Spine

SPINE Volume 36, Number 1, pp 63–73

DIAGNOSTICS

The Accuracy of the Physical Examination for the Diagnosis of Midlumbar and Low Lumbar Nerve Root Impingement

Pradeep Suri, MD,*†‡§ James Rainville, MD,*‡ Jeffrey N. Katz, MD, MS,¶ Cristin Jouve, MD,*‡ Carol Hartigan, MD,*‡ Janet Limke, MD,*‡ Enrique Pena, MD,* Ling Li, MPH,* Bryan Swaim, MS,* and David J. Hunter, MBBS, PhD*

Study Design. Cross-sectional study with prospective recruitment.

Objective. To determine the accuracy of the physical examination for the diagnosis of midlumbar nerve root impingement (L2, L3, or L4), low lumbar nerve root impingement (L5 or S1) and level-specific lumbar nerve root impingement on magnetic resonance imaging, using individual tests and combinations of tests.

Summary of Background Data. The sensitivity and specificity of the physical examination for the localization of nerve root impingement has not been previously studied.

Methods. Sensitivities, specificities, and likelihood ratios (LRs) were calculated for the ability of individual tests and test combinations to predict the presence or absence of nerve root impingement at midlumbar, low lumbar, and specific nerve root levels.

Results. LRs ≥5.0 indicate moderate to large changes from pre-test probability of nerve root impingement to post-test probability. For the diagnosis of midlumbar impingement, the femoral stretch test (FST), crossed FST, medial ankle pinprick sensation, and patellar reflex testing demonstrated LRs ≥5.0 (LR ∞). LRs ≥5.0 were observed with the combinations of FST and either patellar reflex testing (LR 7.0; 95% confidence interval [CI] 2.3–21) or the sit-to-stand test (LR ∞). For the diagnosis of low lumbar impingement, the Achilles reflex test demonstrated an LR ≥5.0 (LR 7.1; 95% CI 0.96–53); test combinations did not increase LRs. For the diagnosis of level-specific impingement, LRs ≥5.0 were observed for anterior thigh sensation at L2 (LR 13; 95% CI 1.8–87); FST at L3 (LR 5.7; 95% CI 2.3–4.4); patellar reflex testing (LR 7.7; 95% CI 1.7–35), medial ankle sensation (LR ∞), or crossed FST (LR 13; 95% CI 1.8–87) at L4; and hip abductor strength at L5 (LR 11; 95% CI 1.3–84). Test combinations increased LRs for level-specific root impingement at the L4 level only.

Conclusion. Individual physical examination tests may provide clinical information that substantially alters the likelihood that midlumbar impingement, low lumbar impingement, or level-specific impingement is present. Test combinations improve diagnostic accuracy for midlum-bar impingement.

Key words: physical examination, diagnostic, accuracy, sensitivity, specificity, lumbar, back pain, radiculop-athy, herniation, impingement. **Spine 2011;36:63–73**

Midlumbar nerve root impingement, or nerve root impingement at the L2, L3, or L4 levels, is a cause of lum-bosacral radicular syndrome found with increasing frequency in older adults.[1,2] Although early studies estimated the prevalence of midlumbar nerve root impingement at 5% to 11%,[2,3] results from recent studies of lumbar disc herniation suggest that the prevalence of midlumbar nerve root impingement may be substantially higher.[4–6] As our population ages, midlumbar nerve root impingement may be observed with increasing frequency in specialty spine clinics.[7]

Although the classic presentation of radicular pain in midlumbar nerve root impingement is in the groin or anterior thigh, pain may also commonly be experienced in the calf, ankle, or foot.[8] The diagnosis of midlumbar nerve root impingement may, therefore, be quite challenging, with a clinical presentation resembling lower lumbar nerve root impingement, hip os-teoarthritis, and other causes of referred pain.[9] The physical examination may be helpful in clinical decision-making by altering the post-test probability that nerve root impingement localized to a region or a specific level is responsible for the production of symptoms. Establishing a clear picture of clinical deficits before obtaining advanced imaging is necessary to avoid a situation where the physical examination

From the *Division of Research, New England Baptist Hospital, Boston, MA; †Spaulding Rehabilitation Hospital, Boston, MA; ‡Depart-ment of Physical Medicine and Rehabilitation, Harvard Medical School, Boston, MA; §VA Boston Healthcare System, Boston, MA; and ¶ Division of Rheumatology, Immunology and Allergy, Department of Medicine and Department of Orthopedic Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA.

Acknowledgment date: August 21, 2009. First revision date: October 20, 2009. Second revision date: October 27, 2009. Acceptance date: November 3, 2009.

The manuscript submitted does not contain information about medical device(s)/drug(s).

Funds were received to support this work from NIH. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript. Supported by the Rehabilitation Medicine Scientist Training Program (RMSTP) and the National Institutes of Health grant K12 HD 01097 (to P.S.), and, in part, by the grant NIH/NIAMS K24 AR 02123 (to J.N.K.).

Address correspondence and reprint requests to Pradeep Suri, MD, Division of Research, New England Baptist Hospital, 125 Parker Hill Ave Boston, MA 02130; E-mail: psuri@caregroup.harvard.edu

becomes biased by prior knowledge of an abnormality on imaging, especially in light of the well-documented prevalence of asymptomatic disc herniations using lumbar spine magnetic resonance imaging (MRI),[10] and the common clinical occurrence of multiple imaging abnormalities in a single patient.

Although prior studies have examined the performance characteristics of the physical examination in patients with radicular pain, there are deficits in the existing literature. First, prior studies have examined the physical examination in reference to lumbar disc herniation, but not in reference to lumbar nerve root impingement. This fact is noteworthy, considering that decompression of nerve impingement (and not discectomy alone) is the primary goal of decompression surgery. Second, the basis of the most commonly used physical examination tests in the evaluation of lumbosacral radicular pain, including the straight leg raise (SLR) test, rests on the assumption that nerve root pathology affects the low lumbar nerve roots (L5 or S1).[11] The performance characteristics of most common physical examination tests in the setting of midlumbar nerve root impingement, where the L5 or S1 nerve roots are not affected, is unclear. Third, although data exist on the accuracy of the physical examination for detecting lumbar disc herniation without regard to herniation level, there are no accuracy studies on the performance of the physical examination tests for the neuroanatomic localization of nerve root impingement. The conceptual localization of nerve root impingement by the examining physician before reviewing MRI results is vital to ensure that imaging is correlated with the clinical picture, rather than vice versa.

