ACRM
AMERICAN CONGRESS OF REHABILITATION MEDICINE

Archives of Physical Medicine and Rehabilitation

journal homepage: www.archives-pmr.org

Archives of Physical Medicine and Rehabilitation 2018;99:726-35

ORIGINAL RESEARCH

Accuracy of Clinical Tests in Detecting Disk Herniation and Nerve Root Compression in Subjects With Lumbar Radicular Symptoms

Harald Ekedahl, PT, PhD,[a] Bo Jönsson, MD, PhD,[a] Mårten Annertz, MD, PhD,[b] Richard B. Frobell, PT, PhD[a]

From the [a]Department of Orthopaedics, Clinical Sciences Lund, Lund University, Lund; and [b]Department of Radiology, Lund University Hospital, Lund, Sweden.

Abstract

Objectives: To investigate the accuracy of 3 commonly used neurodynamic tests (slump test, straight-leg raise [SLR] test, femoral neurodynamic test) and 2 clinical assessments to determine radiculopathy (radiculopathy I, 1 neurologic sign; radiculopathy II, 2 neurologic signs corresponding to 1 specific nerve root) in detecting magnetic resonance imaging (MRI) findings (extrusion, subarticular nerve root compression, and foraminal nerve root compression).

Design: Validity study.

Setting: Secondary care.

Participants: We included subjects (N=99; mean age, 58y; 54% women) referred for epidural steroid injection because of lumbar radicular symptoms who had positive clinical and MRI findings. Positive clinical findings included the slump test (n=67), SLR test (n=50), femoral neurodynamic test (n=7), radiculopathy I (n=70), and radiculopathy II (n=33). Positive MRI findings included extrusion (n=27), subarticular nerve compression (n=14), and foraminal nerve compression (n=25).

Interventions: Not applicable.

Main Outcome Measures: Accuracy of clinical tests in detecting MRI findings was evaluated using sensitivity, specificity, and receiver operating characteristics analysis with area under the curve (AUC).

Results: The slump test had the highest sensitivity in detecting extrusion (.78) and subarticular nerve compression (1.00), but the respective specificity was low (.36 and .38). Radiculopathy I was most sensitive in detecting foraminal nerve compression (.80) but with low specificity (.34). Only 1 assessment had a concurrent high sensitivity and specificity (ie, radiculopathy II) in detecting subarticular nerve compression (.71 and .73, respectively). The AUC for all tests in detecting extrusion, subarticular nerve compression, and foraminal nerve compression showed ranges of .48 to .60, .63 to .82, and .33 to .57, respectively.

Conclusions: In general, the investigated neurodynamic tests or assessments for radiculopathy lacked diagnostic accuracy. The slump test was the most sensitive test, while radiculopathy II was the most specific test. Most interestingly, no relationship was found between any neurodynamic test and foraminal nerve compression (foraminal stenosis) as visualized on MRI.

Archives of Physical Medicine and Rehabilitation 2018;99:726-35

The lifetime prevalence of low back pain was reported to be >70%,[1] while low back pain with accompanying radicular symptoms has a reported prevalence of 2% to 43%.[2] The large variation reflects on the lack of consensus in diagnosing lumbar radicular pain.[3]

To diagnose radicular pain, guidelines suggest an initial clinical examination to evaluate the likelihood of disk herniation (DH), nerve root compression, or both.[4,5] Guidelines also suggest a subsequent magnetic resonance imaging (MRI) study if the symptoms fail to improve and the clinical test findings are steadfastly positive.[4] The clinical test results and the MRI findings are the basis of the diagnosis; thus, the intercorrelations between the clinical tests and the MRI findings are of clinical importance.

Supported by Region Skåne, Sweden.
Disclosures: none.

0003-9993/18/$36 - see front matter © 2017 by the American Congress of Rehabilitation Medicine
https://doi.org/10.1016/j.apmr.2017.11.006

Table 1 Baseline characteristics and clinical test results for the entire sample and for subjects with MRI findings (disk extrusion, subarticular nerve compression, foraminal nerve compression)

Variable	All (N = 99)	Extrusion (n = 27)	High-Grade Subarticular Nerve Compression (n = 14)	High-Grade Foraminal Nerve Compression (n = 25)
Age (y)	58 (54–61)	56 (52–61)	53 (44–62)	61 (56–66)
Sex (men)	45 (45)	12 (44)	6 (43)	15 (60)
BMI	27 (26–28)	27 (25–28)	26 (24–28)	28 (26–30)
Smoker (yes)	14 (14)	1 (4)	0 (0)	2 (8)
Duration of leg pain (mo)	34 (29–39)	33 (23–43)	24 (11–35)	38 (26–49)
Lumbar surgery	17 (17)	8 (29)	2 (14)	6 (24)
Level of TESI L2-3/L4/L5/S1	6/12/68/13	1/6/14/6	1/0/9/4	3/3/19/0
Positive slump test*	67 (67)	21 (78)	14 (100)	12 (48)
Positive SLR test*	50 (50)	16 (59)	13 (93)	8 (32)
Positive femoral neurodynamic test [n(%)]†	18 (18)	7 (26)	5 (36)	3 (12)
Radiculopathy I‡	70 (70)	18 (67)	13 (93)	19 (76)
Radiculopathy II§	33 (33)	13 (48)	10 (71)	7 (28)
Sensory deficit No/L2-4/L5/S1‖	38/5/41/15	11/1/8/7	2/0/7/5	11/1/11/2
Impaired patellar tendon reflex‖	10 (10)	3 (11)	0 (0)	5 (20)
Extensor hallucis longus weakness‖	36 (36)	12 (44)	11 (77)	11 (44)
Impaired Achilles' tendon reflex‖	17 (17)	7 (26)	3 (21)	4 (16)
VAS leg	50 (46–53)	52 (46–59)	57 (49–65)	43 (37–49)
ODI	43 (39–49)	42 (37–47)	44 (39–49)	39 (34–45)

NOTE. Values are mean (95% confidence interval), n (%), or n.
Abbreviations: BMI, body mass index; ODI, Oswestry Disability Index; VAS, visual analog scale.
* Lumbosacral neurodynamic test.
† Neurodynamic test to test L2-4.
‡ Assessed with 1 neurologic sign at the level of the planned TESI.
§ Assessed with 2 neurologic signs corresponding to 1 specific nerve root at the level of the planned TESI.
‖ All neurologic tests assessed at all levels.

