implying that both tests are measuring the same thing. In the study of Homayouni et al.<sup>35</sup>, 506 consecutive patients with a history and physical examination suggesting unilateral L5-S1 radiculopathy took part. Electrodiagnosis (EDx) was the reference standard for this study. Sensitivity, specificity, positive and negative LR, positive and negative predictive values were calculated. The sensitivity for the SLR test was 63.46% (57.85-68.81) and the specificity was 45.88% (38.72-53.16). Using GRADE, there is a moderate level of evidence

to support the use of Walsh and Hall's<sup>36</sup> straight leg raise and a high level of evidence to support the use of Homayouni et al's<sup>35</sup> straight leg raise in diagnosing sciatica.

### *Slump*

The Slump Test is a clinical test used in the examination of low back related leg pain. In this observational cross-sectional study of Walsh and Hall<sup>36</sup> the Slump test was compared to SLR to prove agreement and correlation between the 2 tests. They compared the positive and negative findings of the 2 tests as well as the ROM as it was described previously. The  $\kappa$  score of the agreement for both tests was 0.71 and a strong correlation of the ROM of Slump and SLR was found on the symptomatic side with  $r=0,64$   $p<0.01$  but a weak one of the asymptomatic  $r=0.3$   $p=0.05$ . Using GRADE, there is a moderate level of evidence to support the use of Slump test in diagnosing sciatica.

### *Modified Bragard Test*

Modified Bragard test, was examined in the study of Homayouni et al.<sup>35</sup> Firstly, the SLR was performed in every patient and it was considered as positive if the patient complained of reproduction of symptoms distal to the knee, between 30 - 90 degrees of hip flexion. Secondly, the Modified Bragard test was performed in this way: The examiner began with an SLR, and if no radicular pain or symptoms were reproduced despite 70 degrees hip flexion, the foot was dorsiflexed firmly, and if radiating pain below the knee was produced the test was positive. Electrodiagnosis (EDx) was the reference standard for this study. Sensitivity, specificity, positive and negative LRs, positive and negative predictive values were calculated. The sensitivity for the Modified Bragard test was 69.30% (59.97-77.60) and the specificity was 67.42% (56.66-76.98). The DOR for the Modified Bragard test was 4.63. Using GRADE, there is a low level of evidence to support the use of this test in diagnosing sciatica.

### *Individual Neurologic Examination Tests*

Neurologic examinations tests included pain dermatome/location, reflex tests, sensory loss testing, motor strength and neurologic suspected level were examined in the study of Hancock et al.<sup>28</sup>. The area under the curve (AUC), sensitivity and specificity were used to investigate the diagnostic accuracy of each index test. If any of the index tests produced an AUC under 0.55, then it was excluded. The level of disc herniation from neurologist's examination was more accurate than the index tests. The AUC was 0.79 at L4/L5 and 0.80 at L5/S1. For the L3/L4 level, L4/L5 and L5/S1, respectively, the specificity was 90%, 83% and 94%, but the sensitivity was poor. Using GRADE, there is a low level of evidence to support the use of neurologic examination in assessing sciatica.

### *Standardized Qualitative Sensory Testing (SQST)*

In the study of Lin et al.<sup>30</sup> SQST was found to have low/

moderate sensitivity (62%) and high specificity (95%) when detecting lumbar lateral stenosis of the L5 nerve root. However, indirectness of evidence was highlighted as the participants recruited into this study were all surgical patients and therefore not fully representative of the target population for this review. Using GRADE low level of evidence supports the use of SQST in diagnosing lumbar lateral stenosis of the L5 nerve root.

---

## DISCUSSION

---

The main purpose of this review was to identify which tools have the diagnostic value for assessing patients with sciatica. Many techniques have been explored to treat patients with sciatica, but none of them have been statistically or clinically significant<sup>38,39</sup>. It is thus easily understood that research interest is increasingly focused on exploring more effective evaluation tools to categorize these patients with the aim of selecting the most appropriate technique for their treatment. To this end, reference was made to questionnaires and clinical tools that through their investigations proved their diagnostic value, either individually or through various combinations of tools. Evaluating and categorizing patients through simple procedures is a very useful tool as it can be a guiding element in finding the most appropriate care and ultimately treating patients. The methodological evaluation of the 11 studies that were the sample of this systematic review was made using the criteria presented in the "Cochrane Handbook for Reviews of Diagnostic Test Accuracy"<sup>16</sup>.