We sought to characterize the accuracy of physical examination testing over the range of lumbar nerve root impingement, and to elucidate how the physical examination may aid in neuroanatomic localization of pathology. In this prospective study of patients with lumbosacral radicular pain, we examined the performance characteristics of single physical examination tests in relation to 3 different reference standards using lumbar spine MRI: (1) midlumbar nerve root impingement (L2, L3, or L4), (2) low lumbar nerve root impingement (L5 or S1), and (3) level-specific nerve root impingement at the L2–S1 levels. We also examined the performance characteristics of test combinations to optimize diagnostic performance and to simulate how the physical examination is commonly used in clinical practice.

MATERIALS AND METHODS

Study Participants

This work was an ancillary study to a prospective study of the outcomes of lumbar disc herniation. The study was approved by the hospital's institutional review board. Participants were recruited from a hospital spine center between January 2008 and March 2009. All consecutive patients aged 18 and older with lower extremity radiating pain for <12 weeks were eligible for participation. Inclusion criteria were a history of radicular pain in a L2, L3, L4, L5, or S1 dermatome, with or without neurologic symptoms. Exclusion criteria were known pregnancy and severe active medical or psychiatric comorbidities that would limit study participation. The presence of symptomatic hip arthritis, prior total hip arthroplasty, hip flexion contractures, and other hip disorders were not criteria for exclusion. Patients with lumbar spine MRI available to the examining physician at the time of the physical examination were ineligible for this ancillary study, to ensure blinding to the reference standard and eliminate potential bias from knowledge of nerve root compression before performing the physical examination.

Information on participant demographics and clinical characteristics was collected prospectively, including participant age, gender, and symptom duration. Comorbidity burden was measured by the Self-Acquired Comorbidity Questionnaire, which is a commonly used, valid, and reliable measure.[12] Disability was measured by the Oswestry Disability Index. The Oswestry Disability Index is a condition-specific measure of disability, which has been used extensively in prior studies of radiculopathy, and has demonstrated validity and reliability in these contexts.[13] Pain was measured by the visual analog scale for leg pain and back pain.[14]

Physical Examination

Each participant was examined by 1 of 6 board-certified physiatrists specializing in spine care. The physical examination included a standard battery of tests, administered by the physiatrist in a consistent and stereotyped manner. All physical examination tests were performed bilaterally. Testing results were documented by the examiner in reference to the symptomatic limb; for example, a positive SLR was documented if reproduction of radicular pain was elicited in the symptomatic limb. In a minority of cases, in whom bilateral symptoms existed, the results of testing were documented in reference to the limb that was most symptomatic. The examining physician recorded all information at the time of the examination using a standardized data sheet.

Provocative Manuevers

a. *Straight leg raise/crossed straight leg raise:* The SLR and crossed SLR have been well described previously.[15] Reproduction of the patient's typical lower extremity pain constitutes a positive result for both the SLR and the crossed SLR.
b. Femoral stretch test/crossed femoral stretch test: The femoral stretch test (FST) is performed with the patient lying prone. The examiner grasps the patient's ankle on the symptomatic (ipsilateral) side and facilitates gentle ipsi-lateral knee flexion; reproduction of typical lower extremity pain constitutes a positive test. The crossed femoral stretch test (CFST) is performed similarly, except that contralateral knee flexion instead reproduces typical (ipsilateral) lower extremity pain.

Motor Testing

For assessment of knee extensor strength, ankle dorsiflexor strength, and ankle plantarflexor strength, the examiner performs a maximally challenging and "functional" test of

Figure 1. The single-leg sit-to-stand test. **(A)** Starting position. The seated patient fully extends 1 knee and avoids contacting the floor with that foot during testing. The examiner holds the patient's hands as a balance aid only. **(B)** Finishing position. The patient rises to the standing position using only the strength of the supporting limb. Inability to transfer to standing constitutes a positive test.

strength to identify subtle deficits. If the participant is able to complete the functional test, they receive a grade of normal strength (5), and no further strength testing is performed for that muscle group. If an impairment is detected with the functional test, normal strength is not present, and the degree of weakness is quantified (0 – 4) using manual muscle testing.[16]

a. *Knee extensor strength:* The single-leg sit-to-stand test is performed as the primary assessment of knee extensor strength.[8] The performance of this test is demonstrated in Figure 1. The test is performed first on the unaffected leg and second on the affected leg.
b. *Ankle dorsiflexor strength:* The heel walk test is per formed as the primary assessment of ankle dorsiflexion-strength. The patient walks on heels only while avoiding contacting the floor with the forefeet, using the examiner's hands for balance as needed; inability to maintain the forefoot off the ground is a positive result.
c. *Ankle plantarflexor strength:* The single-leg heel raise test is performed as the primary assessment of ankle plantarflexion strength. The patient stands on 1 foot while flexing the contralateral knee, while holding the examiner's hands or a countertop for balance. The patient plantar-flexes the ankle, raising the heel of the supporting limb off the floor to maximal plantarflexion. Inability to perform 10 successive heel raises is a positive result.
d. *Great toe extensor strength:* The patient fully dorsiflexes the great toe and maintains this position as the examiner applies a plantarflexion force.
e. *Hip abductor strength:* The patient is positioned in side-lying with the asymptomatic limb against the table surface and the symptomatic thigh abducted to 30° from horizontal. The patient maintains the abducted position as the examiner applies an adduction force.

Sensory Testing

Sensory impairments are assessed by pinprick testing at the midanterior thigh, the medial aspect of the knee, the medial aspect of the ankle, the dorsal aspect of the great toe, and the lateral border of the foot. Sensation is graded on a standard 3-grade scale (0 – 2), with "2" representing normal sensation, "1" representing impairment, and "0" representing absent sensation.[17]

Reflex Testing

Deep tendon reflexes are assessed at the patellar tendon and Achilles tendon, using a standard 5-point grading scale.[18]

MRI Studies

MRI was performed for the majority of participants within 7 days of their physical examination. MRI scans consisted at minimum of T1- and T2-weighted images of the lumbar spine in the sagittal and axial planes. Each MRI scan was evaluated by 1 of 8 board-certified neuroradiologists, who were blinded to study design and research questions. The classification of the most severe level of nerve root impingement according to the neuroradiologist final impression was used as the reference standard for this study. Inter-rater reliability of the most severe level of nerve root impingement at the L2–S1 root levels on each MRI was excellent ($\kappa = 0.92$) in a subsample of 18 scans, which were interpreted by an independent and blinded muscu-loskeletal radiologist who did not participate in the primary MRI assessments. The evaluation of nerve root impingement on MRI has previously been validated in reference to operative findings.[19]

Statistical Analysis

To characterize the demographics, clinical characteristics, and radiographic features of the study population, we calculated means and standard deviations for continuous variables and frequencies and proportions for categorical variables. We calculated the frequencies of pain symptom locations by root impingement level. For analytic purposes, the results of physical examination tests with a categorical grading were dichotomized: motor strength testing was dichotomized as negative (5) *versus* positive (0–4); sensory testing was dichotomized as negative (2) *versus* positive (0 or 1); and reflex testing was considered positive if the reflex grade in the symptomatic limb was diminished by at least 1 grade compared with the same reflex in the contralateral limb. We used the reference standard of the most severe level of nerve root impingement on lumbar spine MRI.