The clinical examination consists of neurodynamic tests and neurologic examination tests.[6] The MRI findings to indicate radicular pain are DH and nerve root compression (subarticular or foraminal).[7,8]

The most frequently used neurodynamic test, the straight-leg raise (SLR) test,[6] was found to have moderate accuracy in detecting MRI-verified DH.[6] The other neurodynamic tests to diagnose lumbar radicular pain, the slump test and the femoral neurodynamic test, are less frequently studied and were also found to have moderate accuracy in detecting DH.[9,10] When the neurodynamic tests were correlated to MRI-verified nerve root compression, the accuracy was found to be poor to moderate.[10-12]

The neurologic test findings (sensory deficit, reflex impairment, muscle weakness) contribute to the diagnosis of radiculopathy.[13] However, no guidance was presented to indicate how many positive neurologic tests are required to meet the diagnosis.[3] One positive neurologic test in isolation does not accurately detect an MRI-verified DH or nerve root compression.[6,11,14]

A combination of neurologic tests was compared to MRI-verified DH in only 2 reports,[14,15] and to nerve compression (subarticular or foraminal) in 1 report,[10] and a comparison of the accuracy of the neurodynamic tests and the neurologic tests has to our knowledge never been conducted.

We aimed, in subjects with radicular pain who were referred for epidural steroid injection, to investigate the accuracy of 3 neurodynamic tests (slump test, SLR test, femoral neurodynamic test) and 2 assessments to determine radiculopathy (I, 1 neurologic sign; II, 2 neurologic signs corresponding to 1 specific nerve root) in detecting MRI findings (disk extrusion, subarticular nerve root compression, and foraminal nerve root compression).

List of abbreviations:

AUC area under the curve
DH disk herniation
DOR diagnostic odds ratio
MRI magnetic resonance imaging
SLR straight-leg raise
TESI transforaminal epidural steroid injection

Methods

Study design

This study was a prospective cohort study that investigated the accuracy of clinical test findings in relationship to MRI findings (reference standard) using the same cohort as a previous study.[16]

Participants

During 2011 through 2012, 151 patients older than 18 years who were referred for transforaminal epidural steroid injection (TESI) from the orthopedic clinic at a single hospital in the south of Sweden because of lumbar radicular pain were consecutively invited to participate in this study. The decision to treat with TESI (level and side) was made by an experienced orthopedic surgeon when signs and symptoms were consistent with the MRI findings

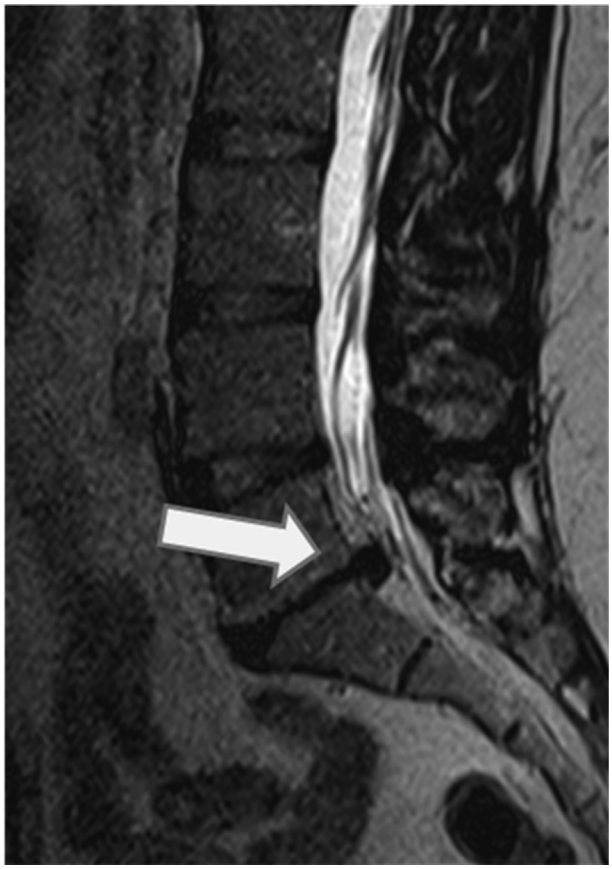

Fig 1 T2-weighted sagittal image revealing an extrusion at level L5-S1 (arrow).

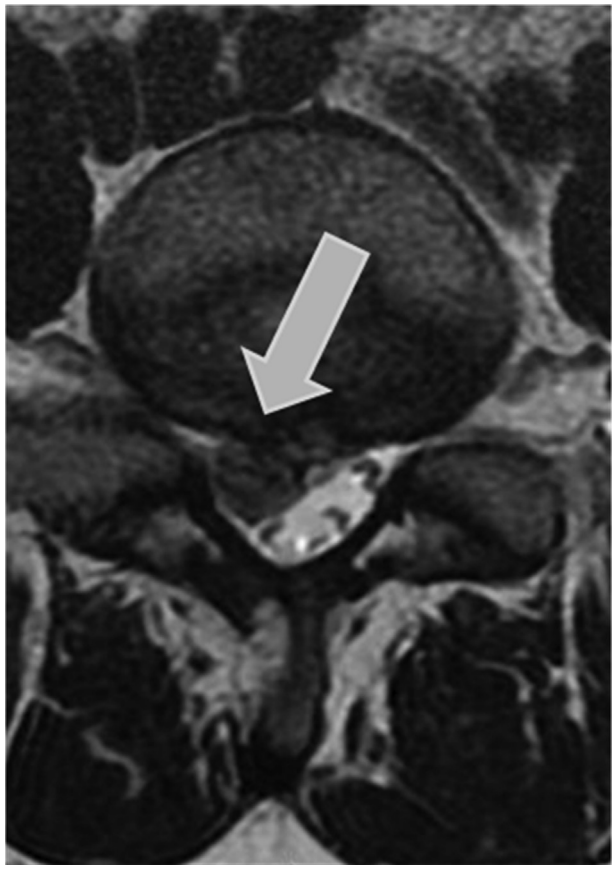

Fig 2 T2-weighted image revealing a right-sided extrusion at level L5-S1 and a high-grade subarticular nerve compression of right S1 nerve root (arrow).

of nerve root involvement secondary to protrusion/extrusion or foraminal nerve root compression.

