A Prospective, Masked 18-Month Minimum Follow-up On Neurophysiologic Changes In Persons with Spinal Stenosis, Low Back Pain, and No Symptoms

Andrew J. Haig, MD, Karen S. J. Yamakawa, MS, Christopher Parres, BS, Anthony Chiodo, MD, Henry Tong, MD, MS

Objectives: To describe neurophysiologic changes over time in persons with and without spinal complaints and to assess whether paraspinal denervation predicts change in stenosis on magnetic resonance imaging (MRI) and clinical course.

Design: Prospective, controlled, masked trial.

Setting: University spine program.

Participants: Persons aged 55 to 80 years, screened for polyneuropathy and determined on clinical examination to have spinal stenosis, mechanical low back pain, or no spinal symptoms.

Interventions: A comprehensive codified history was obtained and subjects underwent physical examination, ambulation testing, masked electrodiagnostic testing including paraspinal mapping, and MRI, repeated at greater than 18 months. This study presents detailed technical information and additional analyses not reported previously.

Main Outcome Measurements: Change in electrodiagnostic findings. Among persons with clinical stenosis, relationship of change in paraspinal mapping scores to MRI findings and clinical changes.

Results: Of 149 initial subjects, 83 (79.3% of eligible subjects) repeated testing at 20 ($\pm$2 SDs) months. No significant change in limb muscle spontaneous activity or motor unit pathology was noted in any group. In 23 persons with initial diagnosis of stenosis, paraspinal mapping electromyography related to change in diagnosis over time (analysis of variance $F = 3.77$, $P = .037$), but not to most initial magnetic resonance imaging measurements or to change in spinal canal diameter.

Conclusions: Clinical spinal stenosis is neurophysiologically stable in most persons. Paraspinal electromyographic changes reflect large changes in clinical course, but neither neurophysiologic nor clinical changes relate to change in spinal geometry over 20 months.

INTRODUCTION

Lumbar spinal stenosis is a commonly treated but poorly understood problem. The hallmark clinical symptom of neurogenic claudication does not occur in all persons thought to have stenosis, and symptoms mimicking neurogenic claudication occur with vascular and even other spinal disorders [1-3]. While the terminology suggests an anatomical lesion, anatomical stenosis is so common in persons without symptoms or with mechanical back pain that magnetic resonance imaging (MRI) or other anatomical tests are not definitive [4-6].

For persons with spinal stenosis, a serious consideration is the possible progression of neurological deficit. This concern may drive them and their physicians toward more invasive treatments. Electrodiagnostic testing (EDX) is commonly used to determine the neurologic deficit associated with spinal stenosis [7-12]. There is a perception that changes in EDX relate to changes in symptoms; however, to the authors' knowledge, no prospective study or case report has validated this belief for any spinal disorder.

In addition to the practical clinical question answered by following electrodiagnostic findings over time, this kind of inquiry can help solve some theoretical questions. The authors have proposed that denervation of paraspinal muscles can cause segmental hyper-

A.J.H. Department of Physical Medicine and Rehabilitation, The University of Michigan, 325 E Eisenhower, Ann Arbor, MI 48108. Address correspondence to A.J.H.; e-mail: andyhaig@umich.edu
Disclosure: nothing to disclose

K.S.J.Y. Department of Physical Medicine and Rehabilitation, The University of Michigan, Ann Arbor, Michigan
Disclosure: nothing to disclose

C.P. Department of Physical Medicine and Rehabilitation, The University of Michigan, Ann Arbor, Michigan
Disclosure: nothing to disclose

A.C. Department of Physical Medicine and Rehabilitation, The University of Michigan, Ann Arbor, Michigan
Disclosure: 2A Medtronic

H.T. Department of Physical Medicine and Rehabilitation, The University of Michigan, Ann Arbor, Michigan
Disclosure: nothing to disclose

This is a publication of the Spine Program of the University of Michigan, which is funded by the U.S. Department of Health and Human Services, National Institutes of Health, under Grant 5 R01 NS41855-02. The opinions contained in this publication are those of the grantee and do not necessarily reflect those of the U.S. Department of Health and Human Services. The funding source had no role in any of the following: the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Disclosure Key can be found on the Table of Contents and at www.pmrjournal.org

Submitted for publication March 11, 2008; accepted October 16.

mobility and loss of kinesthetic sense, leading to more wear on facet joints, joint hypertrophy, and eventually a stenotic spinal canal [13].