We began by calculating sensitivities, specificities, and likelihood ratios (LRs), including 95% confidence intervals (CIs), for all individual tests using the reference standard of MRI midlumbar nerve root impingement. This analytic approach has been recommended by recent research guidelines for studies of diagnostic accuracy.[20,21] Positive LRs were calculated as "sensitivity/(1 – specificity)," and negative LRs were calculated as "(1 – sensitivity)/specificity."[21] We then repeated these calculations for all individual tests using a reference standard of MRI low lumbar nerve root impingement, and a reference standard of MRI level-specific nerve root impingement.

Next, we calculated the performance characteristics of test combinations for the diagnosis of midlumbar nerve root impingement. We combined tests in a manner designed to optimize sensitivity, by considering a test combination to be positive if any test in the combination had a positive result. To limit the number of possible test combinations, we included only individual tests that demonstrated LR 95% CIs ≥1.0, and LR point estimates ≥2.0. We chose an LR point estimate threshold of 2.0 because LRs ≥2.0 reflect small but sometimes important changes in post-test probability.[22] We expected that certain test combinations would result in more substantial LRs ≥5.0, which indicate moderate to large changes from the pre-test probability of nerve root impingement to post-test probability.[22] We then repeated these analyses of test combinations, using a reference standard of low lumbar nerve root impingement, and a reference standard of level-specific root impingement. All analyses were performed using SAS software, version 9.0 (SAS Institute, Cary, NC).

RESULTS

Participant recruitment for this study is depicted in Figure 2. Of the 170 potential participants, 10 individuals either declined to participate or were missed by the recruiting physicians. A total of 160 participants were consented, including 57 participants who had no imaging available and 103 participants who had imaging available at the time of the examination. All 103 participants with imaging available to the examiner were ineligible for this study because of lack of blinding to the reference standard. Of the 57 participants who had no imaging available, 3 participants did not go on to receive MRI because of clinical improvement, leaving 54 participants who received MRI and neuroradiologist interpretation; these individuals reported a slightly shorter duration of symptoms (4.3 ± 2.8 *vs*. 5.2 ± 3.1; $P = 0.08$), but were otherwise not materially different from the 103 individuals who had no imaging available (data not shown).

Demographics and clinical characteristics of the study sample are presented in Table 1. The 3 participants who did not go on to receive MRI were not materially different from those who did go on to MRI (data not shown). Table 2 demonstrates the prevalence of radicular pain experienced by the patient in each anatomic region of the lower extremities, according to the specific level of nerve root impingement. Fifty-one participants had nerve root impingement primarily because of lumbar disc herniation, though concurrent mild to moderate bony stenosis at the affected level was not uncommon. One participant had a synovial cyst, and 2 participants had bony stenosis as the primary cause of impingement. Forty (74%) participants had nerve root compression, 11 (20%) had nerve root deviation, and 3 (6%) had nerve root contact. All 3 participants with nerve root contact had only single-level involvement. Fifteen (28%) individuals also had nerve contact at another level.

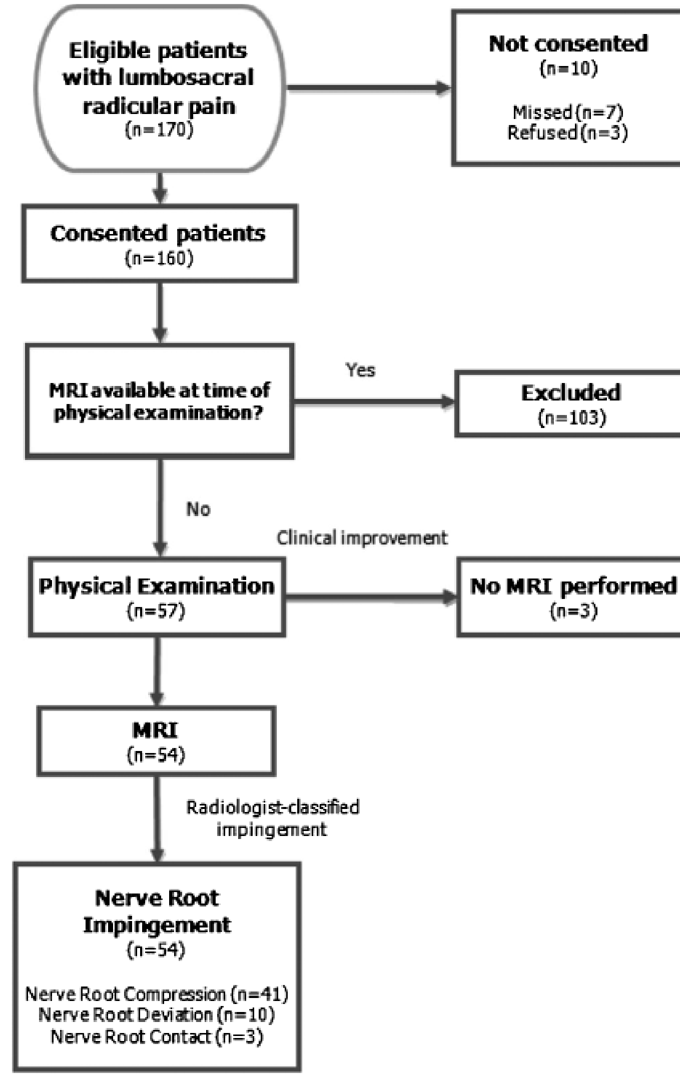

Figure 2. Flowchart of patient recruitment and participation.