Baseline characteristics, pain intensity, and disability of the study cohort are presented in table 1. More detailed information was previously described elsewhere.[16] In brief, 46 persons were excluded because of exclusion criteria (15 had TESI in the previous 12mo, 10 had bilateral radicular pain, 7 had lumbar fusion surgery, and 14 had other diseases), and 5 patients declined participation because they were unwilling to participate in a scientific trial, leaving 100 included patients with unilateral radicular pain.

The decision to treat with TESI, including determining the level and side of injection, was made in the normal clinical setting by an experienced orthopedic surgeon based on pain distribution (dermatome) and MRI findings. The MRI findings representing this specific nerve root and the level were the data included in the validity study.

All subjects provided signed informed consent before inclusion. At the start of this study (2011), in contrast to randomized controlled trials, cohort studies were rarely registered. Therefore, this study was not registered before the start of the study but was conducted in accordance to the study protocol and approved by the ethics committee, Lund University, Sweden (reference no. 2011/481).

Magnetic resonance imaging

All MRI acquisitions were made using a 1.5-tesla scanner (80% using a single Siemens Avanto scanner[a]). The images were obtained at the level of the nerve root engagement and no longer than 2 months before the planned TESI. The imaging protocol included T2-weighted turbo spin-echo sequences obtained in the sagittal and axial plane and a T1-weighted spin-echo sequence obtained in the sagittal plane. All sequences had a maximum of a 4-mm slice thickness.

Assessment of MRI

There was internal loss of MRI data in 1 subject because of poor-quality MRI images and consequently, 99 MRI images were analyzed. All MRI images were analyzed by a well-experienced radiologist (M.A.) using previously published classification systems.[7,17-19] The radiologist was blinded to all clinical information but the level of TESI. At this level, the following was analyzed: (1) whether DH existed; (2) the type of DH (protrusion or extrusion, according to Fardon et al[8]); and (3) the grade of nerve compression, according to Pfirrmann et al.[7]

Disk herniation

The subjects were classified as having (1) no DH (n=15, bulging disk was regarded as no DH), (2) protrusion (n=57), and (3) extrusion (n=27, fig 1).

Nerve compression

The grade of nerve compression, subarticular and foraminal, was assessed on axial T2-weighted images and sagittal T1-weighted

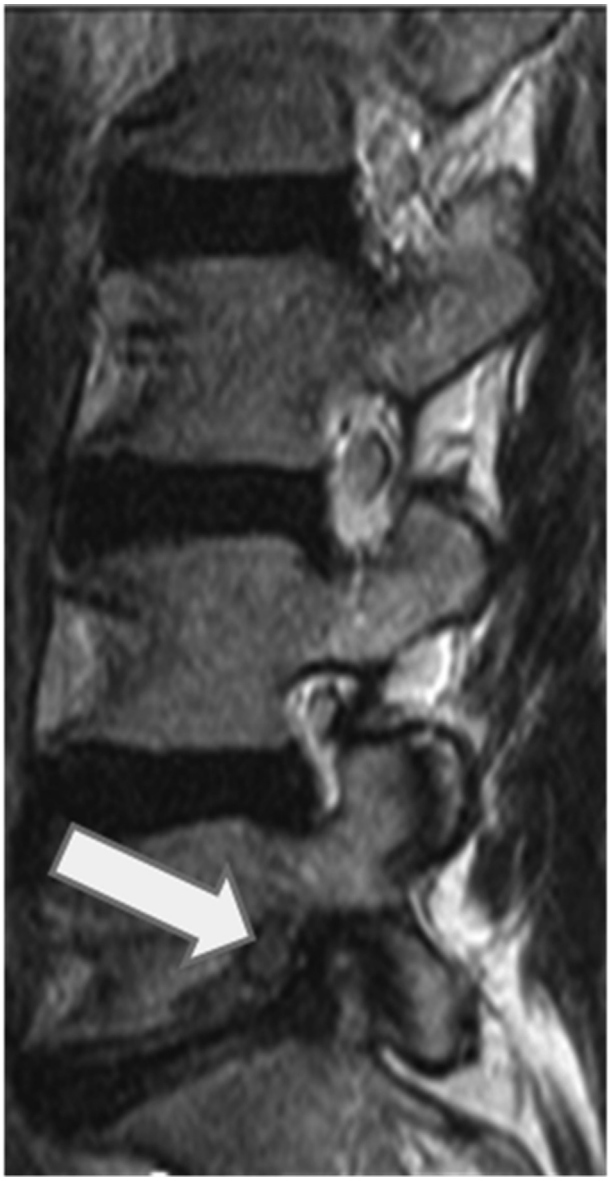

Fig 3 T1-weighted sagittal image revealing a high-grade foraminal nerve compression of L5 nerve root at level L5-S1 (arrow).

images, respectively. Subarticular nerve compression was assessed using the modification of a system described by Pfirrmann.[7,17] Grade I applies when the disk simply contacts the nerve root, grade II when the nerve root is displaced but with preservation of periradicular cerebrospinal fluid or fat, grade III when the periradicular cerebrospinal fluid or fat is obliterated, and grade IV when the nerve root is morphologically distorted. Grades I and II were considered as low-grade nerve compression, and grades III and IV were considered as high-grade nerve compression.[17] The subjects were classified as low-grade (n=85)/high-grade (n=14) subarticular nerve compression (fig 2).

The grade of foraminal nerve compression was assessed using a system introduced by Lee et al.[19] Grade I applies when perineural fat is obliterated in 2 opposing directions (vertical or transverse), grade II when perineural fat is obliterated in 4 directions without morphologic distortion of the nerve root, and grade III when distortion or other morphologic change