The Michigan Spinal Stenosis Study was designed to answer these questions. It is a masked cohort study of EDX, MRI, and clinical examination that follows persons with clinically defined spinal stenosis, low back pain only, or no symptoms for longer than 18 months. Analysis of data on the initial cohort has resulted in a number of publications, including the first masked, controlled trial of electrodiagnosis to establish the validity of the test [14] and subsequent analysis demonstrating the relative benefits of electrodiagnosis over MRI in diagnosing spinal stenosis [5,15]. More pertinent to the current question, the authors have shown that function and pain at 18 months are predicted by baseline function and sleep deficits, not MRI or EDX findings [16]. While previous reports used summarized EDX results, they did not provide sufficient technical detail for EDX specialists to understand the variation in individual tests. Furthermore, data regarding the relationship of paraspinal denervation with radiological changes over time have not been reported sufficiently to address the principal hypothesis of the research project.

The current article has 2 purposes: (1) to determine if specific electrodiagnostic findings change over time in persons with clinical stenosis, low back pain, or asymptomatic volunteers, and (2) to seek any relationship between paraspinal denervation or change in paraspinal denervation with change in clinical diagnosis or radiologic measures over 18 months.

METHODS

Subjects and Testing

The study protocol has been described previously [15]. In summary, the study sought to recruit persons with no back pain, mechanical low back pain without radiologic stenosis, and clinically evident spinal stenosis. Subjects were recruited based on preliminary screening criteria for these 3 categories, but final diagnosis as used in this study is based on a comprehensive history and physical examination as noted later. All subjects underwent masked MRI or review of MRI, masked EDX, and ambulation testing and completed numerous questionnaires, as detailed later. All tests were repeated at more than 18 months after the initial testing. Figure 1 outlines these steps. Specific details are given in the next section.

Recruitment and Preliminary Screening

At a university hospital, serial lumbar MRI reports from the university imager were screened for persons aged 55 to 80 years and for any exclusion criteria (previous surgery, tumor, etc). Radiologist reports were supplemented with review of all images by a study physician to select persons with "preliminary diagnosis of stenosis." Among those with no apparent stenosis on MRI, further review of the university computerized medical records excluded persons with report of pain radiating below the knee. The subsequent group was labeled "preliminary diagnosis of mechanical back pain." People from the community with no back pain complaint, who had none of the exclusion criteria but were within the same age range, were recruited via postings and advertisements.

All potential subjects were then screened by telephone for exclusion criteria, including known polyneuropathy, diabetes, heavy alcohol use, previous lumbar surgery, or relative contraindications to MRI or EDX. Subjects who had plans for surgery were also excluded due to the project's long-term goals, which is to follow the natural history of the disorders over 18 months. Findings from this preliminary process were not revealed to the examining physiatrist who made the final clinical diagnosis. All subjects were given informed consent and were compensated. The university's ethical review board approved the study.

Clinical Evaluation

All subjects filled out an extensive patient questionnaire, including the Pain Disability Index, the Quebec Back Pain Disability Index, and the McGill Pain Questionnaire; a visual analog pain scale (VAS); and a pain drawing, along with a 5-page clinical spine questionnaire that encompassed medical history, review of systems, family history, and social history [17-21]. Each patient performed an ambulation test in which they were instructed to walk at a comfortable speed for 15 minutes and wore a pedometer during waking hours at home for a week.