The distributions of patient-reported radicular pain symptoms are presented in Table 2. Anterior *versus* posterior location of thigh pain symptoms showed notable differences by location of impingement. For example, all patients who reported anterior thigh pain had midlum-bar nerve root impingement, whereas all patients who reported posterior thigh pain had low lumbar nerve root impingement. Anterior *versus* posterior location of calf pain symptoms also showed notable differences by location of impingement. For example, all but 1 patient with anterior calf pain had midlumbar nerve root impingement, whereas all patients with posterior calf pain had low lumbar nerve root impingement. All but 1 patient with pain in the foot or ankle had low lumbar nerve root impingement.

The performance characteristics of individual physical examination tests for the diagnosis of midlumbar nerve root impingement and low lumbar nerve root impingement are presented in Table 3. Positive results on 4 physical examination tests had both

TABLE 1. Characteristics of the Study Sample

Characteristic	Midlumbar Nerve Root Impingement L2, L3, or L4* (n = 25)	Low Lumbar Nerve Root Impingement L5 or S1* (n = 29)	Nerve Root Impingement at Any Level* (n = 54)
Age (yrs)	61.4 (13.8)	48.6 (13.6)	54.5 (15.0)
Female (%)	9 (36.0)	4 (13.8)	13 (24.1)
Symptom duration (wks)	3.9 (2.5)	4.6 (3.0)	4.3 (2.8)
Oswestry Disability Index (0–100)	42 (21)	49 (21)	46 (21)
VAS Leg Pain (0–10)	7.2 (2.2)	7.0 (2.7)	7.1 (2.5)
VAS Back Pain (0–10)	4.8 (3.2)	5.4 (3.4)	5.1 (3.3)
Self-Acquired Comorbidity Questionnaire (0–45)	4.6 (4.0)	1.4 (2.3)	2.8 (3.5)
*Mean (SD) or N (%).			

LR point estimates ≥5.0 and 95% CIs ≥1.0 for the diagnosis of midlumbar nerve root impingement on MRI: the FST (∞), CFST (∞), medial ankle pinprick sensation (∞), and patellar reflex testing (∞). An impaired Achilles reflex was the only test with an LR point estimate ≥5.0 and 95% CIs ≥1.0 for the diagnosis of low lumbar nerve root impingement on MRI. Although the crossed SLR test demonstrated higher specificity compared with the SLR (96% vs. 84%, respectively), the FST demonstrated such high specificity (100%) that the crossed FST could provide no additional gain in specificity. Furthermore, the CFST was positive in only 1 participant, and, therefore, the interpretation of performance characteristics for this test should be viewed in this context (Supplemental Digital Content, Table 1, available at: http://links.lww.com/BRS/A438).

TABLE 2. Distribution of Radicular Pain Symptoms According to Level of Nerve Root Impingement

	Midlumbar Nerve Root Impingement L2, L3, or L4 (n = 25)			Low Lumbar Nerve Root Impingement L5 or S1 (n = 29)	
	L2 (n = 2)	L3 (n = 10)	L4 (n = 13)	L5 (n = 18)	S1 (n = 11)
Groin	1 (50.0%)	2 (20.0%)	0(0%)	1 (5.6%)	—
Thigh					
Anterior	2 (100%)	10 (100%)	5 (38.5%)	—	—
Lateral	—	1 (10.0%)	8 (61.5%)	8 (44.4%)	2 (18.2%)
Posterior	—	—	—	11 (61.1%)	8 (72.7%)
Knee					
Anterior	—	5 (50.0%)	3 (23.1%)	1 (5.6%)	—
Posterior	—	—	—	2 (11.1%)	1 (9.1%)
Calf					
Medial	—	—	1 (7.7%)	—	—
Anterior	—	3 (30.0%)	7 (53.9%)	1 (5.6%)	—
Lateral	—	—	3 (23.1%)	13 (72.2%)	2 (18.2%)
Posterior	—	—	—	8 (44.4%)	7 (63.6%)
Foot/ankle					
Medial	—	—	1 (7.7%)	1 (5.6%)	—
Dorsum	—	—	—	1 (5.6%)	1 (9.1%)
Lateral	—	—	—	2 (11.1%)	1 (9.1%)
Heel	—	—	—	1 (5.6%)	3 (27.3%)

Table 3. Performance Characteristics of Physical Examination Tests for the Diagnosis of Midlumbar Nerve Root Impingement and Low Lumbar Nerve Root Impingement on MRI*

Physical Examination Test	Midlumbar Nerve Root Impingement (L2, L3, L4)				Low Lumbar Nerve Root Impingement (L5, S1)			
	Sens. (95% CI)	Spec. (95% CI)	LR + (95% CI)	LR - (95% CI)	Sens. (95% CI)	Spec. (95% CI)	LR + (95% CI)	LR – (95% CI)
Provocative testing								
Straight Leg Raise (SLR)	16 (6–35)	31 (17–49)	**0.23 (0.09–0.59)**	**2.7 (1.5–4.8)**	69 (51–83)	84 (65–94)	**4.3 (1.7–11)**	**0.37 (0.21–0.65)**
Crossed Straight Leg Raise (CSLR)	4 (1–20)	93 (78–98)	0.58 (0.06–6.0)	1.0 (0.91–1.1)	7 (2–22)	96 (81–99)	1.7 (0.17–18)	0.97 (0.85–1.1)
Femoral Stretch Test (FST)	50 (31–69)	100 (88–100)	**NA (∞)**	**0.50 (0.34–0.75)**	0	50 (31–69)	**0**	**2.0 (1.3–3.0)**
Crossed Femoral Stretch Test (CFST)	5 (1–22)	100 (87–100)	**NA (∞)**	0.96 (0.87–1.1)	0	96 (78–99)	**0**	**1.1 (1.0–1.2)**
Motor testing								
Sit-to-Stand test	48 (30–67)	90 (74–96)	**4.6 (1.5–15)**	**0.58 (0.39–0.86)**	10 (4–26)	52 (34–70)	**0.22 (0.07–0.68)**	**1.7 (1.2–2.6)**
Heel Raise test	4 (1–20)	86 (69–95)	0.29 (0.04–2.4)	1.1 (0.94–1.3)	14 (6–31)	96 (81–99)	3.5 (0.41–29)	0.90 (0.76–1.1)
Heel Walk test	20 (9–39)	86 (69–95)	1.1 (0.33–3.7)	0.98 (0.80–1.2)	14 (6–31)	80 (61–91)	0.69 (0.21–2.3)	1.1 (0.84–1.4)
Great toe extensors	20 (9–39)	62 (44–77)	0.53 (0.21–1.3)	1.3 (0.91–1.8)	38 (23–56)	80 (61–91)	1.9 (0.76–4.7)	0.78 (0.55–1.1)
Hip abductors	4 (1–20)	82 (64–92)	0.22 (0.03–1.8)	**1.2 (0.97–1.4)**	18 (8–36)	96 (81–99)	4.5 (0.56–36)	**0.86 (0.71–1.0)**
Sensory testing								
Anterior thigh	8 (2–26)	96 (82–99)	2.3 (0.23–24)	0.95 (0.83–1.1)	4 (1–18)	92 (74–98)	0.43 (0.04–4.4)	1.1 (0.91–1.2)
Medial knee	17 (7–36)	96 (82–99)	4.7 (0.56–39)	0.86 (0.71–1.1)	4 (1–19)	83 (64–93)	0.21 (0.03–1.8)	**1.2 (1.0–1.4)**
Medial ankle	17 (7–36)	100 (88–100)	**NA (∞)**	**0.83 (0.70–1.0)**	0	83 (64–93)	**0**	**1.2 (1.0–1.4)**
Great toe	13 (5–32)	82 (64–92)	0.73 (0.20–2.7)	1.1 (0.84–1.3)	18 (8–36)	87 (68–96)	1.4 (0.37–5.1)	0.94 (0.75–1.2)
Lateral foot	8 (2–26)	79 (61–90)	0.39 (0.06–1.8)	1.2 (0.93–1.5)	21 (10–40)	92 (74–98)	2.6 (0.57–12)	0.86 (0.68–1.1)
Reflex testing								
Patella	28 (14–48)	100 (88–100)	**NA (∞)**	**0.72 (0.56–0.92)**	0	72 (52–86)	0	**1.4 (1.1–1.8)**
Ankle	4 (1–20)	71 (53–85)	**0.14 (0.02–1.0)**	**1.3 (1.1–1.7)**	29 (15–47)	96 (81–99)	**7.1 (0.96–53)**	**0.74 (0.58–0.95)**

