



Critically Re-Evaluating a Common Technique: Accuracy, Reliability, and Confirmation Bias of EMG

Citation

Narayanaswami, Pushpa, Thomas Geisbush, Lyell Jones, Michael Weiss, Tahseen Mozaffar, Gary Gronseth, and Seward B. Rutkove. 2015. "Critically Re-Evaluating a Common Technique." *Neurology* 86 (3) (December 23): 218–223. doi:10.1212/wnl.0000000000002292.

Published Version

doi:10.1212/WNL.0000000000002292

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:33974334>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Critically re-evaluating a common technique

Accuracy, reliability, and confirmation bias of EMG



Pushpa Narayanaswami, MD
Thomas Geisbush, BA
Lyell Jones, MD
Michael Weiss, MD
Tahseen Mozaffar, MD
Gary Gronseth, MD
Seward B. Rutkove, MD

Correspondence to
Dr. Narayanaswami:
pnarayan@bidmc.harvard.edu

ABSTRACT

Objectives: (1) To assess the diagnostic accuracy of EMG in radiculopathy. (2) To evaluate the intrarater reliability and interrater reliability of EMG in radiculopathy. (3) To assess the presence of confirmation bias in EMG.

Methods: Three experienced academic electromyographers interpreted 3 compact discs with 20 EMG videos (10 normal, 10 radiculopathy) in a blinded, standardized fashion without information regarding the nature of the study. The EMGs were interpreted 3 times (discs A, B, C) 1 month apart. Clinical information was provided only with disc C. Intrarater reliability was calculated by comparing interpretations in discs A and B, interrater reliability by comparing interpretation between reviewers. Confirmation bias was estimated by the difference in correct interpretations when clinical information was provided.

Results: Sensitivity was similar to previous reports (77%, confidence interval [CI] 63%–90%); specificity was 71%, CI 56%–85%. Intrarater reliability was good (κ 0.61, 95% CI 0.41–0.81); interrater reliability was lower (κ 0.53, CI 0.35–0.71). There was no substantial confirmation bias when clinical information was provided (absolute difference in correct responses 2.2%, CI –13.3% to 17.7%); the study lacked precision to exclude moderate confirmation bias.

Conclusions: This study supports that (1) serial EMG studies should be performed by the same electromyographer since intrarater reliability is better than interrater reliability; (2) knowledge of clinical information does not bias EMG interpretation substantially; (3) EMG has moderate diagnostic accuracy for radiculopathy with modest specificity and electromyographers should exercise caution interpreting mild abnormalities.

Classification of evidence: This study provides Class III evidence that EMG has moderate diagnostic accuracy and specificity for radiculopathy. *Neurology*® 2016;86:218–223

GLOSSARY

CD = compact disc; CI = confidence interval.

EMG is frequently used to assess radiculopathy due to degenerative disc disease. Several factors influence the diagnostic accuracy of EMG. Reliability measures the extent to which EMG provides consistent results, when repeated by the same electromyographer at different time points (intrarater reliability), or by different electromyographers (interrater reliability). While there are no studies evaluating EMG intrarater reliability, interrater reliability has been previously reported to be 60.5% among faculty-level electromyographers.¹

Validity reflects the ability of EMG to identify radiculopathy accurately, and is often measured as sensitivity and specificity. The sensitivity of EMG for the diagnosis of radiculopathy ranges from 36% to 64% in patients with only pain,^{2–4} and 51%–86% in patients with an abnormal examination.^{4–9} Studies of specificity are limited. One study reports 54%–58% specificity for any EMG abnormality.¹⁰

Finally, EMG interpretation may be influenced by clinical information; this is termed confirmation bias. For example, without history or examination findings suggestive of

Supplemental data
at Neurology.org

From the Department of Neurology (P.N., T.G., S.B.R.), Beth Israel Deaconess Medical Center/Harvard Medical School, Boston, MA; the Department of Neurology (L.J.), Mayo Clinic, Rochester, MN; the Department of Neurology (M.W.), University of Washington, Seattle; the Department of Neurology (T.M.), University of California, Irvine; and the Department of Neurology (G.G.), University of Kansas Medical Center, Kansas City.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

radiculopathy, mild motor unit polyphasia may be interpreted as normal; if the history/examination suggests radiculopathy, the same finding may be interpreted as mild chronic radiculopathy. To our knowledge, confirmation bias in EMG has not been studied.