Physiatrists reviewed the questionnaires and performed a comprehensive and codified spinal history and physical examination. Four of the physiatrists were board certified in all 3 specialties of physical medicine and rehabilitation, pain medicine, and electrodiagnostic medicine, and 5 others were in a clinical fellowship designed to qualify them for all 3 boards. Their primary residency training was at 8 different universities, suggesting a diverse background. The physiatrist's impression as to whether the subject had low back pain, had spinal stenosis, or was asymptomatic is termed the "clinical diagnosis" throughout this report.

The diagnostic standard for the current study was the conclusion of the examining physician, who was not restricted by any a priori criteria. To establish face validity of the clinical diagnosis, a number of potential associations were examined with findings often thought to relate to spinal stenosis. Portions of the clinical examination that are thought to be consistent with the clinical syndrome of spinal stenosis were found to relate to the clinician's diagnosis. These included pain below the knee, pain severity, difficulty with walking, 15-minute ambulation velocity, spine tenderness, strength deficits, reflex deficits, straight leg raise test, and femoral stretch test. Also, a senior academic spine surgeon who was masked from the radiologist's and physiatrist's diagnoses reviewed the MRI images, history and physical examination data, and patient questionnaire to independently arrive at a neurosurgical diagnosis. This was compared with the clinical diagnosis. Both the spine surgeon's impression and the examination details supported the reasonableness of the clinical diagnosis [5].