*Physical examination tests with positive likelihood ratio 95% confidence intervals ≥1.0 or ≤1.0 in bold.

Sens. indicates sensitivity (%); Spec., specificity (%); LR+, positive likelihood ratio; LR–, negative likelihood ratio; CI, confidence interval.

The performance characteristics of selected physical examination tests for the diagnosis of level-specific nerve root impingement are presented in Table 4 (Supplemental Digital Content, Table 2, available at: http://links.lww.com/BRS/A438). Only physical examination tests with LR 95% CIs >1.0 for the diagnosis of level-specific nerve root impingement are presented because of space limitations. Six tests had positive LRs ≥5.0 for level-specific impingement: anterior thigh sensation for L2 impingement (12.5), the FST for L3 impingement (5.7), patellar reflex testing for L4 impingement (7.7), medial ankle sensation for L4 impingement (∞), CFST for L4 impingement (∞), and hip abductor strength for L5 impingement (10.6).

The performance characteristics of selected combinations of physical examination tests for the diagnosis of midlumbar nerve root impingement are presented in Table 5 (Supplemental Digital Content, Table 3, available at: http://links.lww.com/BRS/A438). Multiple test combinations demonstrated LRs ≥5.0, as well as increased sensitivity over individual tests; we have presented data for the most parsimonious combinations that either demonstrated the greatest improvements in sensitivity over individual tests while preserving high specificity (≥90%), or were more feasible to administer as screening tests in a clinical setting. The combination of the SLR and Achilles reflex testing for the diagnosis of low lumbar nerve root impingement demonstrated increased sensitivity (79%), but a decreased LR+ (4.0) compared with the individual tests because of decreased specificity (80%). In general, the progressive addition of more tests in combination for the diagnosis of either midlumbar or low lumbar impingement did not improve LRs + because increases in sensitivity were offset by decreased specificity (data not shown).

The performance characteristics of selected combinations of physical examination tests for the diagnosis of level-specific nerve root impingement are presented in Table 6 (Supplemental Digital Content, Table 4, available at: http://links.lww.com/BRS/A438). Only those combinations with LR+ point estimates ≥5.0 are presented. Combining tests substantially increased LRs, increased sensitivity, and maintained high specificity for L4 nerve root impingement only.

DISCUSSION

This study differs from prior studies of diagnostic accuracy in radicular pain in that it uses a specific reference standard of either a region of impingement (for example, midlumbar impingement at the L2, L3, or L4 levels) or impingement at a specific level (for example, the L5 nerve root), and not a reference standard of nerve root impingement or disc herni-

TABLE 4. Performance Characteristics of Physical Examination Tests for Level-Specific Nerve Root Impingement on MRI*

Physical Examination Test	Level of Nerve Root Impingement on MRI	Sens.(95% CI)	Spec.(95% CI)	Positive LR (95% CI)	Negative LR (95% CI)
Provocative testing	L3	70 (40–89)	88 (75–95)	**5.7 (2.3–4.4)**	0.34 (0.13–0.89)
Femoral Stretch Test (FST)	L4	9 (2–38)	100 (91–100)	**NA (∞)**	0.91 (0.75–1.1)
Crossed Femoral Stretch Test (CFST)	L5	67 (44–84)	67 (50–80)	2.0 (1.1–3.5)	0.50 (0.25–1.0)
	S1	73 (43–90)	63 (48–76)	2.0 (1.2–3.3)	0.43 (0.16–1.2)
Straight Leg Raise (SLR)					
Straight Leg Raise (SLR)					
Motor testing					
Sit-to-Stand test	L3	50 (24–76)	77 (63–87)	2.2 (1.0–5.0)	0.65 (0.34–1.23)
Sit-to-Stand test	L4	54 (29–77)	81 (66–90)	2.8 (1.2–6.1)	0.57 (0.31–1.1)
Great toe extensor	L5	61 (39–80)	86 (71–94)	4.4 (1.8–11)	0.45 (0.25–0.82)
Hip abductor	L5	29 (13–53)	97 (86–100)	**11 (1.3–84)**	0.73 (0.53–0.99)
Sensory testing					
Anterior thigh	L2	50 (10–91)	96 (87–99)	**13 (1.8–87)**	0.52 (0.13–2.1)
Medial ankle	L4	31 (13–57)	100 (91–100)	**NA (∞)**	0.69 (0.48–1.0)
Reflex testing					
Patella	L4	39 (18–65)	95 (84–99)	**7.7 (1.7–35)**	0.65 (0.42–1.0)
Achilles	L5	33 (16–56)	91 (78–97)	3.9 (1.1–14)	0.73 (0.52–1.0)

*Including only physical examination tests with positive likelihood ratio 95% confidence intervals ≥1.0 for level-specific nerve root impingement; tests with positive likelihood ration point estimates ≥5.0 in bold.
Sens. indicates sensitivity (%); Spec., specificity (%); LR +, positive likelihood ratio; LR–, negative likelihood ratio; CI, confidence interval.