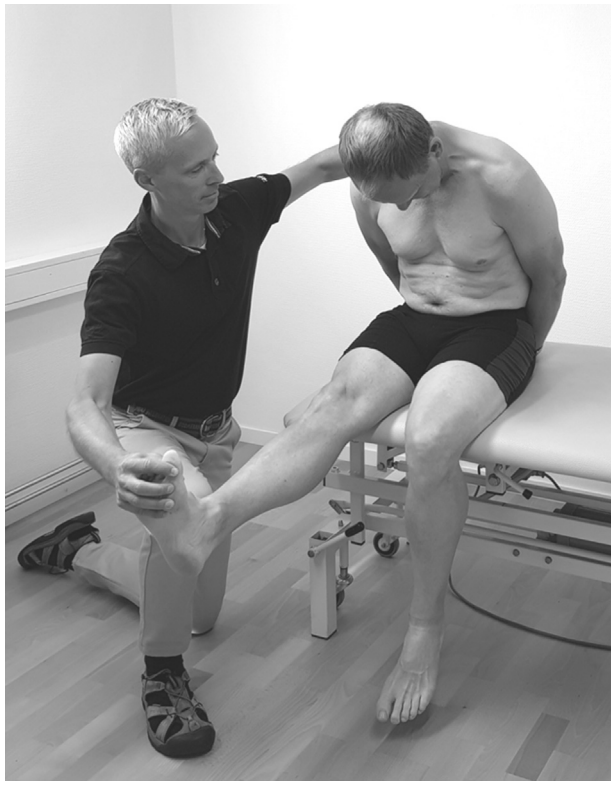

Fig 4 Slump test.

in the nerve root is evident. Grade I was considered as low-grade nerve compression, and grades II and III were considered as high-grade nerve compression.[19] The subjects were classified as low-grade (n=74)/high-grade (n=25) foraminal nerve compression (fig 3).

The sample was classified into 3 groups of nerve compression: low-grade nerve compression (neither high-grade subarticular nor high-grade foraminal nerve compression, n=61), high-grade subarticular nerve compression (n=14), and high-grade foraminal nerve compression (n=25). One subject was included in both of the latter groups.

Fig 5 SLR test.

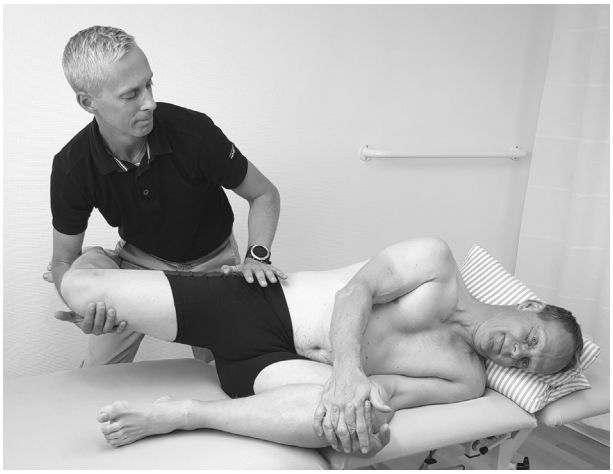

Fig 6 Femoral neurodynamic test (slump—knee bend test).

Clinical assessment

The clinical assessment was performed by the same experienced physiotherapist (H.E.) who was blinded to all MRI information. The neurologic examination was performed first, followed by the slump test, the SLR test, and the femoral neurodynamic test. After the clinical assessment, pain measures (visual analog scale leg pain) and demographic history were reported by the patient, and lastly the self-reported disability questionnaire (Oswestry Disability Index)[20] was filled out.

Clinical assessment, as well as collection of self-reported pain and self-reported disability scores, was performed preinjection in identical order for each subject and is described in a published report.[16]

Radiculopathy I and II was determined if the patellar reflex, Achilles' reflex, strength of the large toe in dorsiflexion, or sensibility (sensory deficit) in a specific dermatome area was asymmetrically deranged. Sensory deficit was assessed using light touch, pinprick, and cold stimuli. Radiculopathy I was considered when one of the neurologic signs above was present and corresponded to the nerve root of the planned TESI. Radiculopathy II was considered when 2 neurologic signs (sensory deficit + reflex impairment or muscle weakness) were present and corresponded to the specific nerve root of the planned TESI. No subject had a radiculopathy II at a different level from the planned TESI.

All neurologic tests were assessed at all levels and are presented in table 1. The motor function test for the L4 (musculus quadriceps) and S1 (musculus triceps surae) myotomes were discarded because of the risk of provoking excessive pain.

The slump test (fig 4) is a validated dichotomous test to assess the presence/absence of lumbosacral neural mechanosensitivity.[21] The test was performed with the patient sitting and was assessed through a combination of sitting thoracolumbar flexion, cervical flexion, ankle dorsiflexion, and knee extension. With the use of sensitizing maneuvers, beginning with the ankle and continuing with the neck, the test was considered positive if one of the maneuvers reproduced the symptoms and the symptoms were different from the contralateral side.[21]

The SLR test (fig 5) was performed with the patient supine according to the published instructions.[22] The straight leg was slowly raised and the test was classified as positive or negative, using sensitizing maneuvers, beginning with the ankle and continuing with the neck.[23]

The femoral neurodynamic test (slump knee bend, fig 6), a validated test performed with the patient side-lying on the nonaffected side, assesses the presence/absence of neural mechanosensitivity (L2-4) using a combination of thoracolumbar flexion, cervical flexion, knee flexion, and hip extension.[24] The test was classified as positive or negative, using sensitizing maneuvers, beginning with the knee and continuing with the neck.[24] The test was considered positive if one of the maneuvers reproduced the symptoms and the symptoms were different from the contralateral side.[16,24]

Table 2 A 2×2 contingency table comparing the clinical tests (slump test, SLR test, radiculopathy I, radiculopathy II) with MRI findings (disk extrusion, high-grade subarticular compression, high-grade foraminal compression)

	Extrusion		Subarticular Compression		Foraminal Compression		
Clinical Test	Yes	No	High Grade	Low Grade	High Grade	Low Grade	Total
Slump test							
Pos	21	46	14	53	12	55	67
Neg	6	26	0	32	13	19	32
SLR test							
Pos	16	34	13	37	8	42	50
Neg	11	38	1	48	17	32	49
Radiculopathy I							
Yes*	18	51	13	56	20	49	69
No	9	21	1	29	5	25	30
Radiculopathy II							
Yes†	13	20	10	23	7	26	33
No	14	52	4	62	18	48	66
Total	27	72	14	85	25	74	99

NOTE. Values are n.
Abbreviations: Neg, negative; Pos, positive.
* Assessed with 1 neurologic sign.
† Assessed with 2 neurologic signs corresponding to 1 specific nerve root.