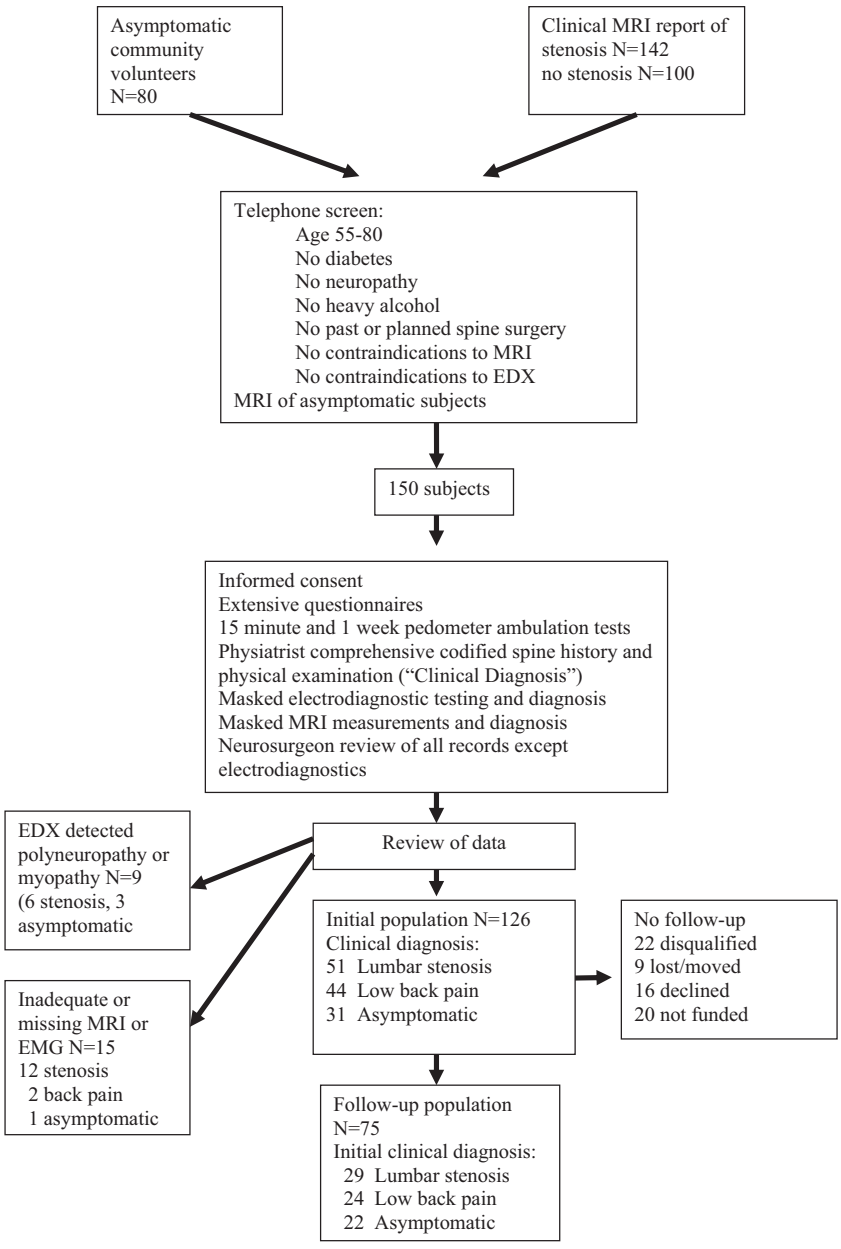

Figure 1. Recruitment and testing of subjects in the Michigan Spinal Stenosis Study.

MRI

The asymptomatic volunteers underwent lumbar MRI, and MRI of the others, performed within 6 months of the study, were all reviewed. All scans were noncontrast lumbosacral spinal MRI performed on a GE Signa Horizon LX (General Electric Medical Systems, Milwaukee, WI). They including sagittal T2-weighted scans (field of view [FOV]: 30; scan thickness [ST]: 3.0 mm; interscan spacing [IS]: 0.5 mm; matrix 384 × 192; TR: 3000; TE: 102; pulse: fast-spin echo [FSE]), sagittal T1-weighted scans (FOV: 30; ST: 3 mm; IS: 0.5 mm; matrix: 256 × 192; TR: 400-700; TE: min full; pulse: SE), and axial T2-weighted scans (FOV: 20; ST: 4 mm; IS 5 mm; 5 slices through each disc space T12-L1 through L5-S1; matrix: 256 × 256; TR: 3000-5000; TE: 102; pulse: FSE).

All images were masked and reviewed by a neuroradiologist at a workstation (Windows Advantage Workstation; General Electric Medical Systems). Anatomic measurements were made using an electronic cursor at each lumbar intervertebral disc level. Measurements included midline anteroposterior osseous spinal canal diameter, midline anteroposterior thecal sac diameter, osseous spinal canal cross-sectional area, thecal sac cross-sectional area, osseous interfacet distance (measured between the medial osseous margins of each facet joint), distance between the medial joint capsular margins of each facet joint, and the anteroposterior lateral recess diameter on the right and left side. Previous work suggested symptoms occur because vascular compromise happens when the vasa nervorum are compressed at

2 levels [22,23]. Therefore, composite scores of the average of the smallest 2 canal diameters, smallest 2 thecal sac diameters, and so on, were developed. Test-retest validity of the radiologists was excellent, as documented elsewhere [24].

Electrodiagnostic Testing

A masked EDX specialist performed a detailed electrodiagnostic study. The adequacy of masking is described elsewhere, but in summary, at most, unmasking could have affected 6% or fewer of the examinations [25]. Electrodiagnostics were performed with a Nicolet Viking II (Nicolet Biomedical, Madison, WI) using a 50- or 75-mm monopolar needle. Skin temperature was monitored and heat was applied when necessary to keep skin temperature above 32° C.

As recommended by Dillingham et al [26], the testing included exploration of 5 muscles with overlapping root innervation—the tensor fascia lata (L4, L5, and S1 innervated), vastus medialis (L2, L3, and L4), tibialis anterior (L4 and L5), extensor hallucis longus (L5 and S1), and the medial gastrocnemius (S1 and S2) on either the most symptomatic side—or if symptoms were absent or symmetrical, a leg was chosen by the assistant by coin toss [26]. In each muscle, the presence of fibrillation potentials (a sign of denervation) was scored after 6 insertions in 4 different directions as 0 to 4+ using Daube's definitions [27]. Ten motor units were sampled in each muscle. Because it is not the usual clinical practice to actually capture and measure each motor unit, examiners were asked to make informal estimates of typical motor unit amplitude, number of polyphasic motor units ("polys"), and motor unit recruitment (firing rate of the first motor unit when a second motor unit was recruited). The number of polyphasic motor units seen per muscle was recoded as none, as 1 to 2/10 polys, and as more than 2/10 polys. Sural sensory response and peroneal motor responses from the ankle, fibular head, and popliteal space were measured. Bilateral H waves and peroneal F waves were performed by a technician and interpreted by the electromyodiagnostician.

