ULITARY OF ELECTRODIAGNOSTIC TESTING IN EVALUATING PATIENTS WITH LUMBOSACRAL RADICULOPATHY: AN EVIDENCE-BASED REVIEW

S. CHARLES CHO, MD, 1 MARK A. FERRANTE, MD, 2 KERRY H. LEVIN, MD, 3 ROBERT L. HARMON, MD, MS, 4 and YUEN T. SO, MD, PhD 1

1 Stanford University School of Medicine, Stanford, California, USA
2 Ohio State University School of Medicine, Columbus, Ohio, USA
3 Cleveland Clinic, Cleveland, Ohio, USA
4 Pain Management and Rehabilitation Specialists, PC and Medical College of Georgia, Augusta, Georgia, USA

ABSTRACT: This is an evidence-based review of electrodiagnostic (EDX) testing of patients with suspected lumbosacral radiculopathy to determine its utility in diagnosis and prognosis. Literature searches were performed to identify articles applying EDX techniques to patients with suspected lumbosacral radiculopathy. From the 355 articles initially discovered, 119 articles describing nerve conduction studies, electromyography (EMG), or evoked potentials in adequate detail were reviewed further. Fifty-three studies met inclusion criteria and were graded using predetermined criteria for classification of evidence for diagnostic studies. Two class II, 7 class III, and 34 class IV studies described the diagnostic use of EDX. One class II and three class III articles described H-reflexes with acceptable statistical significance for use in the diagnosis and confirmation of suspected S1 lumbosacral radiculopathy. Two class II and two class III studies demonstrated a range of sensitivities for use of muscle paraspinal mapping. Two class II studies demonstrated the utility of peripheral myotomal limb electromyography in radiculopathies.

The diagnosis of lumbosacral radiculopathy is based on clinical history and examination, imaging studies, and electrodiagnostic (EDX) testing. Converging lines of evidence localize the lesion, establish the diagnostic certainty, and assist with prognostication. In those cases with negative imaging findings or atypical clinical presentations, EDX is especially useful; whereas, in those with abnormal imaging findings, it serves a complementary diagnostic role. A correct diagnosis of lumbosacral radiculopathy is important for implementation of timely and appropriate treatments.

The primary difficulty in this systematic review was the lack of an established reference standard for the diagnosis of lumbosacral radiculopathy. This lack of a standard makes comparison of reported sensitivity and specificity impossible. Structural abnormalities by imaging modalities had high rates of false positives, 1–4 and occasional false negatives. 5, 6 Surgical visualization also did not consistently translate into clinical symptoms or physiological changes. 5 In addition, the clinical presentation itself may be inconclusive. For example, injuries affecting the sacroiliac and zygapophysial joints, ligaments, muscles, and the peripheral disk annulus may cause referred pain suggestive of radiculopathy. 7–9

The Lumbosacral Radiculopathy Task Force was charged by the American Association of Neuromuscular and Electrodiagnostic Medicine (AANEM) to perform a systematic review of the available literature to evaluate the utility of EDX studies in the diagnosis of lumbosacral radiculopathies. Task force members were selected on the basis of their EDX and methodological expertise. This practice parameter describes the results and recommendations of this systematic review.

METHODS

A search of the United States National Library of Medicine’s MEDLINE database was performed in October 2003 for articles in English pertaining to humans, using the medical subject headings (MeSHs) of lumbosacral radiculopathy, radiculitis, low back pain, sciatica, herniated disk, root compression (limited to the lumbosacral region), electrophysiology, neurophysiology, nerve conduction studies, electromyography, and spinal nerve root compression. A second search was conducted in October 2005, a third in January 2006, and a fourth in August 2006, using the expanded search criteria based on the original search, additional MeSH terms from the final articles, bibliographies, relevant textbooks, and relevant articles presented.
In the literature review, the authors did not find any consensus for a reference case definition of lumbosacral radiculopathy. Case definitions considered potentially acceptable included an appropriate combination of clinical findings, imaging studies demonstrating a structural lesion and surgical confirmation of root pathology, and a consensus diagnosis based on unanimous independent clinical opinions of a team consisting of a physiatrist, a radiologist, and a neurosurgeon.

RESULTS

Nine articles providing Class II or Class III evidence were identified. Three articles studied the use of F-wave latencies and evoked potentials, and four articles addressed H-reflexes and EMG. Four studies applied the EDX test in a blinded evaluation. The outcome assessments in the remaining articles were not blinded, but were classified as Class III as the EDX measures were objective (i.e., F-wave, H-reflex, and SEP studies, or EMG interpretation limited to the presence or absence of fibrillations potentials and positive sharp waves). Class IV studies are not considered further in this review.