TABLE 5. Selected Combinations of Physical Examination Tests for the Diagnosis of Midlumbar Nerve Root Impingement on MRI*

Physical Examination Test Combinations	Midlumbar Nerve Root Impingement (L2, L3, L4)			
	Sens. (95% CI)	Spec. (95% CI)	LR + (95% CI)	LR – (95% CI)
Two test combinations				
Patellar reflex or FST	60 (41–77)	100 (88–100)	NA (∞)	0.40 (0.25–0.65)
FST or Sit-to-stand test	72 (52–86)	90 (74–96)	7.0 (2.3–21)	0.31 (0.17–0.59)
Multiple test combinations				
Medial ankle sensation or patellar reflex or FST	64 (45–80)	100 (88–100)	NA (∞)	0.36 (0.21–0.61)
Medial ankle sensation or FST or Sit-to-Stand test	76 (57–89)	90 (74–96)	7.4 (2.5–22)	0.27 (0.13–0.54)
Medial ankle sensation or patellar reflex or FST or Sit-to-Stand test	80 (61–91)	90 (74–96)	7.7 (2.6–23)	0.22 (0.10–0.49)
Clinically useful combinations (able to be performed in seated position)				
Patellar reflex or Sit-to-Stand test	60 (41–77)	90 (74–96)	5.8 (1.9–18)	0.45 (0.27–0.73)
Medial ankle sensation or patellar reflex	40 (23–59)	100 (88–100)	NA (∞)	0.60 (0.44–0.83)
Medial ankle sensation or patellar reflex or Sit-to-Stand test	68 (48–83)	90 (74–96)	6.6 (2.2–20)	0.36 (0.20–0.64)

**Selected combinations with positive likelihood ratio point estimates ≥5.0, chosen for either improved sensitivity over individual tests while preserving high specificity, or for feasibility of administration in a clinical setting.*

Sens. indicates sensitivity (%); Spec., specificity (%); LR+, positive likelihood ratio; LR–, negative likelihood ratio; CI, confidence interval; FST, Femoral Stretch test.

ation at any level. This should not affect our understanding of the estimates of sensitivity (which is based on prevalence in the cases) in this study; however, it does have implications for estimates of specificity (which is based on prevalence in the non-cases). In prior studies, the question of relevance with regards to specificity has been "Does the finding occur in people who do not have nerve root impingement?"; in this study, the relevant question is "Does the finding occur in people who have nerve root impingement at a different level?".

Our findings demonstrate that individual physical examination tests may provide valuable clinical information that substantially alter the post-test probability that midlumbar nerve root impingement is present in patients presenting with lumbosacral radicular pain. However, although many physical examination tests are highly specific for midlumbar nerve root impingement, no single test is highly sensitive. The FST (sensitivity 50%) and the single-leg sit-to-stand test (sensitivity 48%) showed the highest sensitivities for the diagnosis of midlumbar nerve root impingement. Therefore, screening

TABLE 6. Combinations of Physical Examination Tests Chosen to Optimize Performance Characteristics for the Diagnosis of Level Specific Nerve Root Impingement on MRI*

Physical Examination Test Combinations	Sens. (95% CI)	Spec. (95% CI)	LR + (95% CI)	LR – (95% CI)
L4 nerve root				
Ankle sensation or patellar reflex	62 (36–82)	95 (84–99)	12 (3.0–51)	0.41 (0.20–0.81)
Ankle sensation or patellar reflex or CFST	62 (36–82)	95 (84–99)	13 (3.1–52)	0.40 (0.20–0.81)
CFST or patellar reflex	38 (18–64)	95 (84–99)	7.9 (1.7–36)	0.65 (0.42–1.0)
L5 nerve root				
Achilles reflex or hip abduction	61 (39–80)	89 (75–96)	5.5 (2.0–15)	0.44 (0.24–0.79)

**Combinations with positive likelihood ratio point estimates ≥5.0.*

Sens. indicates sensitivity (%); Spec., specificity (%); LR +, positive likelihood ratio; LR–, negative likelihood ratio; CI, confidence interval; FST, Femoral Stretch test; CFST, Crossed Femoral Stretch test.

with individual physical examination tests may not detect many cases of midlumbar nerve root impingement. This is expected because many cases of nerve root impingement do not have detectable impaired neurologic function or pain produced with nerve root tension signs. Certain combinations of physical examination tests increase sensitivity while maintaining high LRs, but other test combinations do not increase overall diagnostic accuracy, because of a trade-off of decreasing specificity with the addition of more tests.

The SLR showed moderate sensitivity (69%) for low lumbar nerve root impingement. Although the diagnostic accuracy of the SLR in our study showed decreased sensitivity compared with some prior studies of diagnostic accuracy in surgical populations with lumbar disc herniation,[2,23] it is consistent with prior findings in non-surgical populations with lumbar disc herniation.[24] This systematic difference between performance characteristics from surgical and nonsurgical studies has been previously noted.[25] We did not find any improvement in diagnostic accuracy with the addition of other tests in combination with the SLR.

Once midlumbar or low lumbar nerve root impingement is suspected, the physical examination may help to distinguish which specific nerve root is involved. A positive result on the FST, CFST, hip abduction strength, anterior thigh sensation, medial ankle sensation, and pa-tellar reflex testing indicates a moderate to large increase in the post-test probability of nerve root impingement at a specific level. A positive finding on 1 of these tests in practical terms also substantially decreases the probability that another root level is responsible for symptoms. Test combinations are only useful for the diagnosis of level-specific root impingement at the L4 level. From a clinical perspective, localization of nerve root involvement allows the physician to characterize relevant functional limitations at the baseline visit, and monitor for progression of deficits or new deficits at follow-up.