This study used the MiniPM abbreviated version of the original Paraspinal Mapping (PM) to study the paraspinal muscles bilaterally [28-31]. Convention has led others to use the term "paraspinal mapping" for this abbreviated version, so the authors will use that term throughout this document. PM is an anatomically validated, quantified scoring system for the paraspinal muscles. The range of normal values for PM has been defined, good interrater reliability has been established ($r = 0.830$, $P = .041$), and limited clinical evidence suggests that it can localize the root level of a lesion [24,32,33].

The PM technique is described in detail elsewhere [30,31]. Briefly, it includes palpation of the inferior border of the 3 lowest lumbar spinous processes and the midpoint between the posterior superior iliac spines, measuring 2.5 cm laterally and (for the L3, L4, and L5 spinous processes) 1 cm cranial. From each of these 4 locations, a 50- to 75-mm monopolar electromyographic (EMG) needle is inserted at a 45- to 60-degree angle to the surface in 3 different directions—cranial 45 degrees, directly across to the spinous process, and caudal 45 degrees—and advanced through the muscle in 5-mm movements to detect abnormal muscle membrane instability. Any fibrillations must last longer than 1 second and be reproducible. Scores for the medial-most 1 cm are scored separately from the more lateral components of the insertion. Depending on the severity of findings, scores ranged from 0 to 4+ in any of 24 total locations.

A total score for the side (number of +'s) is determined, resulting in a potential range of 0 to 96, but 95% of asymptomatic younger persons score 2 or less. A cutoff above 4 for abnormal on the PM score was established prior to this analysis based on data from the entire asymptomatic population presented elsewhere, which shows that the range of normal in older asymptomatic persons is higher than the 95% cutoff of greater than 2 that the authors had earlier reported in younger persons [30,33]. PM was performed bilaterally, but for the current study we used only the data for the side on which limb EDX was performed.

Follow-up and Disqualifications

Subjects were invited back for retesting at 18 months after initial testing. This time frame was chosen primarily from a pragmatic standpoint. The authors did not think that the federal funding source would support a longer-term study. Those who had a newly diagnosed exclusion criterion, including lumbar surgery, were eliminated. Initial testing occurred late enough in the funding cycle that a number of subjects were not included in the follow-up cohort. A few others declined participation in the 18-month follow-up.

At more than 1 year after testing, a quality check on electrodiagnostic data on all subjects was performed, in conjunction with review of available medical records when needed. It was determined whether there was evidence of polyneuropathy or myopathy, which would have excluded them from recruitment, if it had been known. Neuromuscular disorders were determined by characteristic needle examination findings not seen in radiculopathy (eg, brief, short, early recruited polys in proximal muscles for myopathy) and nerve conduction findings (eg, sural nerve or peroneal nerve slowing for polyneuropathy). These subjects were removed from further analysis.

A few MRI examinations were not completed or had technical errors. All except 1 subject completed the EDX, but persons who had poorly relaxed muscles at 2 or more of the 4 levels explored on either side of the PM grid were eliminated. A total of 15 subjects were eliminated from analysis because of these testing issues. Three subjects had some unreliable paraspinal EMG data at only 1 spinal level. These levels were assumed to be normal.

Statistical Methods

Data were initially entered into a Microsoft Excel (Microsoft, Redmond, Washington) database where errors were checked and cleaned. SPSS Version 12.0 (SPSS, Chicago, Illinois) was used for statistical analysis. Significance was accepted at $P < .05$.