Study Characteristics. The study design and cohort assembly method varied. Two studies were prospective, masked, and controlled; five studies were prospective and case-controlled; and three studies were retrospective and case-controlled. All of these studies, except two, have a narrow spectrum of patients who had lumbosacral radiculopathy, both clinically and by various other tests. Conversely, the two studies that employed a broad spectrum of patients included asymptomatic patients and those with low back pain who did not have spinal stenosis. The number of patients with lumbosacral radiculopathy in each study ranged from 16 to 206. Mean ages ranged from 18 to 80 years, although one study did not state the inclusive ages.

In two Class II prospective, controlled studies, the EDX physician was blinded to the clinical diagnosis. In one study, subjects were designated normal control, low back pain without spinal stenosis, or spinal stenosis. The reference standard was an independent consensus agreement of the physiatrist, neuroradiologist, and neurosurgeon for the presence of spinal stenosis. However, the definition of spinal stenosis was not provided, and no attempt was made to correlate the EMG findings to the level or the side of radiologic or clinical abnormality. Another Class II study utilized data from Knutsson’s 1961 article to evaluate the lumbosacral EMG screen. The clinical data were presented in adequate detail to permit this reformulation. The data compared EMG, myelogram, and...
Clinical symptoms to the intraoperative findings of a narrow spectrum of 206 patients who had sciatica. One hundred eighty-five patients had intraoperative confirmation of root compromise. The study group included subjects with sciatica and intraoperative abnormalities, whereas the control group consisted of patients with sciatica and no abnormal intraoperative findings. The EMG assessments were blinded, performed prospectively, and compared with an objective “gold standard.”

The spectrum of clinical presentations varied. Two studies mandated dermatomal sensory deficits and an L5 or S1 pattern of weakness (dorsiflexion or plantarflexion), whereas another allowed clinical symptoms to include or exclude sensory paresthesias in a dermatomal distribution. Clinical presentations included back pain, pain radiation, reflex abnormalities, strength deficits, and ambulation ability in both the radiculopathy cohort and the group with low back pain that did not have a radiculopathy. Other symptoms were sciatica over a certain amount of time, and 60% of patients had a reduced Achilles tendon reflex, 17% had a reduced patellar reflex, 61% had weakness of great toe extension, 30% had reduced sensation, 40% had muscular atrophy, and 95% had positive straight leg raises. Most of the studies used radicular pain as an initial screen; however, one article did not explicitly state the clinical features of presentation. Four articles included diminished ankle jerks and positive straight leg raises as part of the clinical criteria. Three studies did not detail the clinical examination.

Two studies explicitly excluded patients with previous surgery. Another two studies implied exclusion because surgery was part of the reference standard, but the authors did not explicitly state the surgery exclusion. Although the spectrum of disease and the associated presentations varied, the common feature among the prospective patient assembly was radiating lower back pain. Four groups included surgical observations as part of the “gold standard.” Others used some form of EDX, and most implemented various imaging studies. A recent study implemented a unanimous consensus among the physiatrist, neuroradiologist, and neurosurgeon as the case definition of spinal stenosis.

**EMG/Paraspinal Mapping.** Several studies investigated the use of limb myotomal EMG and PM, using abnormal spontaneous activities as the definition of abnormalities (Table 1). A Class III study compared PM with imaging (either computed tomography or magnetic resonance imaging) in 43 patients and extremity EMG in 110 patients. The patient groups were classified by the degree of certainty of lumbosacral radiculopathy. Imaging was classified as normal and abnormal with abnormalities classified as possible, probable, single level, or multilevel. Extremity EMG was given in mean and standard deviation scores of abnormal spontaneous activity, motor unit potential changes, and composite EMG rather than raw data. Comparing only the certain (and excluding the possible and probable) single-level abnormality imaging findings to PM, the sensitivity was 63%, specificity 92%, positive predictive value (PPV) 87%, and negative predictive value (NPV) 75%; the values were 71%, 92%, 83%, and 85%, respectively, for the multilevel imaging abnormalities. Compared with imaging, the false-positive rate was 8% (1 of 13) and the false-negative rate was 33% (6 of 18). The false-negative rate decreased to 5% with the addition of lower extremity EMG. All 7 patients with an isolated S1 radiculopathy had normal PM studies.