No prior study has examined the performance characteristics of the blinded physical examination in reference to midlumbar nerve root impingement. The results of this study, however, are consistent with the limited prior reports of physical examination characteristics in midlumbar disc herniation, when methodologic differences are taken into account. Although early reports using a surgical reference standard have stated a prevalence of 84% to 95% for a positive FST in known high lumbar disc herniation,[26–28] these estimates may have been affected by the well-documented overestimation of sensitivity observed because of spectrum bias.[25,29] Indeed, our findings of FST sensitivity are concordant with the results of a prior study using a reference standard of impingement on MRI, which reported a prevalence of 43% to 60% of the FST in nonsurgical patients with midlumbar radiculopathy.[8]

A notable exception to the trend of poor sensitivities with single tests for midlumbar nerve root impingement in our study was with the assessment of quadriceps strength using the single-leg sit-to-stand test, which demonstrated a sensitivity of 48% for the diagnosis of midlumbar nerve root impingement. This is likely explained by the superiority of the sit-to-stand test over manual muscle testing. A prior study of manual muscle testing found that a 50% loss of quadriceps strength was necessary to be detectable by manual testing.[30] The single-leg sit-to-stand test detects many cases of quadriceps weakness missed on manual testing.[8] In addition, the sit-to-stand test has other advantages that may favor its use as a screening test for midlumbar nerve root impingement. First, the sit-to-stand test not only provides highly specific diagnostic information, but also informs as to functional limitations. Poor performance on the sit-to-stand test may be associated with alterations in stair climbing ability, and impaired quadriceps strength is associated with poor balance and mobility.[31] Second, the sit-to-stand test has high reliability (k = 0.85),[8] which may be because of the objectiveness of the test result (ability to stand *vs*. inability). This means that the sit-to-stand test can be readily compared between different examiners. Furthermore, a deterioration in performance of the test is likely to represent a true change, and unlikely to represent the variability of the test itself. Finally, the sit-to-stand test can be efficiently integrated into the standard office examination and is easily performed after completion of the history, at the point when the patient needs to come to a standing position for the examination. Other tests described above may be used in combination with the sit-to-stand test to improve diagnostic accuracy and can also be readily performed with the patient seating. These combinations may have clinical utility as a screening tool in patients where the history, including location of radicular pain, suggests midlumbar nerve root impingement.

This study has other methodologic features that distinguish it from prior studies. We used nerve root impingement on MRI as the reference standard, in contrast to many prior studies, which used a surgical reference standard. The value of using an imaging reference standard in diagnostic studies of radiculopathy has been affirmed in recent publications, including an upcoming Co-chrane review.[29,32] Although the use of an imaging reference standard creates the potential for bias because of false-positive test results, it should be noted that prior imaging studies have found the prevalence of incidental nerve root impingement in asymptomatic subjects to be considerably lower than the prevalence of incidental disc herniation.[33,34] Furthermore, there are known disadvantages to the use of a surgical reference standard. A surgical reference standard is most affected by spectrum bias, in which patients studied have more severe manifestations of disease.[25] Verification bias, in which the reference standard is preferentially ordered as a result of the diagnostic test result, is likely to be more pronounced in studies using operative findings as the reference standard.[25,34] Finally, the accuracy and consistency of operative findings is unknown. It is unclear whether intraoperative observations reflect true pathologic states of the nerve root, or whether varying degrees of exposure of the operative field in different surgical techniques affect the validity of observations made during the procedure.[25]

The use of nerve root impingement as the anatomic reference of interest in this study, as opposed to lumbar disc herniation, is also worthy of further discussion. Although the majority of prior studies have used a disc herniation reference standard, a reference standard of nerve root impingement has greater validity from a conceptual standpoint. Given that the most sensitive and specific physical examination tests for the evaluation of radicular pain are conceptually based on the detection of neural tension (*i.e.*, SLR) or nerve root dysfunction (sen-sorimotor deficits), and are only indirectly related to the disc herniation itself, the appropriateness of a disc herniation reference standard is questionable.[29]

We chose to highlight LRs as a measure of diagnostic accuracy in this study. The reporting of LRs has long been advocated by experts and is a central recommendation in established guidelines for diagnostic studies.[21] The penetrance of LRs in the spine literature has been less than in other areas of clinical research. LRs are able to summarize in a single number how the initial assessment of the likelihood of disease ("pretest probability") is changed by a test result ("post-test probability").[35] Positive LRs describe how the likelihood of disease is changed by a positive test result, and negative LRs describe how the likelihood of disease is changed by a negative test result. LRs have the advantage over predictive values in that they are independent of disease prevalence and can be used to quickly calculate a discrete probability of disease that is contingent on the physician's pre-test suspicion of disease for any given patient.[20]

This study has limitations. First, this study was performed in a referral center for spine disorders. As such, our patient population has a higher prevalence of nerve root impingement and a different spectrum of severity than in primary care settings. Furthermore, although the neuroanatomic localization is a matter of great importance to the spine specialist, it may be beyond the scope of a standard diagnostic workup in a busy primary care practice. Our findings are likely not generalizable to the primary care setting. Second, only patients with lower extremity radiating pain were considered for entry into this study. Given that a history of lower extremity pain is itself sensitive for the diagnosis of sciatica,[36] the sensitivities and specificities determined in this study should be viewed in this context. However, it could be argued that these physical examination tests for the localization of nerve root impingement should only be performed in situations where there is some *a priori* suspicion of impingement, such as may be suggested by a history of sciatica. Therefore, our use of lower extremity pain as a criterion for inclusion is consistent with clinical practice in specialty spine clinics. A final limitation of this study is that the relatively small sample size studied resulted in wide CIs for some estimates of accuracy; future studies should include larger sample sizes.

In conclusion, this study demonstrates that the physical examination may yield useful diagnostic information for the detection of midlumbar nerve root impingement, low lumbar nerve root impingement, or level-specific nerve root impingement on MRI. As our population ages and the cost of care rises, the development and refinement of diagnostic tools that can be applied cheaply and broadly is of great importance.[37] Optimizing the accuracy of the physical examination is a goal worthy of further study, to improve the array of cost-effective diagnostic tools available in specialty spine care.

➢ Key Points

- ❑ Individual physical examination testing may provide clinical information that substantially alters the post-test probability that midlumbar, low lumbar, or level-specific nerve root impingement is present.
- ❑ Sensitivities of individual tests for the diagnosis of midlumbar nerve root impingement and level-specific nerve root impingement are modest.
- ❑ Test combinations may optimize sensitivity while maintaining high likelihood ratios for the diagnosis of midlumbar nerve root impingement, but not for low lumbar or level-specific nerve root impingement.
- ❑ Supplemental digital content is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site (www.spinejournal.com).