The aims of this study were to assess the accuracy of EMG to diagnose radiculopathy (Class III); evaluate the intrarater reliability (reproducibility) and interrater reliability (agreement between reviewers) of EMG in radiculopathy; and assess the presence of confirmation bias in EMG.

METHODS Standard protocol approvals, registrations, and patient consents. This study was approved by the Committee on Clinical Investigations, BIDMC, Boston, and all participants signed approved written informed consent forms.

Study design. Consecutive participants with clinically diagnosed cervical/lumbosacral radiculopathy were recruited in the EMG laboratory; healthy controls were recruited via advertisement. The reference standard for diagnosis of radiculopathy was based on history and examination, defined by the presence of 2 or more of the following: neck/low back pain, radicular pain or sensory symptoms, radicular sensory loss, myotomal weakness, and myotomal reflex abnormality. Participants with clinical suspicion for common entrapment neuropathies such as carpal tunnel syndrome or ulnar neuropathy at the elbow were excluded. If available, radiologic evidence for degenerative disc disease at the symptomatic level was used as supporting evidence.

The severity of radiculopathy was categorized clinically as follows: mild, neck or back pain with radicular pain; moderate, radicular pain and one of the following: radicular sensory loss, reflex change, myotomal weakness; moderate to severe: radicular pain and 2 of radicular sensory loss, reflex change, and myotomal weakness; severe: radicular pain with radicular sensory loss, reflex change, and myotomal weakness.

Concentric needle EMG was performed and video-recorded in 10 participants with clinically diagnosed cervical or lumbosacral radiculopathy and in 10 healthy controls. Two investigators, S.B.R. and P.N., both fellowship-trained, board-certified electromyographers with over 15 years of experience, performed all EMGs. A standard EMG root screen was performed in all participants, comprising the following muscles: upper extremity: deltoid, biceps, triceps, brachioradialis, flexor carpi radialis, extensor digitorum communis, extensor indicis, first dorsal interosseous; lower extremity: tibialis anterior, medial gastrocnemius, vastus medialis, vastus lateralis, biceps femoris (long head), tensor fasciae lata. Radiculopathy was diagnosed when 2 or more muscles innervated by the same root and different peripheral nerves showed active denervation or chronic reinnervation.

Three compact discs (CDs), CD-A, B, and C, were created, each with 24 EMG videos lasting 4–6 minutes, with all identifiers removed. Each CD consisted of 10 EMG recordings that were consistent with radiculopathy and 10 normal recordings, as determined by the investigator who performed the EMG (S.B.R., P.N.). Additionally, each CD had 4 dummy recordings of either healthy controls or radiculopathy participants, which were not included in the analysis. These 4 dummy recordings were the same in discs A

and C, and different in disc B. They were used to introduce some differences between CDs to minimize recollection of the earlier dataset. Changes in sensitivity or sweep speed settings during the recording were marked.

Three experienced, board-certified electromyographers from different academic medical centers across the United States with experience ranging from 10 to 20 years postfellowship training (L.J., M.W., T.M.) evaluated the recordings. They were only informed that the CDs contained 24 EMG videos from individuals with cervical/lumbosacral radiculopathy or healthy controls; the aims of the study were not provided. They were requested to use a standardized recording sheet to evaluate the EMGs (appendix e-1 on the *Neurology*[®] Web site at Neurology.org). They used their usual methods to evaluate the EMGs. We did not standardize EMG interpretation to avoid artificially increasing interrater reliability. We asked the reviewers to grade denervation on an ordinal scale from 0 to 4, where 0 was normal and 4 was interference pattern of fibrillation potentials. Recruitment and reinnervation were graded qualitatively on a scale from 0 to 3 (0: normal and 3: severely reduced recruitment or severe reinnervation). We also asked the reviewers to interpret the EMG as either normal or radiculopathy, and if radiculopathy, to state the level, duration, and severity. Duration of radiculopathy on EMG was classified as acute, subacute, chronic, and ongoing/chronic active. Severity was graded as mild, moderate, or severe. Finally, we asked reviewers to provide standardized feedback regarding the quality of the videos (audio and video quality, clarity of waveforms, length of recording and reasons if a video could not be evaluated), graded on a 1–10 ordinal scale.