In another Class III study, PM was compared with imaging, extremity EMG, physical examination, and combined EMG/imaging as separate reference standards. Of note, the authors calculated different values than what were recorded, and they presented those calculations. Using imaging as the reference standard the sensitivity of PM was 66% and the specificity was 92%; using extremity EMG as the reference standard, these values were 50% and 85%, respectively; using physical examination as the gold standard, they were 54% and 69%, respectively; and using combined EMG/imaging as the gold standard, they were 58% and 90%, respectively.

A Class II study documented PM of four or more muscles. Abnormalities defined as membrane instability on one side demonstrated sensitivity of 30.4% and specificity of 100%, meeting statistical significance compared with control subjects (P = 0.004). Additional EMG examination in an overlapping myotomal distribution in limb muscles (tensor fascia lata, vastus medialis, tibialis anterior, extensor hallucis longus, medial gastrocnemius) had sensitivity of 47.8% and specificity of 87.5% (P = 0.008), when only fibrillation potentials were accepted as abnormal. When a combination of any needle EMG abnormalities was used (fibrillation potential, typical motor unit potential amplitude, number of polyphasic motor unit potential, motor unit potential recruitment), the sensitivity increased to 79.2%, and specificity fell to 50% (P = 0.035).

In another Class II study, a four- and five-muscle extremity EMG screen that included PM had sensitivity in the range of 89–92%. Four- and five-muscle screens without PM were not as sensitive (77–90%). Specificity could not be calculated because the data did not include false-positive or true-negative values.
H-Reflex. One Class II\textsuperscript{13} and two Class III studies\textsuperscript{14,20} investigated the tibial nerve H-reflex in the diagnosis of lumbosacral radiculopathy (Table 2). The diagnostic sensitivity and specificity varied widely. One study\textsuperscript{8,14} noted 100% sensitivity and specificity, whereas others\textsuperscript{13,14,20} reported sensitivity of 51%, specificity of 91%, PPV of 64%, and NPV of 84% in the S1 radiculopathy group. The sensitivity dropped to 6% in the L5 group. The Class II study\textsuperscript{13} did not specify the root level of abnormality or describe stimulating or recording techniques. Using H-wave absence (ipsilateral or bilateral) or asymmetry as markers of abnormality, sensitivities were 36.4% and 18.2%, and specificities were 91.3% and 100%, respectively.

F-Wave Studies. Two Class III studies used peroneal nerve (L5 radiculopathy) and posterior tibial nerve (S1 radiculopathy) F-wave latency and interside latency differences as markers of abnormality and compared these with normal values (Table 3).\textsuperscript{14,17} One used the abductor digiti minimi for posterior tibial and extensor digitorum brevis for peroneal recording.\textsuperscript{17} The sensitivity was 65% using the peroneal nerve, and 56% using the tibial nerve. The data did not allow the specificity, PPV, and NPV to be calculated. The other Class III study did not state the muscles used and reported a combined peroneal/tibial cumulative sensitivity of 25%, specificity of 62%, PPV of 57%, and NPV of 29% in detecting lumbosacral radiculopathies.\textsuperscript{14} A Class II study\textsuperscript{13} mentioned peroneal F studies, but did not report tibial studies. Markers of abnormality included F-wave absence (ipsilateral or bilateral) or asymmetry. Asymmetry was not defined. Sensitivity was 4.8%, and specificity was 95.5%.

Motor Evoked Potential with Root Stimulation. One Class III study\textsuperscript{21} tested motor evoked potentials (MEPs) stimulating with surface electrodes over the midline of the lumbosacral region and recording from the tibialis anterior (L5) and soleus (S1) muscles. The MEP latency was considered abnormal when it was 3 standard deviations greater than the mean of normal control subjects, or when an interside latency difference exceeded 0.8 ms. The L5 MEP latencies were 72% sensitive and 100% specific in detecting lumbar radiculopathies, with a PPV of 100% and an NPV of 83%. Similarly, the S1 latencies were 66.7% sensitive and 100% specific in detecting sacral radiculopathies, with a PPV of 100% and an NPV of 74%.