One of the most common assessment tools for sciatica is SLR. In this review SLR was investigated in two separate studies. In the study of Walsh and Hall<sup>36</sup> the SLR has an increased sensitivity but low specificity. This study suggests that low diagnostic accuracy may be due to the fact that the SLR may be positive for causes other than the direct mechanical effects of the intervertebral disc, such as in the presence of nerve root inflammation. Therefore, it is recommended that the SLR with dorsal flexion in the ankle, may an indicator of nerve root motility, which requires further research. In the study by Homayouni et al.<sup>35</sup>, it is argued that SLR alone cannot provide more accurate guidance for the clinical picture of patients, which is consistent with the research of Walsh and Hall<sup>36</sup>. It is argued that the use of the modified Bragard sigh in combination with the SLR - as has been demonstrated - reinforces its distinction. It also offers results even in the acute phase. SLR, Bragard, and the Slump test were part of our review, with positive results, either independently<sup>36</sup>, or after combination<sup>35</sup>.

SQST demonstrated low/moderate sensitivity and high specificity when diagnosing lumbar lateral stenosis involving the L5 nerve root in the study of Lin et al.<sup>30</sup>. The population of patients used were all surgical and therefore not fully rep-

representative of the target population for this review, thus the applicability of these findings is poor. There is evidence to support the use of quantitative sensory testing (QST) in diagnosing small fibre nerve degeneration in entrapment neuropathies<sup>40</sup>. However, SQST differs to QST as it describes tests which are inexpensive and accessible within a clinical setting (e.g. coin for testing temperature). Evidence to support SQST to detect small fibre nerve degeneration is limited<sup>41</sup> and yet to be investigated in participants with LBLP. The sensory profiles of those with NP in LBLP is not known and therefore support for SQST in identifying NP in LBLP is inconclusive. Screening tools A range of low/moderate to high sensitivity and specificity values were found for S-DN4, ID Pain and PDQ in a study investigating CLBP with or without leg pain<sup>29</sup>. Scholz et al.<sup>31</sup> found high sensitivity and specificity in their study investigating the use of the StEP tool in identifying lumbar radicular pain, this study was at low RoB. Moderate level of evidence supports the diagnostic utility of the StEP tool in diagnosing lumbar radicular pain. However clinical judgement was used as a reference standard which was not adequately described, furthermore clinical judgement is not a validated means to identify NP in LBLP. There is no further research to support the use of the StEP tool in identifying NP in LBLP, further research is needed to support its use.

In this review demonstrated that the diagnosis of sciatica, as shown by the study of Hancock et al.<sup>28</sup> and Taylor et al.<sup>34</sup>, cannot be based on the patient's description of the area of pain and in the distribution of pins and needles feeling as it seems that in most cases, they do not follow the expected dermatomal distribution. In addition, sensitivity, strength, and reflex tests, as shown in the study by Hancock et al.<sup>28</sup> are not of great diagnostic value both individually and in combination. Therefore, for the diagnosis of patients with sciatica we cannot rely solely on neurological examination as shown above.

In the study of Konstantinou et al.<sup>27</sup> some self-reported questions were investigated for their diagnostic value. Pain below knee was the best single item for diagnostic accuracy with an area under curve (AUC) of 0.67–0.68, which however is slightly less than the 'acceptable discrimination'. A cluster of three items, including distribution of pain below the knee, leg pain that is worse than back pain, and feeling of numbness or pins and needles in the leg, did improve discrimination to an 'acceptable' level with an AUC of 0.72–0.74 in relation to confirmatory and indicative references, respectively. However, the likelihood ratios from the models were reflective of a 'small' amount of discrimination. In conclusion in this primary care population seeking treatment for LBP with or without leg pain, was found no clear set of self-report items that accurately identified patients with nerve root pain. When accurate case definition is important, clinical assessment should be the method of choice for identifying LBP with possible nerve root involvement. The results suggest that self-report is not an accurate method for identifying individuals with the condi-

tion and it may overestimate or underestimate its prevalence. Certain self-report indicators particularly pain radiating below, leg pain worse than back pain and numbness, pins and needles in the leg can be useful at a very crude level. However, when accuracy in case definition is of importance, clinical examination is the recommended method.

The results of this review have shown that the assessment of sciatica it is a very complex and challenging task in clinical practice. History taking has a major role and for this reason, a holistic approach is needed when assessing these patients. Although this review did not find the diagnostic utility of questionnaires used for assessing sciatica, there are some very commonly used questionnaires in clinical practice. Questionnaires are critical for the assessment of patients with sciatica by evaluating many aspects of these patients such as: general health, psychology, mental health, neuropathic pain, disability, depression, anxiety and kinesiophobia. As a result of the indirectness highlighted regarding applicability concerns as well as the highly heterogenous data, the studies have been largely assessed individually and the limited synthesis made between studies have been suggested with caution. Due to the general low level of evidence and indirectness of evidence we believe that further research is needed to address the title of this review.

#### STRENGTHS OF REVIEW

The review was strengthened by following a predetermined and registered protocol (PROSPERO), by following established and standard reporting criteria (PRISMA), and the use of independent and blinded reviewers for filtering, data extraction, and QUADAS 2 assessment. For QUADAS 2 assessment, previously established standard and transparent criteria have been used. Furthermore, the GRADE approach for rating the level of evidence was used and reasons for downgrading have been explained, which increases transparency around the decisions that have been made.