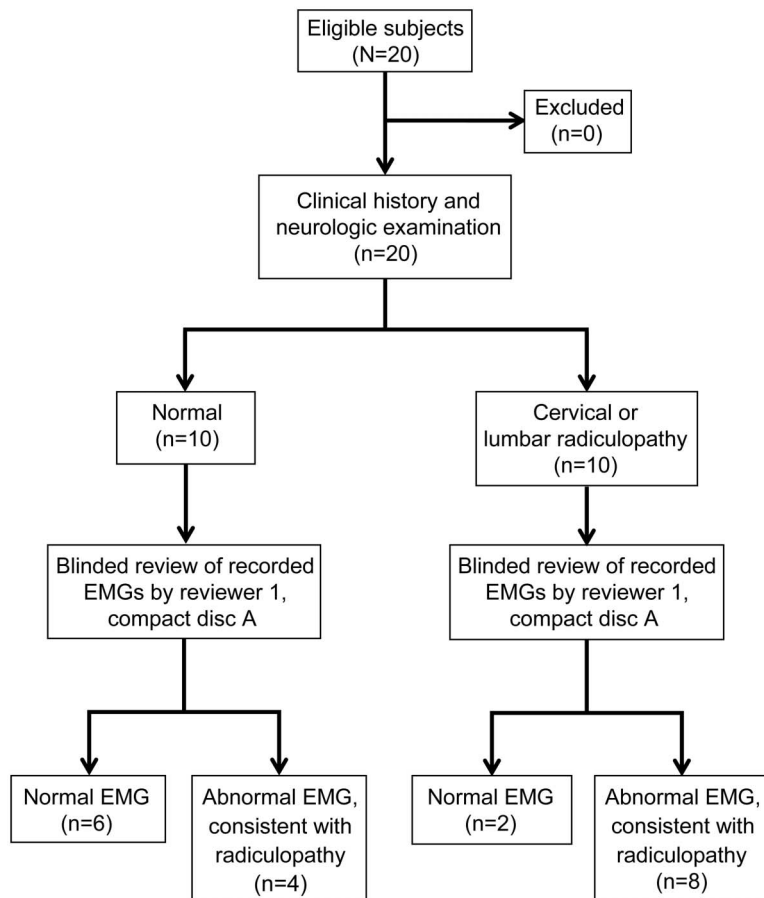
The reviewers first evaluated CD-A, and were given 1 month to return their evaluations. CD-B was sent out 1 month after CD-A was returned, to minimize the possibility of recollection of the earlier dataset. The order of the EMG videos was changed in CD-B for the same reason. One month after CD-B was returned, CD-C was mailed. CD-C had EMG videos in the same order as CD-A; a de-identified clinical history and examination document was also provided. Figures 1 and 2 summarize the study design and work flow.

Statistical analysis. EMG diagnoses were dichotomized into normal or radiculopathy categories. Diagnostic accuracy was measured using sensitivity and specificity. The specificity and sensitivity of each reviewer was represented on a receiver operating characteristic (ROC) plot. Intrarater reliability was measured by comparing a reviewer's EMG interpretations from CD-A to CD-B using Cohen κ (0, no agreement beyond chance; 1, perfect agreement). Interrater reliability was measured by comparing EMG interpretation between raters using Fleiss κ . We then dichotomized radiculopathy EMGs by the presence or absence of ongoing denervation. Intrarater and interrater reliability were calculated for these groups. The degree of agreement was interpreted based on the κ coefficients as follows: 0–0.2 poor; 0.21–0.4 fair; 0.41–0.6 moderate; 0.61–0.8 good; 0.81–1 very good.¹¹

Confirmation bias was assessed by comparing the average proportion of correct diagnoses (based on the clinical reference standard) between CD-A and CD-B to the proportion of correct diagnoses in CD-C (absolute difference). Statistical precision was measured using 95% confidence intervals (CI). To provide a summary estimate of effect and to increase precision, ratings across reviewers and across CDs were combined using inverse-variance weighted averages, where the point estimate for each reviewer was weighted by the inverse of the variance.