Dermatomal/Segmental SEP. There were two Class III studies using dermatomal SEP (Table 4).\textsuperscript{14,16} One study used a site of stimulation 6 cm above the medial malleolus for L4, the medial side of the second metatarsal bone for L5, and the lateral side...
of the fifth metatarsal bone for S1.\textsuperscript{14} Of the 5 patients with herniation and an abnormal SEP, 1 had an S1 lesion coupled to an L5 SEP abnormality, 1 had an L5 lesion coupled to an L4 SEP abnormality, and the other 3 showed concordance with the actual localization. The sensitivity was 29\%, specificity 67\%, PPV 63\%, and NPV 33\%. In the other study,\textsuperscript{14,16} the sural and superficial peroneal nerves were stimulated first, and then a site at the superomedial margin of the first metatarsophalangeal joint was employed for L5. A site at the midsection of the fifth metatarsal bone on the lateral foot was used for S1 dermatomal SEP. The L5 dermatomal SEP was 50\% sensitive and 97.7\% specific at both the 90\% and 95\% confidence intervals (CIs). The S1 dermatomal SEP was 20\% sensitive and 97\% specific at 90\% CI, and 10\% sensitive and 97.9\% specific at 95\% CI. The superficial peroneal nerve for L5 radiculopathy of 70\% and 60\%, respectively, at 90\% and 95\% CI. Of note, there was a high false-negative rate. Twenty percent of the patients with a confirmed L5 radiculopathy had an S1 dermatomal abnormality, and 10\% of the patients with a confirmed S1 radiculopathy had an L5 dermatomal abnormality.

CONCLUSIONS

1: In patients with suspected lumbosacral radiculopathy, the following EDX studies probably aid the clinical diagnosis:

(a) Peripheral limb EMG (Class II evidence, Level B recommendation).

(b) PM with needle EMG in lumbar radiculopathy (Class II evidence, Level B recommendation).

(c) H-reflex in S1 radiculopathy (Class II and III evidence, Level C recommendation).

2: Evidence suggests a low sensitivity of peroneal and posterior tibial F-waves (Class II and III evidence, Level C recommendation).

3: There is inadequate evidence to reach a conclusion on the utility of the following EDX studies:

(a) Dermatomal/segmental SEP of the L5 or S1 dermatomes (Class III evidence, Level C recommendation)

\[\text{Table 2. Design characteristics and outcomes in nerve conduction studies for lumbosacral radiculopathy.}\]

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Test</th>
<th>Class</th>
<th>Blind</th>
<th>Prospec/retro</th>
<th>Cohort size</th>
<th>*Gold Standard* Dx</th>
<th>Test screening yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albeck (2000)</td>
<td>H-reflex (S1 radiculopathy)</td>
<td>3 (9 control subjects)</td>
<td>Yes</td>
<td>Prospective</td>
<td>16</td>
<td>Surgical, secondary other test SEP, F-wave, H-reflex, NCS, EMG</td>
<td>100 100 100 100</td>
</tr>
<tr>
<td>Haig (2005)</td>
<td>H-reflex</td>
<td>2 (30 control subjects, 30 pain but no MRI)</td>
<td>Yes</td>
<td>Prospective</td>
<td>60</td>
<td>History, exam, MRI; 3 consensus Dxs</td>
<td>33.3 91.7 80 57.9</td>
</tr>
<tr>
<td>Marin (1995)</td>
<td>H-reflex</td>
<td>3 (53 normal control subjects)</td>
<td>No</td>
<td>Prospective</td>
<td>17 L5, 18 S1</td>
<td>EMG, F-wave, SEP</td>
<td>50 91</td>
</tr>
</tbody>
</table>

Dx, diagnosis; EMG, electromyography; PM, paraspinal mapping; MRI, magnetic resonance imaging; NCS, nerve conduction study; NPV, negative predictive value; PPV, positive predictive value; SEP, somatosensory evoked potential.

\[\text{Table 3. Design characteristics and outcomes in nerve conduction studies for lumbosacral radiculopathy.}\]

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Test</th>
<th>Class</th>
<th>Blind</th>
<th>Prospec/retro</th>
<th>Cohort size</th>
<th>*Gold standard* Dx</th>
<th>Test screening yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albeck (2000)</td>
<td>F-waves</td>
<td>3 (9 control subjects)</td>
<td>Yes</td>
<td>Prospective</td>
<td>16</td>
<td>Surgical, secondary other test SEP, F-wave, H-reflex, NCS, EMG</td>
<td>25 62 57 29</td>
</tr>
<tr>
<td>Eisen (1977)</td>
<td>F-wave latency</td>
<td>3</td>
<td>No</td>
<td>Prospective</td>
<td>60 normal, 25 subjects</td>
<td>Myelography; 6 surgically proven History, exam, MRI; 3 consensus Dxs</td>
<td>50</td>
</tr>
<tr>
<td>Haig (2005)</td>
<td>F-waves absent (either or both sides)</td>
<td>2 (30 control subjects, 30 pain but no MRI)</td>
<td>Yes</td>
<td>Prospective</td>
<td>60</td>
<td></td>
<td>4.2 95.8 50 50</td>
</tr>
</tbody>
</table>