#### LIMITATIONS OF THE REVIEW

However, as mentioned only 11 studies were included in the review. Seventeen studies initially appeared to be relevant, but after Full-text articles screened, 6 studies were excluded because there were no conclusions regarding diagnostic value were found. The small number of studies on which our conclusions are based is clearly a limitation, but this also reflects the small amount of the literature that was available. As with any review missing studies or failure to be comprehensive is a potential weakness, and the potential difficulty of finding relevant articles meant that search terms had to be reasonably wide. Any research to assess the accuracy, quality, and reliability of the tool under investigation requires comparing it with another tool that already has these components. As already mentioned, the tools used in our study included magnetic resonance imaging, electrodiagnosis, clinical examination and in some cases

other reliable questionnaire or clinical trial tools. As there is no gold standard for diagnosing sciatica, there is a limitation to compare diagnostic models for sciatica. Different choice of reference standard each time limits the accuracy of each tool.

---

### CONCLUSION

---

This systematic review revealed a number of results that may have a positive impact on the evaluation of sciatica as detailed above. From these results it is easily understood that history taking has a major role as assessment tool of sciatica in clinical practice. The most promising diagnostic tools include a cluster of 8 patient history/clinical examination signs, the StEP tool and the SLR test. Overall low-moderate level evidence supports the diagnostic utility of the tools examined in this review in diagnosing sciatica. The weak evidence base is largely due to methodological flaws and indirectness regarding applicability of the included studies. There are some studies that have not demonstrated the diagnostic value of the tools tested. For this reason, low risk of bias and high level of evidence diagnostic utility studies are needed, in order for stronger recommendations to be made.

### REFERENCES

- Hill JC, Konstantinou K, Egbewale BE, et al. Clinical outcomes among low back pain consultants with referred leg pain in primary care. *Spine* 2011;36:2168-2175.
- Koes BW, Van Tulder M, Peul WC. Diagnosis and treatment of sciatica. *BMJ* 2007;334:1313-1317.
- Konstantinou K, Dunn MK. Sciatica. Review of Epidemiological Studies and Prevalence Estimates. *Spine* 2008;33:2464-2472.
- Tubach F, Beute J, Leclerc A. Natural history and prognostic indicators of sciatica. *J Clin Epidemiol* 2004; 57:174-179.
- Molano SM, Burdorf A, Elders LA. Factors associated with medical care-seeking due to low-back pain in scaffolders. *Am J Ind Med* 2001;40:275-281.
- Hagen KB, Thune O. Work incapacity from low back pain in the general population. *Spine* 1998;23:2091-2095.
- Vroomen PC, Krom MC, Slofstra PD, et al. Conservative treatment of sciatica: a systematic review. *J Spinal Disorder* 2000;13:463-469.
- Konstantinou K, Hider SL, Jordan JL, et al. The impact of low Back-related leg pain on outcomes as compared with low Back pain alone: a systematic review of the literature. *Clin J Pain* 2013;29:644-654.
- Van Tulder MW, Koes B, Bouter LM. A cost of illness study of back pain in the Netherlands. *Pain* 1995;62:233-240.
- Nicholson B, Verma S. Comorbidities in chronic neuropathic pain. *Pain Med* 2004;5(9-27).
- Heliövaara M, Makela M, Knekt P, et al. Determinants of sciatica and low back pain. *Spine* 1991;16:608-614.
- Miranda H, Viikari-Juntura E, Martikainen R, et al. Individual factors, occupational loading, and physical exercise as predictors of sciatic pain. *Spine (Phila Pa 1976)* 2002; 27:1102-1109.
- Adams MA, Bogduk N, Burton K, et al. Biology of spinal tissues. The biomechanics of backpain. *Edinburgh: Churchill Livingstone* 2002;(49-71).
- Vroomen P, De Krom M, Knottnerus J. Diagnostic value of history and physical examination in patients suspected of sciatica due to disc herniation: a systematic review. *J Neurol* 1999;246:899-906.
- Moher D, Liberati A, Tetzlaff J, et al; Group TP. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine* 2009;6:194-197.
- Higgins J, Green S. *Cochrane Handbook for systematic reviews of interventions*. The Cochrane Collaboration 2011.
- Cooke A, Smith D, Booth A, et al. Beyond PICO: the spider tool for qualitative evidence synthesis. *Qual Health Res* 2012;22:1435-1443.
- Cohen J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 1960;20:37-46.
- Leeflang MMG, Deeks JJ, Takwoingi Y, et al. Cochrane diagnostic test accuracy reviews. *Syst Rev* 2013;2:82.
- Brozek JL, Akl EA, Jaeschke R, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines: Part 2 of 3. The grade approach to grading quality of evidence about diagnostic tests and strategies. *Allergy* 2009;64:1109-1116.
- Grøvle L, Haugen AJ, Natvig B, et al. The prognosis of self-reported paresthesia and weakness in disc-related sciatica. *Eur Spine J* 2013;22:2488-2495.
- De Vet HCV, Foumani M, Scholten MA, et al. Minimally important change values of a measurement instrument depend more on baseline values than on type of intervention. *Journal of Clinical epidemiology* 2014;68(518-524).
- Christiansen D, Larsen K, Jensen OK, et al. Pain response classification does not predict long-term outcome in patients with low back pain who are sick-listed. *J Orthop Sports Phys Ther* 2010;40:606-615.
- Hong YS, Hou WH, Liang HW. Responsiveness of the modified lower extremity functional scale in patients with low back pain and sciatica: A comparison with pain intensity and the modified Roland-Morris Disability Scale. *Medicine* 2019; 98:e15105.
- Hartvigsen L, Hestbaek L, Leboeuf-Yde C, et al. Leg pain location and neurological signs relate to outcomes in primary care patients with low back pain. *BMC Musculoskeletal Disorders* 2017; 989:e214
- Verwoerd AJH, Mens J, El Barzouhi A, et al. A diagnostic study in patients with sciatica establishing the importance of localization of worsening of pain during coughing, sneezing and straining to assess nerve root compression on MRI. *European Spine Journal* 2016; 25:1389-1392.
- Konstantinou K, Lewis M, Dunn, KM. Agreement of self-reported items and clinically assessed nerve root involvement (or sciatica) in a primary care setting. *European Spine Journal* 2012; 21:2306-2315.