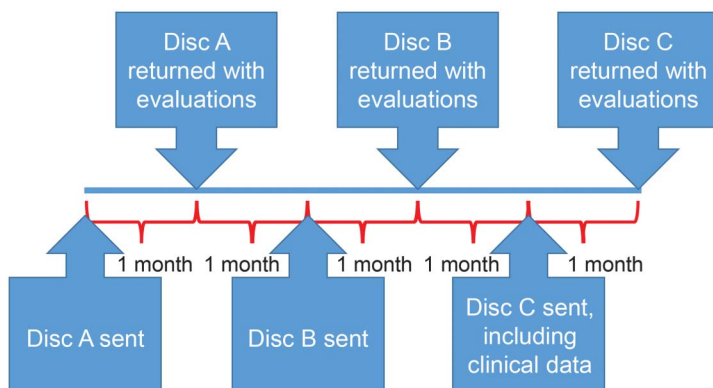
The EMGs where all 3 raters agreed on a diagnosis of radiculopathy were analyzed for agreement on radiculopathy level,

Figure 1 STARD flowchart



duration, and severity. For level, agreement was established if one level was mentioned by all 3 reviewers, even if they mentioned additional levels in their interpretation (e.g., interpretations of C5, C6 and C5 and C5, C6, C7 were analyzed as agreement for C5). When reviewers graded severity as mild to moderate, it was included in the mild group, and moderate to severe was

Figure 2 Study design and time flow



Three compact discs (CDs A, B, and C) consisting of the same 20 EMG videos and 4 dummy EMG videos were sent to 3 blinded electromyographers, who were given a month to review each CD. One month after each CD was returned, the subsequent CD was mailed. A brief clinical history and examination was provided only with CD-C. Each electromyographer recorded his interpretation on a standard EMG data sheet (appendix e-1).

included in the moderate group. Descriptive statistics were applied.

RESULTS Demographics and clinical data. Mean age was 48.8 years (24–85 years, 4 men) for healthy controls and 50.1 years (32–64 years, 7 men) for radiculopathy participants. Six had lumbosacral radiculopathy and 4 had cervical radiculopathy. Four radiculopathy participants had only sensory symptoms; 6 had motor weakness and sensory symptoms. The myotomes affected on EMG were as follows: single level: 5 (2 C7, 2 L5, 1 S1); 2 levels: 3 (2 L5-S1, 1 C5-6); and multilevel: 2 (1 C7-8, T1, 1 L4-5, S1). On EMG, the radiculopathy was graded mild in 5, moderate in 3, and severe in 2 participants. Active denervation was noted in 8 of the 10 radiculopathy EMGs.

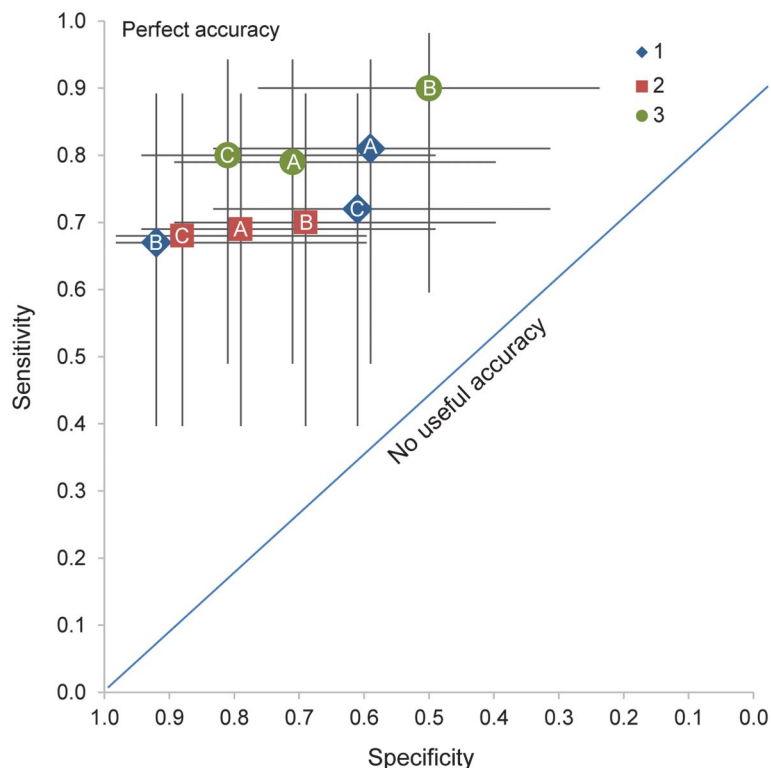
EMG analysis. All reviewers confirmed that they diagnosed radiculopathy when ≥ 2 muscles innervated by the same root but different peripheral nerves showed active denervation or chronic reinnervation. The raw data for all 3 reviewers for all 3 CDs are presented in table e-1.