Table 1. *Spontaneous activity changes in the limb muscles in persons with no symptoms, low back pain, and clinical spinal stenosis*

	Tensor Fascia Lata	Vastis Medialis	Tibialis Anterior	Extensor Hallucis Longus	Medial Gastrocnemius	Whole Subject
			Asymptomatic (n = 23)			
Abnormal	0	0	0	0	0	0 (0%)
Increased	0	0	1	1	0	2 (8.7%)
Decreased	0	0	0	0	0	0 (0%)
			Low Back Pain (n = 28)			
Abnormal	3	0	2	3	2	6 (21.4%)
Increased	2	0	1	2	3	6 (21.4%)
Decreased	2	0	1	0	2	3 (10.7%)
			Clinical spinal stenosis (n = 32)			
Abnormal	4	2	2	3	4	7 (21.9%)
Increased	0	0	2	2	4	5 (15.6%)
Decreased	4	2	2	3	4	5 (15.6%)

For each diagnosis and muscle, the number of subjects abnormal on initial testing is indicated in the first row. The number of muscles with an increase in spontaneous activity over 18 months is in the next row, and the number of muscles with a decrease in spontaneous activity is in the third row. Increase and decrease are based on any change in the 0-4+ scoring of spontaneous activity. Thus, a person whose spontaneous activity changed from 1+ to 2+ in a muscle would have an increase, while a person whose spontaneous activity changed from 2+ to normal would have a decrease.

RESULTS

Subjects

Figure 1 shows the process of subject recruitment, selection, and follow-up. After initial testing, certain subjects were disqualified, including 9 subjects due to spinal surgery, 1 who was deceased, 12 who had new contraindications (anticoagulation, cancer, diabetes, neuropathy), and 2 who had moved out of state. Seven did not respond to requests for testing and 16 were not interested in repeat testing. Unfortunately, study funding ran out before 20 other subjects became eligible. Eighty-three persons (79% of subjects who were eligible for repeated testing), including 32 persons with clinically defined stenosis, underwent follow-up testing. Average time to follow-up was 20.0 ± 2.0 months (range 18.3-29.2 months).

Compared with the subjects who did not complete follow-up testing, when stratified by clinical diagnosis the final cohort was in general not any different in terms of EMG findings (limb fibrillations, limb polys, and PM scores). Two exceptions include the extensor hallicus longus in the asymptomatic group, where a significantly higher percentage of dropouts (58.3%) had more than 2/10 polys (58% vs 13%, $\chi^2 = 7.926$, $P = .005$) and "any fibrillations in the whole subject" among the stenosis group, which approached significance (dropouts 44.8% vs follow-up subjects 21.9%, $\chi^2 = 3.637$, $P = .057$).

Limb EMG Findings

Abnormal spontaneous activity is considered a fairly concrete finding of pathology. Changes in limb muscle spontaneous activity in the 3 diagnostic groups over time are reported in Table 1. A large percentage of the subjects across all 3 groups, ranging from 67.9% to 100.0%, had unchanged scores in all muscles. None of the asymptomatic subjects had abnormal spontaneous activity on initial evaluation, but at follow-up, 1 had 2+ fibrillations in the tibialis anterior and another had 3+ positive waves in the extensor hallucis longus. Among persons with low back pain but no stenosis, 22 (78.6%) were initially completely normal. The vastus medialis remained normal in all subjects, but each of the other muscles was abnormal in about 10% of instances. As a majority (greater than 50%) of the cells have expected count less than 5, χ^2 analyses were not performed to assess the differences in the distribution (0-4+ according to Daube's definitions) of the initial versus follow-up scores [27]. There was no significant increase in spontaneous activity in any muscle or in the subjects as a whole over time. Although severity changed in a few individuals (0-4 rating), most were unchanged. The limb examination for fibrillations was abnormal in 7 (21.9%), but essentially unchanged over time, with an equal number of subjects (15.6%) worsening or improving somewhat electrophysiologically.