28. Hancock MJ, Koes B, Ostelo R, et al. Diagnostic accuracy of the clinical examination in identifying the level of herniation in patients with sciatica. *Spine* 2011; 36:E712-9.
29. Gudala K, Ghai B, Bansal D. Usefulness of four commonly used neuropathic pain screening questionnaires in patients with chronic low back pain: a cross-sectional study. *Korean J Pain* 2017;30:51.
30. Lin JH, Hsieh YC, Chen YC, et al. Diagnostic accuracy of standardised qualitative sensory test in the detection of lumbar lateral stenosis involving the L5 nerve root. *Sci Rep* 2017;7:10598.
31. Scholz J, Mannion RJ, Hord DE, et al. A Novel Tool for the Assessment of Pain: Validation in Low Back Pain. *PLoS Med* 2009;6:e1000047.
32. Verwoerd AJ, Peul WC, Willemsen SP, et al. Diagnostic accuracy of history taking to assess lumbosacral nerve root compression. *Spine J* 2014;14:2028-2037.
33. Vroomen P, de Krom MCTFM, Wilmink J, et al. Diagnostic value of history and physical examination in patients suspected of lumbosacral nerve root compression. *J Neurol Neurosurg Psychiatry* 2002; 72:630-634.
34. Taylor CS, Coxon, AJ, Watson PC, et al. Do L5 and S1 nerve root compressions produce radicular pain in a dermatomal pattern? *Spine* 2013;38:995-998.
35. Homayouni, K, Jafari SH, Yari H. Sensitivity and Specificity of Modified Bragard Test in Patients with Lumbosacral Radiculopathy Using Electrodiagnosis as a Reference Standard. *Journal of Chiropractic Medicine* 2018;17:36-43.
36. Walsh J, Hall T. Agreement and Correlation Between the Straight Leg Raise and Slump Tests in Subjects with Leg Pain. *Journal of Manipulative and Physiological Therapeutics* 2009; 32:184-192.
37. Kalichman L, Kim DH, Li L, et al. Computed tomography evaluated features of spinal degeneration: prevalence, inter-correlation, and association with self-reported low back pain. *Spine J* 2010;10:200-208.
38. Murphy RD, Hurwitz LE, McGovern EE. A nonsurgical approach to the management of patients with lumbar radiculopathy secondary to herniated disk: a prospective observational cohort study with follow-up. *Journal of Manipulative and Physiological Therapeutics* 2009;32:723-733.
39. Hahne JA, Jon J, McMeeken JM. Conservative Management of Lumbar Disc Herniation with Associated Radiculopathy. A Systematic Review. *Spine* 2010;35:488-504.
40. Schmid AB, Hailey L, Tampin B. Entrapment neuropathies: challenging common beliefs with novel evidence. *J Orthop Sports Phys Ther* 2018; 48:58-62.
41. Ridehalgh C, Sandy-Hindmarch OP, Schmid A. Validity of clinical small fibre sensory testing to detect Small-Nerve Fibre Degeneration. *J Orthop Sport Phys Ther* 2018;48:767-774.