Sensitivity and specificity of EMG in the diagnosis of radiculopathy. Sensitivities and specificities for each reviewer for each CD are shown in figure 3. Sensitivity of EMG for the diagnosis of clinically confirmed radiculopathy ranged from 70% to 90% and specificity ranged from 50% to 90% across the 3 reviewers and 3 CDs. Combining the 3 reviewers, the pooled sensitivity was 77% (95% CI 63%–90%) and specificity was 71% (95% CI 56%–85%) in CDs A and B (without clinical information), indicating moderate diagnostic accuracy. There were no significant differences in diagnostic accuracy between reviewers or among the 3 CDs. When clinical information was provided (CD-C), the sensitivity was 73%, specificity 77%.

When radiculopathy participants were stratified based on clinical severity into mild and moderate/severe categories based on the absence or presence of motor weakness, sensitivity was comparable at 80% when motor weakness was present (clinically moderate to severe radiculopathy), but dropped to only 40% in the absence of motor weakness (clinically mild radiculopathy).

Intrarater and interrater reliability. The intrarater reliability (Cohen κ) ranged from 0.62 to 0.70 (table 1) and was not significantly different between reviewers. The combined Cohen κ for all 3 reviewers was 0.61, 95% CI 0.41 to 0.81, indicating substantial intrarater reliability. Interrater reliability across all 3 reviewers (table 1) was identical in both CD-A and CD-B (Fleiss κ 0.53, 95% CI 0.28–0.78), indicating moderate interrater reliability. For radiculopathy

Figure 3 EMG in radiculopathy: Receiver operating characteristic (ROC) plot for individual reviewers ($\pm 95\%$ confidence intervals)



1, 2, and 3 represent the blinded EMG reviewers; A, B, and C represent the 3 compact discs (CDs) with 20 EMG videos each. Each reviewer evaluated the EMG videos in all 3 CDs. The error bars represent 95% confidence intervals. Sensitivity ranges from 70% to 90% across reviewers and discs; specificity ranges from 50% to 90%.

EMGs with ongoing denervation (8/10 participants), the combined intrarater reliability (Cohen κ) was 0.78 (95% CI 0.34–1), indicating substantial agreement, and combined interrater reliability (Fleiss κ) was 0.49 (95% CI 0.22–0.75), indicating moderate agreement. Because there were only 2 participants in the group with reinnervation alone, meaningful analysis was not possible.

Table 1 Intrarater reliability (Cohen κ) and interrater reliability (Fleiss κ)

Intrarater reliability (reproducibility)			
Reviewer	Cohen κ	95% CI	p Value
1	0.62	0.28–0.95	0.003
2	0.7	0.39–1.00	0.002
3	0.48	0.08–0.87	0.024
Pooled	0.61	0.41–0.81	
Interrater reliability (agreement)			
Disc	Fleiss κ	95% CI	p Value
A	0.53	0.28–0.78	0.000004
B	0.53	0.28–0.78	0.000004
Pooled	0.53	0.35–0.71	

Abbreviation: CI = confidence interval.

Confirmation bias in EMG. There were no significant differences in the number of correct EMG diagnoses (radiculopathy vs normal) between CD-A/B (averaged across the 2 CDs) and CD-C for any of the 3 reviewers, indicating no significant confirmation bias resulting from the inclusion of clinical information (table 2). Across all 3 reviewers, the change in the number of correct responses ranged from -10% (fewer EMGs correct with clinical information) to $+7.5\%$ (more EMGs correct). The pooled change in percentage of correct EMGs when clinical information was provided was 2.2% (95% CI -13.3% to 17.7%).