Table 3 A 2×2 contingency table comparing the femoral neurodynamic test with MRI findings (disk extrusion, high-grade subarticular compression, high-grade foraminal compression) in subjects with midlumbar radicular pain (n=18)

	Extrusion		Subarticular Compression		Foraminal Compression		
Clinical Test	Yes	No	High Grade	Low Grade	High Grade	Low Grade	Total
Femoral neurodynamic test*							
Pos	3	4	1	6	1	6	7
Neg	4	7	0	11	5	6	11
Total	7	11	1	17	6	12	18

NOTE. Values are n.
Abbreviations: Neg, negative; Pos, positive.
* Neurodynamic test to test L2-4.

Interrater reliability evaluation was not performed, as all MRI assessments and clinical tests have shown adequate reliability previously.[17,19,21,24-26]

Statistical analysis

Statistical analyses were made using SPSS (version 23.0)[b] and R software.[c] A 2×2 contingency table was created to compare each clinical test (slump test, SLR test, femoral neurodynamic test, and neurologic examination) to MRI findings (disk extrusion, high-grade subarticular compression, and high-grade foraminal compression). The accuracy of the clinical tests in detecting the MRI findings was primarily evaluated using sensitivity, specificity, and receiver operating characteristics analysis with area under the curve (AUC), including 95% confidence intervals, and secondarily evaluated using positive and negative predictive values, positive and negative likelihood ratios, and diagnostic odds ratios (DORs).[27]

As references for the interpretation of sensitivity and specificity, the following classification was used: 0 to 0.4, low; 0.4 to 0.7, moderate; and 0.7 to 1.0, high.[28,29] For the respective interpretation of AUC, the following classification was used: 0 to 0.7, poor; 0.7 to 0.8, moderate; 0.8 to 0.9, good; and 0.9 to 1.0, excellent.[30]

MRI data were available for 99 subjects (1 subject excluded). The clinical tests–the slump test, the SLR test, and the neurologic examination–were analyzed for the entire sample (N=99), while the femoral neurodynamic test was analyzed only for the subjects who received a midlumbar (L2-4) injection (n=18).

Results

In total, 99 patients with chronic, unilateral radicular symptoms were included. Their clinical and demographic characteristics are summarized in table 1.

The comparisons of slump test, SLR test, radiculopathy I, and radiculopathy II in relation to MRI-verified disk extrusion and high-grade nerve compression (foraminal and subarticular) are shown in table 2 (N=99). The comparisons of the femoral neurodynamic test in relation to the MRI findings are shown in table 3 (n=18).

Sensitivity, specificity, positive and negative predictive values, positive and negative likelihood ratios, and AUC of the clinical tests in relationship to disk extrusion, high-grade subarticular nerve compression, and high-grade foraminal nerve compression are displayed in tables 4 through 6.

The slump test (.78) was the only assessment showing high sensitivity (>.70) in detecting disk extrusion, whereas radiculopathy II (.72) was the only assessment showing high specificity (>.70; see table 4). AUC for all assessments ranged from .48 to .60 (see table 4).

All assessments (.71−1.00) had high sensitivity in detecting high-grade subarticular nerve compression, whereas radiculopathy II (.73) was the only assessment showing high specificity (see table 5). AUC ranged from .63 to .82 (see table 5).

Radiculopathy I (.80) was the only assessment showing high sensitivity in detecting high-grade foraminal nerve compression, whereas no assessment had high specificity (see table 6). AUC ranged from .33 to .57 (see table 6).

Discussion

The general findings of this study were that individual clinical tests lack diagnostic accuracy in detecting MRI-verified disk extrusion or high-grade nerve compression. However, specific findings showed that radiculopathy II had both high sensitivity and specificity in detecting high-grade subarticular nerve compression. Moreover, the neurodynamic slump test had high sensitivity in detecting disk extrusion and high-grade subarticular nerve compression. By contrast and in line with all neurodynamic tests, the slump test had a sensitivity of <0.5 in detecting high-grade foraminal compression.

This study differs from most prior studies of diagnostic accuracy regarding lumbar radicular pain[6,31] because we used a level-specific reference standard of either a disk extrusion or a nerve compression and not a reference standard of MRI findings at any level. Thus, it enabled us to compare the MRI findings at 1 level with the clinical tests to assess this particular nerve root and distinguish between 1 or more neurologic signs corresponding to the same nerve root (radiculopathy I or II). Such a distinction has been postulated by Lin et al[3] but was rarely evaluated.

The diagnostic accuracy of clinical tests in identifying radicular pain caused by DH was recently presented in a systematic review.[6] Neurodynamic tests such as the SLR test and the slump test had mostly high sensitivity but low or moderate specificity for detecting DH.[6,9,32,33] Our results mainly support these reports. Interestingly and in accordance with our results, the slump test was reported to be more sensitive than the SLR test.[9,32]

The SLR test was previously reported to have low to moderate sensitivity (.16−.51) in detecting high-grade foraminal nerve compression.[34-37] Our study supports these results as we found low or moderate sensitivity and specificity (<.50) for all neurodynamic tests in detecting high-grade foraminal nerve

Table 4 Concurrent validity of slump test, SLR test, femoral neurodynamic test, radiculopathy I, and radiculopathy II using MRI-verified disk extrusion as reference standard

	Diagnostic Test							
L2-S1 (N=99)	Sens	Spec	AUC	PPV	NPV	LR+	LR−	DOR
Pos slump test	.78 (.59−.89)	.36 (.26−.48)	.57 (.45−.69)	.31 (.21−.44)	.81 (.64−.93)	1.22 (0.93−1.59)	0.61 (0.29−1.33)	2.00 (0.71−5.53)
Pos SLR test	.59 (.41−.75)	.53 (.41−.64)	.56 (.43−.69)	.31 (.20−.47)	.78 (.63−.88)	1.26 (0.84−1.87)	0.77 (0.47−1.28)	1.63 (0.66−3.98)
Radiculopathy I*	.67 (.46−.83)	.29 (.19−.41)	.48 (.35−.61)	.26 (.16−.38)	.70 (.51−.85)	0.94 (0.69−1.28)	1.14 (0.60−2.18)	0.82 (0.32−2.13)
Radiculopathy II†	.48 (.29−.68)	.72 (.60−.82)	.60 (.47−.73)	.39 (.23−.58)	.79 (.67−.88)	1.71 (1.01−2.98)	0.72 (0.49−1.06)	2.41 (0.97−6.02)
L2-4 (n=18)								
Pos femoral neurodynamic test	.43 (.16−.75)	.64 (.35−.85)	.53 (.25−.81)	.43 (.10−.82)	.64 (.31−.89)	1.18 (0.37−3.75)	0.90 (0.41−1.96)	1.31 (0.19−9.10)

NOTE. Values in each cell are calculated values and 95% confidence intervals.