Dx, diagnosis; EMG, electromyography; MRI, magnetic resonance imaging; NCS, nerve conduction study; NPV, negative predictive value; PPV, positive predictive value; SEP, somatosensory evoked potential.
recommendation).
(b) PM with needle EMG in sacral radiculopathy (one small Class II study, Level U).
(c) MEP with root stimulation in making an independent diagnosis of lumbosacral radiculopathy (Class III evidence, Level U).

**DISCUSSION**

The available evidence is limited by the lack of a universally accepted case definition of lumbosacral radiculopathy. Furthermore, none of the studies presented methodological data in adequate detail to allow unbiased comparison of an objective reference standard to the EDX diagnosis. Also, with the exception of one study, a narrow spectrum of patients were used.

In the setting of suspected lumbosacral radiculopathy, EDX testing is often applied as part of an evaluation that also includes clinical history, physical examination, and imaging study. It would be useful to determine the independent contribution of EDX to the diagnosis. The studies reviewed did not address the optimal combination or sequence of testing in the overall evaluation process. Finally, none of the reviewed studies addressed the utility of EDX in prognosticating outcome or response to treatments.

**RECOMMENDATIONS FOR FUTURE RESEARCH**

Future studies should eliminate potential biases and provide data sufficient to determine the independent contribution of EDX techniques used in diagnosing suspected lumbosacral radiculopathy. These elements are recommended for diagnostic accuracy:

1. The study design should be a prospective cohort survey.
2. The cohort should include a wide spectrum of patients with alternative diagnoses, allowing the EDX to independently assess and determine the presence of radiculopathy.
3. A consensus-based reference standard (gold standard) of lumbosacral radiculopathy must be developed for research purposes.
4. Studies should employ a wide spectrum of patients with alternative diagnoses and with varying degrees of severity of radiculopathy.
5. All patients enrolled should complete the EDX. The presence of radiculopathy should be measured, and the level and side documented.
6. An investigator who is unaware of the EDX results should determine the final diagnosis.
7. Studies may be performed to assess the utility of EDX in combination with or in isolation from alternative diagnostic studies and the sequence of studies and study combinations that provide the highest yield for identifying lumbosacral radiculopathies.

**DEFINITIONS FOR STRENGTH OF EVIDENCE**

**Class I**: Evidence provided by a prospective study in a broad spectrum of persons with the suspected condition, using a reference (gold) standard for...
case definition, where the test is applied in a blinded evaluation, and enabling the assessment of appropriate tests of diagnostic accuracy. All patients undergoing the diagnostic test have the presence or absence of the disease determined.

Class II: Evidence provided by a prospective study of a narrow spectrum of patients with the suspected condition, or a well-designed retrospective study of a broad spectrum of persons with an established condition (by gold standard) compared with a broad spectrum of controls, where testing is applied in a blinded evaluation, and enabling the assessment of appropriate tests of diagnostic accuracy.

Class III: Evidence provided by a retrospective study where either persons with the established condition or controls are of a narrow spectrum, and where the reference standard, if not objective, is applied by someone other than the person who performed the test.

Class IV: Any design where test is not applied in an independent evaluation OR evidence provided by expert opinion alone or in descriptive case series (without control subjects).

Rating System for Strength of Recommendations

- **Level A**: Established as effective, ineffective, or harmful (or established as useful/predictive or not useful/predictive) for the given condition in the specified population.
- **Level B**: Probably effective (or probably useful/predictive or not useful/predictive) for the given condition in the specified population.
- **Level C**: Possibly effective (or possibly useful/predictive or not useful/predictive) for the given condition in the specified population.
- **Level U**: Data inadequate or conflicting. Given current knowledge, the test is unproven.

### REFERENCES


This guideline is greater than 5 years old. Every five years, an interim literature search is performed and the guideline reviewed. While new studies have been published since this guideline was last reviewed, the Practice Issue Review Panel Committee of the AANEM has determined that these studies are not sufficient to mandate a revision of this guideline at the present time. The information contained in this guideline is current knowledge, the test is unproven.