Interrater reliability for level, duration, and severity of radiculopathy. All 3 reviewers agreed upon the level of radiculopathy (i.e., chose the same level, regardless of additional levels also being mentioned) in 6/10 (60%) radiculopathy EMGs in both CDs A and B, and 5/10 (50%) EMGs in CD-C. In a second-level analysis, when designations of ≥ 3 roots were excluded (i.e., C5 and C5, 6 would be considered as agreement for C5, but C5 and C5, 6, 7 would not), the agreement dropped to 4/10 (40%). Among the EMGs where all 3 reviewers agreed on the same level of radiculopathy, agreement for severity (mild, moderate, severe) varied across the 3 CDs: 4/6 (67%) in CD-A, 2/6 (33%) in CD-B, and 1/5 (20%) in CD-C. Agreement for duration (acute, subacute, chronic, ongoing/chronic active) also varied across the 3 CDs: 4/6 (67%) in CD-A, 3/6 (50%) in CD-B, and 1/5 (20%) in CD-C.

Interrater reliability for individual EMG parameters. There was 90.6% agreement among the reviewers for the presence of denervation (fibrillation potentials or positive sharp waves) in any muscle. In contrast, there was only 60.4% agreement for the presence of reinnervation (defined as polyphasia, long-duration motor unit potentials, or large amplitude motor unit potentials). The reviewers had not been given

Table 2 Confirmation bias (absolute difference in proportion of correct EMG interpretations between averages of CDs A and B vs CD-C)

Reviewer	Absolute difference in number of correct responses, %	95% CI
1	-10	-38 to $+18.2$
2	$+7.5$	-18.8 to $+33.8$
3	$+7.5$	-18.8 to $+33.8$
Pooled	$+2.2$	-13.3 to $+17.7$

Abbreviations: CD = compact disc; CI = confidence interval; $-$ = Fewer EMGs correct with clinical data; $+$ = more EMGs correct with clinical data.

any standard criteria for grading polyphasicity, duration, or amplitude.

False-positive EMGs and EMGs with disagreement. In 6 cases, the diagnosis of radiculopathy in normal EMGs (false-positive EMGs) was based on reviewers' impressions of mild chronic reinnervation in more than one muscle of a myotome and was not related to participants' age. Only one false-positive EMG was based on the finding by one reviewer of fibrillations in the tibialis anterior and tensor fascia lata in 2 CDs.

In 4 EMG videos of radiculopathy participants, there was disagreement either within or between reviewers (3 lumbosacral and 1 cervical radiculopathy). No specific level of radiculopathy was associated with disagreement. Disagreement also could not be consistently ascribed to any specific motor unit potential parameter (polyphasia, duration, amplitude, or recruitment) or to a specific muscle. One reviewer noted that one video was too short for assessing spontaneous activity. However, in the next CD on that same video, the same reviewer identified denervation. In the 3 other EMGs, differences were likely due to mild abnormalities. One EMG had 1+ fibrillations in gastrocnemius and mild S1 reinnervation; the 2 others had very mild reinnervation.

Reviewers' impressions of EMG recording quality. The median score (range) for audio quality was 7/10 (1–8), for video quality 7/10 (4–8), clarity of waveforms 5/10 (4–8), and length of recording 6/10 (3–8).

Specific comments included the following:

1. Patient age and needle size/type would have been helpful to know.
2. Sound reproduction was a bit limiting...low-amplitude discharges and low-frequency elements were often visible but not audible.
3. This is a "style" difference, but I usually assess motor unit potentials at a lower level of activation than in these recordings.
4. Video looking for spontaneous activity was short for some muscles.
5. 60-Hz artifact.
6. Prolonged unstable baseline with needle movements.
7. Some artifact that may have obscured spontaneous activity.
8. Gain set too high to exclude increased amplitude motor unit action potentials.