Abbreviations: LR+, positive likelihood ratio; LR−, negative likelihood ratio; NPV, negative predictive value; Pos, positive; PPV, positive predictive value; Sens, sensitivity; Spec, specificity.

* Assessed with 1 neurologic sign.

† Assessed with 2 neurologic signs corresponding to a specific nerve root.

Table 5 Concurrent validity of slump test, SLR test, femoral neurodynamic test, radiculopathy I, and radiculopathy II using MRI-verified high-grade subarticular nerve root compression as reference standard

	Diagnostic Test							
L2-S1 (N=99)	Sens	Spec	AUC	PPV	NPV	LR+	LR−	DOR
Pos slump test	1.00 (0.77−1.00)	.38 (.27−.49)	.70 (.58−.81)	.21 (.12−.33)	1.00 (0.89−1.00)	1.64 (1.36−1.89)	0	N/A
Pos SLR test	0.93 (0.66−1.00)	.57 (.45−.67)	.75 (.63−.86)	.26 (.15−.40)	0.98 (0.89−1.00)	2.13 (1.61−2.83)	.13 (.02−.84)	16.9 (2.11−134)
Radiculopathy I*	0.93 (0.66−1.00)	.34 (.24−.45)	.63 (.49−.77)	.19 (.10−.30)	0.97 (0.83−1.00)	1.41 (1.14−1.74)	.21 (0.03−1.42)	6.73 (0.84−54)
Radiculopathy II†	0.71 (0.42−0.92)	.73 (.62−.82)	.72 (.58−.87)	.30 (.16−.49)	0.94 (0.85−0.98)	2.62 (1.63−4.27)	.39 (.17−.91)	6.74 (1.92−24)
L2−4 (n=18)								
Pos femoral neurodynamic test	1.00 (0.21−1.00)	.65 (.41−.83)	.82 (.56−1.00)	.14 (.01−.58)	1.00 (0.72−1.00)	2.83 (1.49−5.39)	0	N/A

NOTE. Values in each cell are calculated values and 95% confidence intervals.

Abbreviations: LR+, positive likelihood ratio; LR−, negative likelihood ratio; N/A, not available; NPV, negative predictive value; Pos, positive; PPV, positive predictive value; Sens, sensitivity; Spec, specificity.

* Assessed with 1 neurologic sign.

† Assessed with 2 neurologic signs corresponding to a specific nerve root.

Table 6 Concurrent validity of slump test, SLR test, femoral neurodynamic test, radiculopathy I, and radiculopathy II using MRI-verified high-grade foraminal nerve root compression as reference standard

	Diagnostic Test							
L2-S1 (N=99)	Sens	Spec	AUC	PPV	NPV	LR+	LR−	DOR
Pos slump test	.48 (.29−.69)	.26 (.16−.37)	.37 (.24−.50)	.18 (.10−.29)	.59 (.41−.76)	0.66 (0.42−0.99)	2.03 (1.18−3.48)	0.32 (0.12−0.82)
Pos SLR test	.32 (.17−.52)	.43 (.33−.55)	.38 (.25−.51)	.16 (.07−.29)	.65 (.50−.78)	0.56 (0.31−1.03)	1.55 (1.08−2.29)	0.36 (0.14−0.93)
Radiculopathy I*	0.80 (.61−.91)	.34 (.23−.46)	.57 (.44−.69)	.29 (.19−.41)	.83 (.65−.94)	1.21 (0.94−1.56)	0.59 (0.25−1.38)	2.04 (0.68−6.08)
Radiculopathy II†	.28 (.14−.48)	.65 (.53−.75)	.47 (.34−.60)	.21 (.09−.39)	.73 (.60−.83)	0.80 (0.40−1.61)	1.11 (0.82−1.49)	0.72 (0.26−1.94)
L2-4 (n=18)								
Pos femoral neurodynamic test	.17 (.03−.56)	.50 (.25−.75)	.33 (.07−.60)	.14 (.04−.56)	.55 (.23−.83)	0.33 (0.05−2.18)	1.67 (0.85−3.26)	0.20 (0.02−2.26)

NOTE. Values in each cell are calculated values and 95% confidence intervals.
Abbreviations: LR+, positive likelihood ratio; LR−, negative likelihood ratio; NPV, negative predictive value; Pos, positive; PPV, positive predictive value; Sens, sensitivity; Spec, specificity.
* Assessed with 1 neurologic sign.
† Assessed with 2 neurologic signs corresponding to a specific nerve root.

compression. Thus, a positive neurodynamic test seems not to be an indicator of foraminal nerve compression.

Prior studies of the femoral neurodynamic test in detecting nerve compression caused by DH reported either moderate sensitivity and high specificity or vice versa.[10,38] Our results showed only moderate sensitivity and specificity in detecting midlumbar (L2-4) disk extrusion, and high sensitivity in detecting subarticular nerve compression. However, the latter result must be interpreted with caution because of low positive findings at the L2-4 level (n=1).

Individual neurologic tests have previously been presented with insufficient accuracy in detecting radicular symptoms caused by DH.[6,11,31,39,40] The tests lack sensitivity both in detecting DH[31] and in detecting high-grade foraminal nerve compression.[34,35,37] In accordance with these results, we found that individual neurologic tests (as seen in table 1) lacked sensitivity in detecting disk extrusion and high-grade nerve compression (data not shown). However, radiculopathy I in our report reflects on 1 neurologic deficit, sensory deficit, or motor/reflex impairment, and consequently, higher sensitivity was found for radiculopathy I than for the individual test on its own.

The second radiculopathy variable, radiculopathy II, lacked in sensitivity but showed the highest specificity (.65−.73) of all tests in detecting all MRI findings, and has to our knowledge never been tested for validity.

Other combinations of tests were previously reported. Majlesi et al[9] improved the specificity by combining the SLR test and neurologic deficits; the sensitivity decreased accordingly, however. Others have combined the patient's history with clinical tests and thus somewhat improved the diagnostic accuracy of detecting DH.[11,14,41] However, no proposal was suggested of how to best detect DH. In a similar manner, Hancock et al[15] found slightly better accuracy by using combinations of tests compared with individual tests to identify the level of DH.