References

1. Blower PW. Neurologic patterns in unilateral sciatica. A prospective study of 100 new cases. *Spine* 1981;6:175–9.
2. Spangfort EV. The lumbar disc herniation. A computer-aided analysis of 2,504 operations. *Acta Orthop Scand Suppl* 1972;142:1–95.
3. Hsu K, Zucherman J, Shea W, et al. High lumbar disc degeneration. Incidence and etiology. *Spine* 1990;15:679–82.
4. Pearson AM, Blood EA, Frymoyer JW, et al. SPORT lumbar intervertebral disk herniation and back pain: does treatment, location, or morphology matter? *Spine* 2008;33:428–35.
5. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. *JAMA* 2006;296:2451–9.
6. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *JAMA* 2006;296:2441–50.
7. Bressler HB, Keyes WJ, Rochon PA, et al. The prevalence of low back pain in the elderly. A systematic review of the literature. *Spine* 1999;24: 1813–9.
8. Rainville J, Jouve C, Finno M, et al. Comparison of four tests of quadriceps strength in L3 or L4 radiculopathies. *Spine* 2003;28:2466–71.
9. Nadler SF, Campagnolo DI, Tomaio AC, et al. High lumbar disc: diagnostic and treatment dilemma. *Am J Phys Med Rehabil* 1998;77:538–44.
10. Battie MC, Videman T, Parent E. Lumbar disc degeneration: epidemiology and genetic influences. *Spine* 2004;29:2679–90.
11. Goddard MD, Reid JD. Movements induced by straight leg raising in the lumbo-sacral roots, nerves and plexus, and in the intrapelvic section of the sciatic nerve. *J Neurol Neurosurg Psychiatry* 1965;28:12–8.
12. Katz JN, Chang LC, Sangha O, et al. Can comorbidity be measured by questionnaire rather than medical record review? *Med Care* 1996;34: 73–84.
13. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine* 2000;25: 2940–52; discussion 2952.
14. Collins SL, Moore RA, McQuay HJ. The visual analogue pain intensity scale: what is moderate pain in millimetres? *Pain* 1997;72:95–7.
15. Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? *JAMA* 1992;268:760–5.
16. Brain. Introduction. In: O'Brien MD, ed. *Aids to the Examination of the Peripheral Nervous System*. 4th ed. London: WB Saunders for The Guarantors of Brain; 2000:1–2.

17. Standard Neurological Classification of Spinal Cord Injury. 2006. Accessed April 4, 2009. Available at: http://www.asia-spinalinjury.org/publications/ 2006_Classif_worksheet.pdf.
18. Manschot S, van Passel L, Buskens E, et al. Mayo and NINDS scales for assessment of tendon reflexes: between observer agreement and implications for communication. *J Neurol Neurosurg Psychiatry* 1998;64: 253–5.
19. Pfirrmann CW, Dora C, Schmid MR, et al. MR image-based grading of lumbar nerve root compromise due to disk herniation: reliability study with surgical correlation. *Radiology* 2004;230:583– 8.
20. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ* 2004;329: 168–9.
21. Simel DL, Rennie D, Bossuyt PM. The STARD statement for reporting diagnostic accuracy studies: application to the history and physical examination. *J Gen Intern Med* 2008;23:768–74.
22. Pinksy LE, Wipf JE, Ramsey SD. Evidence-based medicine glossary. In: Geyman JP, Deyo RA, Ramsey SD, eds. *Evidence-Based Clinical Practice: Concepts and Approaches*. Boston, MA: Butterworth-Heinemann; 2000: 165–72.
23. Uden A, Landin LA. Pain drawing and myelography in sciatic pain. *Clin Orthop Relat Res* 1987;216:124–30.
24. Haldeman S, Shouka M, Robboy S. Computed tomography, electrodiagnos-tic and clinical findings in chronic workers' compensation patients with back and leg pain. *Spine* 1988;13:345–50.
25. 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.
26. Abdullah AF, Wolber PG, Warfield JR, et al. Surgical management of extreme lateral lumbar disc herniations: review of 138 cases. *Neurosurgery* 1988;22:648–53.
27. Christodoulides AN. Ipsilateral sciatica on femoral nerve stretch test is pathognomonic of an L4/5 disc protrusion. *J Bone Joint Surg Br* 1989; 71:88–9.
28. Porchet F, Fankhauser H, de Tribolet N. Extreme lateral lumbar disc herni-ation: clinical presentation in 178 patients. *Acta Neurochir (Wien)* 1994; 127:203–9.
29. Vroomen PC, de Krom MC, Wilmink JT, et al. Diagnostic value of history and physical examination in patients suspected of lumbosacral nerve root compression. *J Neurol Neurosurg Psychiatry* 2002;72:630–4.
30. Beasley WC. Quantitative muscle testing: principles and applications to research and clinical services. *Arch Phys Med Rehabil* 1961;42:398 – 425.
31. Visser M, Goodpaster BH, Kritchevsky SB, et al. Muscle mass, muscle strength, and muscle fat infiltration as predictors of incident mobility limitations in well-functioning older persons. *J Gerontol A Biol Sci Med Sci* 2005;60:324–33.
32. van den Windt DAWM, Simons E, Riphagen I, et al. Physical examination for lumbar radiculopathy due to disc herniation in patients with low-back pain (Protocol). *Cochrane Database Syst Rev* 2010 Feb 17;2:CD007431.
33. Boden SD, Davis DO, Dina TS, et al. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg* 1990;72:403–8.
34. Weishaupt D, Zanetti M, Hodler J, et al. MR imaging of the lumbar spine: prevalence of intervertebral disk extrusion and sequestration, nerve root compression, end plate abnormalities, and osteoarthritis of the facet joints in asymptomatic volunteers. *Radiology* 1998;209:661–6.
35. Dawson B, Trapp RG. Methods of evidence-based medicine and decision analysis. In: *Basic & Clinical Biostatistics*. 4th ed. New York: Lange Medical Books/McGraw-Hill; 2004:326.
36. Kerr RS, Cadoux-Hudson TA, Adams CB. The value of accurate clinical assessment in the surgical management of the lumbar disc protrusion. *J Neu-rol Neurosurg Psychiatry* 1988;51:169–73.
37. Weiner DK, Kim YS, Bonino P, et al. Low back pain in older adults: are we utilizing healthcare resources wisely? *Pain Med* 2006;7:143–50.