DISCUSSION Our study found good intrarater reliability for EMG interpretation but lower interrater agreement. Hence, serial EMG studies should ideally be performed by the same electromyographer. Variability in reviewers' criteria for radiculopathy may have affected interrater reliability; we did not formally assess this. We did

not find evidence for confirmation bias. However, because of our small sample size, our results lacked precision to exclude moderate confirmation bias (up to 18%; CI –13.3% to +17.7%). Our study confirms the previously reported sensitivity of EMG for radiculopathy diagnosis (77%). But our specificity was considerably lower than has been commonly believed,¹² at only 71%. Interpretation of our results of agreement for severity and duration of radiculopathy is difficult because we had only 6 cases in which all evaluators agreed, but overall, the EMG analysis appears to be fairly variable among the 3 reviewers. We did not detect specific levels of radiculopathy, specific muscles, or motor unit parameters to be associated with disagreement. Although 3 of 4 EMGs with disagreement were lumbosacral radiculopathy, this likely represents the distribution of radiculopathy EMGs in the study.

Importantly, the previous studies of specificity have evaluated only lumbar radiculopathy and are methodologically limited by the spectrum of controls. One included asymptomatic controls and controls with back pain and reported a specificity of 54%–58% for any EMG abnormality, and 87.5%–100% for fibrillation potentials in any muscle. This included paraspinal mapping, an EMG technique not routinely used.¹⁰ Another study, using root compression on MRI as the reference standard, reported a specificity of 86% for fibrillation potentials, but whether this referred to abnormalities in a single or multiple muscles was unstated.¹³ The specificity of reinnervation, defined as long-duration motor units, was 81%.¹³ Since both reinnervation and denervation are used together to diagnose radiculopathy, calculation of diagnostic accuracy of these features separately is of uncertain value. Two other studies of specificity were methodologically limited by their choice of controls, making results difficult to interpret, one using different reference standards for radiculopathy participants and controls, and the other using only symptomatic participants, but with incorporation bias.^{14–16}

Our study has limitations. First, we used a clinical reference standard. While that may be of practical significance, radiculopathies frequently present with only pain or sensory symptoms, and in such cases, EMG may be normal. In keeping with this, we found that sensitivity dropped to only 40% in radiculopathy without clinical motor findings. Second, this was an artificial experimental situation and there may have been variability in the reviewers' comfort level with interpretation of prerecorded EMGs. A learning effect may explain the cases of disagreement in the first CD but not in subsequent CDs. There was also minor inconsistency in sweep speed and sensitivity between the 2 investigators performing the EMGs. Although

these were clearly marked on the videos, this may have affected the reviewers' evaluations.

The main methodologic limitation of our study is spectrum bias, because we did not have controls with other conditions that mimic radiculopathy. However, in practice, the most important clinical mimic of radiculopathy is musculoskeletal pain, wherein EMG is expected to be normal. Other differential diagnoses such as brachial plexopathy would likely be suspected by clinical history and examination and may affect confirmation bias but would not be expected to affect diagnostic accuracy.

Surprisingly, we found no evidence that knowledge of clinical information biases EMG interpretation, although we cannot exclude subtle confirmation bias. Additionally, the electromyographers knew that participants either had radiculopathy or were normal. It is possible that confirmation bias plays a greater role when clinical information will help differentiate 2 diagnostic possibilities (e.g., cervical radiculopathy vs brachial plexopathy). Finally, confirmation bias may have been low because the reviewers guessed the nature of the study and ignored the clinical information on the last CD.

Perhaps most importantly, this study shows that although EMG has been used in clinical practice for more than 50 years, its specific strengths and weaknesses remain challenging to evaluate, even when being performed and interpreted by experts. Only through better understanding of the intricacies of the technique, including some of the factors evaluated here, can we hope to apply the test most effectively for the care of our patients.

AUTHOR CONTRIBUTIONS

P. Narayanaswami: study concept and design, acquisition of data, analysis or interpretation of data, drafting/revising the manuscript, critical revision of the manuscript for important intellectual content, study supervision. T. Geisbush: study concept and design, acquisition of data, analysis or interpretation of data, drafting/revising the manuscript. L. Jones: analysis or interpretation of data, drafting/revising the manuscript, critical revision of the manuscript for important intellectual content. M. Weiss: analysis or interpretation of data, drafting/revising the manuscript, critical revision of the manuscript for important intellectual content. T. Mozaffar: analysis or interpretation of data, drafting/revising the manuscript. G. Gronseth: analysis and interpretation of data, drafting/revising the manuscript. S. Rutkove: study concept and design, acquisition of data, analysis or interpretation of data, drafting/revising the manuscript, critical revision of the manuscript for important intellectual content, study supervision.