Overall, the results of this study were disappointing because the accuracy of the clinical assessments was generally poor to fair (AUC<0.8). A likely interpretation of the results is that the clinical examination and MRI assess different aspects of nerve root involvement. A positive clinical test might be a sign of a chemically induced radicular pain that cannot be visualized using MRI. Inflammation of the nerve root may have caused changes to nerve fibers, resulting in increased stretch mechanosensitivity.[42] This might suggest that a positive neurodynamic test is a sign of inflammation rather than a high-grade nerve compression.

In validity studies such as the present study, the likelihood ratio and DOR are recommended as complement outcomes to sensitivity, specificity, and AUC.[43,44] However, in the present study, the positive likelihood ratio and DOR were comparatively low. The highest positive likelihood ratio and DOR of this report were 2.6 and 16.9, respectively. Thus, the results correspond to the results of the AUC.

Study limitations

There are limitations to this study. First, our subjects had chronic symptoms and that was postulated as a confounding factor for low accuracy in the clinical testing.[39] Moreover, our subjects all had radicular symptoms and were referred for TESI but not referred for surgery. This might be explained by a lack of precise clinical findings in some cases. On the other hand, in comparison to an unselected primary care population, the likelihood of clinical findings was proposed to be higher in a specialized care sample such as ours.[6] In fact, results from our cohort cannot be

generalized to all subjects with lumbar radicular pain. Second, the choice of using an imaging reference standard in diagnostic studies of nerve root compression and DH has been questioned.[45] MRI as a reference standard creates the potential for bias because of false-positive test results, as the prevalence of nerve root compression and DH in an asymptomatic population was reported to be 5% to 25%.[46,47] However, to use surgical findings as a reference standard is considered to be a less optimal alternative because the selection bias increases.[6,31] Finally, the MRI analysis was level-specific and was focused on DH or nerve root compression. Pathology of the nonaffected side and at other levels was disregarded, and thus it enabled us to compare the MRI findings at the specific level. However, other MRI findings such as Modic changes and DH at other levels were not assessed and might have affected the results.

In patients with a longer symptom duration than 3 months, the clinical testing and treatment decision-making are essential. We thus believe our results are of clinical interest. In summary, radiculopathy II had the highest diagnostic accuracy in detecting disk extrusion and subarticular nerve compression because of its comparatively high specificity. From a clinician's viewpoint, the slump test appears to be the most useful test (because of its high sensitivity) in selecting patients for MRI in the process of diagnosing disk extrusion and subarticular nerve compression. However, based on our results, we cannot recommend any neurodynamic tests for detecting foraminal nerve compression visualized on MRI, but recommend neurologic signs as indicators.

Conclusions

In general, the investigated neurodynamic tests or assessments for radiculopathy lacked diagnostic accuracy. The neurodynamic slump test was the most sensitive test and radiculopathy II was the most specific in detecting extrusion and high-grade subarticular nerve compression. Radiculopathy I was the most sensitive in detecting high-grade foraminal nerve compression (foraminal stenosis), whereas no relationship was found between any neurodynamic test and foraminal stenosis.

Suppliers

a. 1.5-tesla Siemens Avanto scanner; Siemens Corp.
b. SPSS (version 23.0); IBM Corp.
c. R software; The R Project for Statistical Computing. Available at: http://expasy.org/tools/pROC/ under the GNU General Public License.

Keywords

Magnetic resonance imaging; Neurologic examination; Radiculopathy; Rehabilitation; Validity of results

Corresponding author

Harald Ekedahl, PT, PhD, Clinical Sciences Lund, Lund University, Box 117, 221 00 Lund, Sweden. *E-mail address:* harald.ekedahl@med.lu.se.

Acknowledgments

We thank Johan Tidstrand, PT, MSC, for performing the clinical test in figures 4-6, and a special thanks to the orthopedic clinic and the radiology department at Ängelholm Hospital.

References

1. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016;388:1545-602.
2. Konstantinou K, Dunn KM. Sciatica: review of epidemiological studies and prevalence estimates. Spine (Phila Pa 1976) 2008;33:2464-72.
3. Lin CW, Verwoerd AJ, Maher CG, et al. How is radiating leg pain defined in randomized controlled trials of conservative treatments in primary care? A systematic review. Eur J Pain 2014;18:455-64.
4. Bernstein IA, Malik Q, Carville S, Ward S. Low back pain and sciatica: summary of NICE guidance. BMJ 2017;356:i6748.
5. Kreiner DS, Hwang SW, Easa JE, et al. An evidence-based clinical guideline for the diagnosis and treatment of lumbar disc herniation with radiculopathy. Spine J 2014;14:180-91.
6. van der Windt D, Simons E, Riphagen II, et al. Physical examination for lumbar radiculopathy due to disc herniation in patients with low-back pain. Cochrane Database Syst Rev 2010;(2):CD007431.
7. Pfirrmann CW, Dora C, Schmid MR, Zanetti M, Hodler J, Boos N. MR image-based grading of lumbar nerve root compromise due to disk herniation: reliability study with surgical correlation. Radiology 2004;230:583-8.
8. Fardon DF, Milette PC; Combined Task Forces of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. Nomenclature and classification of lumbar disc pathology. Recommendations of the Combined task Forces of the North American Spine Society, American Society of Spine Radiology, and American Society of Neuroradiology. Spine (Phila Pa 1976) 2001; 26:E93-113.
9. Majlesi J, Togay H, Unalan H, Toprak S. The sensitivity and specificity of the slump and the straight leg raising tests in patients with lumbar disc herniation. J Clin Rheumatol 2008;14:87-91.
10. Suri P, Rainville J, Katz JN, et al. The accuracy of the physical examination for the diagnosis of midlumbar and low lumbar nerve root impingement. Spine (Phila Pa 1976) 2011;36:63-73.
11. Iversen T, Solberg TK, Romner B, et al. Accuracy of physical examination for chronic lumbar radiculopathy. BMC Musculoskelet Disord 2013;14:206.
12. Capra F, Vanti C, Donati R, Tombetti S, O'Reilly C, Pillastrini P. Validity of the straight-leg raise test for patients with sciatic pain with or without lumbar pain using magnetic resonance imaging results as a reference standard. J Manipulative Physiol Ther 2011;34:231-8.
13. Merskey H, Bogduk N. Classification of chronic pain: descriptions of chronic pain syndromes and definition of pain terms. Seattle: IASP Pr; 1994.
14. Vroomen PC, de Krom MC, Wilmink JT, Kester AD, Knottnerus JA. Diagnostic value of history and physical examination in patients suspected of lumbosacral nerve root compression. J Neurol Neurosurg Psychiatry 2002;72:630-4.
15. Hancock MJ, Koes B, Ostelo R, Peul W. Diagnostic accuracy of the clinical examination in identifying the level of herniation in patients with sciatica. Spine (Phila Pa 1976) 2011;36:E712-9.
16. Ekedahl H, Jonsson B, Annertz M, Frobell RB. Three week results of transforaminal epidural steroid injection in patients with chronic unilateral low back related leg pain: the relation to MRI findings and clinical features. J Back Musculoskelet Rehabil 2016; 29:693-702.