STUDY FUNDING

Supported by NIH K24 NS060951.

DISCLOSURE

P. Narayanaswami, T. Geisbush, and L. Jones report no disclosures relevant to the manuscript. M. Weiss has received personal compensation for speaking for Walgreens and NuFactor and received research support from ALS therapy Alliance and the Northeast ALS Consortium.

T. Mozaffar, G. Gronseth, and S. Rutkove report no disclosures relevant to the manuscript. Go to Neurology.org for full disclosures.

Received May 19, 2015. Accepted in final form September 16, 2015.

REFERENCES

1. Kendall R, Werner RA. Interrater reliability of the needle examination in lumbosacral radiculopathy. *Muscle Nerve* 2006;34:238–241.
2. Haldeman S, Shouka M, Robboy S. Computed tomography, electrodiagnostic and clinical findings in chronic workers' compensation patients with back and leg pain. *Spine* 1988;13:345–350.
3. Khatri BO, Baruah J, McQuillen MP. Correlation of electromyography with computed tomography in evaluation of lower back pain. *Arch Neurol* 1984;41:594–597.
4. Nardin RA, Patel MR, Gudas TF, Rutkove SB, Raynor EM. Electromyography and magnetic resonance imaging in the evaluation of radiculopathy. *Muscle Nerve* 1999;22:151–155.
5. Aminoff MJ, Goodin DS, Parry GJ, Barbaro NM, Weinstein PR, Rosenblum ML. Electrophysiologic evaluation of lumbosacral radiculopathies: electromyography, late responses, and somatosensory evoked potentials. *Neurology* 1985;35:1514–1518.
6. Kuruoglu R, Oh SJ, Thompson B. Clinical and electromyographic correlations of lumbosacral radiculopathy. *Muscle Nerve* 1994;17:250–251.
7. Leblhuber F, Reisecker F, Boehm-Jurkovic H, Witzmann A, Deisenhammer E. Diagnostic value of different electrophysiologic tests in cervical disk prolapse. *Neurology* 1988;38:1879–1881.
8. Tonzola RF, Ackil AA, Shahani BT, Young RR. Usefulness of electrophysiological studies in the diagnosis of lumbosacral root disease. *Ann Neurol* 1981;9:305–308.
9. Wu ZA, Tsai CP, Yang DA, Chu FL, Chang T. Electrophysiologic study and computerized tomography in diagnosis of lumbosacral radiculopathy. *Zhonghua Yi Xue Za Zhi* 1987;39:119–125.
10. Haig AJ, Tong HC, Yamakawa KS, et al. The sensitivity and specificity of electrodiagnostic testing for the clinical syndrome of lumbar spinal stenosis. *Spine* 2005;30:2667–2676.
11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174.
12. Robinson LR. Electromyography, magnetic resonance imaging, and radiculopathy: it's time to focus on specificity. *Muscle Nerve* 1999;22:149–150.
13. Coster S, de Bruijn SF, Tavy DL. Diagnostic value of history, physical examination and needle electromyography in diagnosing lumbosacral radiculopathy. *J Neurol* 2010;257:332–337.
14. Albeck MJ, Taher G, Lauritzen M, Trojaborg W. Diagnostic value of electrophysiological tests in patients with sciatica. *Acta Neurol Scand* 2000;101:249–254.
15. Cho SC, Ferrante MA, Levin KH, Harmon RL, So YT. Utility of electrodiagnostic testing in evaluating patients with lumbosacral radiculopathy: an evidence-based review. *Muscle Nerve* 2010;42:276–282.
16. Tong HC. Specificity of needle electromyography for lumbar radiculopathy in 55- to 79-yr-old subjects with low back pain and sciatica without stenosis. *Am J Phys Med Rehabil* 2011;90:233–238; quiz 239–242.