17. Ghahreman A, Bogduk N. Predictors of a favorable response to transforaminal injection of steroids in patients with lumbar radicular pain due to disc herniation. Pain Med 2011;12:871-9.
18. Lurie JD, Tosteson AN, Tosteson TD, et al. Reliability of magnetic resonance imaging readings for lumbar disc herniation in the Spine Patient Outcomes Research Trial (SPORT). Spine (Phila Pa 1976) 2008;33:991-8.
19. Lee S, Lee JW, Yeom JS, et al. A practical MRI grading system for lumbar foraminal stenosis. AJR Am J Roentgenol 2010;194:1095-8.
20. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry Low Back Pain Disability Questionnaire. Physiotherapy 1980;66:271-3.
21. Philip K, Lew P, Matyas TA. The inter-therapist reliability of the slump test. Aus J Physiother 1989;35:89-94.
22. Waddell G, Somerville D, Henderson I, Newton M. Objective clinical evaluation of physical impairment in chronic low back pain. Spine (Phila Pa 1976) 1992;17:617-28.
23. Shacklock M. Neurodynamics. Physiotherapy 1995;81:9-16.
24. Trainor K, Pinnington MA. Reliability and diagnostic validity of the slump knee bend neurodynamic test for upper/mid lumbar nerve root compression: a pilot study. Physiotherapy 2011;97:59-64.
25. Vroomen PC, de Krom MC, Knottnerus JA. Consistency of history taking and physical examination in patients with suspected lumbar nerve root involvement. Spine (Phila Pa 1976) 2000;25:91-6; discussion 97.
26. Walsh J, Hall T. Agreement and correlation between the straight leg raise and slump tests in subjects with leg pain. J Manipulative Physiol Ther 2009;32:184-92.
27. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005;58:982-90.
28. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 1994;271:703-7.
29. Scaia V, Baxter D, Cook C. The pain provocation-based straight leg raise test for diagnosis of lumbar disc herniation, lumbar radiculopathy, and/or sciatica: a systematic review of clinical utility. J Back Musculoskelet Rehabil 2012;25:215-23.
30. Roelen CA, van Rhenen W, Groothoff JW, van der Klink JJ, Bultmann U, Heymans MW. The development and validation of two prediction models to identify employees at risk of high sickness absence. Eur J Public Health 2013;23:128-33.
31. Vroomen PC, de Krom MC, Knottnerus JA. 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.
32. Stankovic R, Johnell O, Maly P, Willner S. Use of lumbar extension, slump test, physical and neurological examination in the evaluation of patients with suspected herniated nucleus pulposus. A prospective clinical study. Man Ther 1999;4:25-32.
33. Poiraudeau S, Foltz V, Drape JL, et al. Value of the bell test and the hyperextension test for diagnosis in sciatica associated with disc herniation: comparison with Lasegue's sign and the crossed Lasegue's sign. Rheumatology (Oxford) 2001;40:460-6.
34. Yamada K, Aota Y, Higashi T, et al. Lumbar foraminal stenosis causes leg pain at rest. Eur Spine J 2014;23:504-7.
35. Jönsson B, Strömqvist B. Symptoms and signs in degeneration of the lumbar spine. A prospective, consecutive study of 300 operated patients. J Bone Joint Surg Br 1993;75:381-5.
36. Lyle MA, Manes S, McGuinness M, Ziaei S, Iversen MD. Relationship of physical examination findings and self-reported symptom severity and physical function in patients with degenerative lumbar conditions. Phys Ther 2005;85:120-33.
37. Konno S, Kikuchi S, Tanaka Y, et al. A diagnostic support tool for lumbar spinal stenosis: a self-administered, self-reported history questionnaire. BMC Musculoskelet Disord 2007;8:102.
38. Kido T, Okuyama K, Chiba M, et al. Clinical diagnosis of upper lumbar disc herniation: pain and/or numbness distribution are more useful for appropriate level diagnosis. J Orthop Sci 2016;21:419-24.
39. Bertilson B, Brosjö E, Billing H, Strender LE. Assessment of nerve involvement in the lumbar spine: agreement between magnetic resonance imaging, physical examination and pain drawing findings. BMC Musculoskelet Disord 2010;11:202.
40. Al Nezari NH, Schneiders AG, Hendrick PA. Neurological examination of the peripheral nervous system to diagnose lumbar spinal disc herniation with suspected radiculopathy: a systematic review and meta-analysis. Spine J 2013;13:657-74.
41. Vucetic N, Astrand P, Guntner P, Svensson O. Diagnosis and prognosis in lumbar disc herniation. Clin Orthop Relat Res 1999;(361):116-22.
42. Dilley A, Lynn B, Pang SJ. Pressure and stretch mechanosensitivity of peripheral nerve fibres following local inflammation of the nerve trunk. Pain 2005;117:462-72.
43. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. BMJ 2004;329:168-9.
44. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. J Clin Epidemiol 2003;56:1129-35.
45. Tawa N, Rhoda A, Diener I. Accuracy of magnetic resonance imaging in detecting lumbo-sacral nerve root compromise: a systematic literature review. BMC Musculoskelet Disord 2016;17:386.
46. Jarvik JG, Hollingworth W, Heagerty PJ, Haynor DR, Boyko EJ, Deyo RA. Three-year incidence of low back pain in an initially asymptomatic cohort: clinical and imaging risk factors. Spine (Phila Pa 1976) 2005;30:1541-8; discussion 1549.
47. Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic resonance imaging of the lumbar spine in people without back pain. N Engl J Med 1994;331:69-73.