Motor unit configuration changes over time are presented in Table 2. Among the asymptomatic group, 26.1% had more than 2/10 polys. The tensor fascia lata and the extensor hallucis longus, which share L5 innervation, both had an abnormal proportion of polys in over 10% of asymptomatic subjects. Almost 40% of persons thought to have mechanical low back pain had more than 2/10 polys, again with the 2 L5 muscles most involved. A similar number (43.8%) of subjects with stenosis had abnormal polyphasics with a similar distribution. The majority of all 3 groups increased or decreased somewhat in the extent of polyphasicity, but there was no trend toward worsening or improving.

There was a significant change in the mean number of polys in the vastus medialis, but none of the other muscles. χ^2 analyses were not performed to assess the differences in the distributions of the initial vs follow-up scores broken down by 3 groups (0/10 polys, 1-2/10 polys, and more than 2/10 polys), as a majority of the cells have an expected count of less than 5.

Table 2. *Motor unit morphology changes in persons with no symptoms, low back pain, and clinical spinal stenosis*

	Tensor Fascia Lata	Vastis Medialis	Tibialis Anterior	Extensor Hallicus Longus	Medial Gastrocnemius	Whole Subject
Asymptomatic (n = 23)						
>2/10 polyphasic initial	3	1	1	3	0	6 (26.1%)
Follow-up	4	1	4	3	0	7 (30.4%)
Polyphasic motor units initial (SD)	0.86 (1.13)	0.43 (0.95)	0.65 (0.78)	1.26 (1.14)	0.35 (0.57)	1.74 (1.25)
Follow-up (SD)	0.55 (2.74)	0.26 (0.69)	0.78 (1.54)	0.87 (1.22)	0.35 (0.71)	1.43 (1.78)
Increase in % polyphasic	6	3	5	4	4	6 (26.1%)
Decrease in % polyphasic	8	5	10	11	6	11 (47.8%)
Low Back Pain (n = 28)						
>2/10 polyphasic initial	5	1	3	7	0	11 (39.3%)
Follow-up	5	3	5	8	1	10 (35.7%)
Polyphasic motor units initial (SD)	1.07 (1.49)	0.36 (0.78)	1.00 (1.39)	1.33 (1.59)	0.21 (0.50)	2.07 (1.51)
Follow-up (SD)	0.89 (1.58)	0.50 (1.11)	0.79 (1.57)	1.63 (2.12)	0.36 (0.78)	1.82 (2.09)
Increase in % polyphasic	6	4	4	7	5	7 (25.0%)
Decrease in % polyphasic	9	4	11	7	4	14 (50.0%)
Clinical Spinal Stenosis (n = 32)						
>2/10 polyphasic initial	4	2	5	10	4	14 (43.8%)
Follow-up	5	2	6	8	3	13 (40.6%)
Polyphasic motor units initial (SD)	1.16 (1.22)	0.81 (1.06)	1.22 (1.34)	1.88 (1.72)	0.94 (1.16)	2.25 (1.67)
Follow-up (SD)	0.72 (1.11)	0.31 (0.82)	0.94 (1.44)	1.31 (1.64)	0.56 (1.37)	1.94 (1.80)
Increase in % polyphasic	8	3	7	9	6	12 (37.5%)
Decrease in % polyphasic	12	13	11	14	14	14 (43.8%)

For each muscle and diagnosis, the number of subjects with more than 2/10 polyphasic motor units and the percentage of motor units that were polyphasic are noted at initial assessment and at 18-month follow-up. The number of muscles with an increase or decrease in polyphasia over 18 months is also noted.

Paraspinal Denervation

As expected, the PM scores of patients with spinal stenosis averaged higher than those with back pain and these were higher than PM scores in the asymptomatic group (Table 3). There were no statistically significant differences in PM score among subjects who underwent follow-up and those who did not (independent samples t test $P > .20$ all cases). The paired-samples t test did not find a significant mean difference between the initial PM score and the follow-up PM score. Nonsignificant κ results indicate lack of agreement between 2 categorical measures of normal vs abnormal PM score and initial vs follow-up clinical finding. There was substantial variability (although generally within the authors' "normal" range) in PM scores from initial to follow-up in the asymptomatic population.

The ability of initial PM score to predict change in clinical diagnosis is reported in Table 4. The χ^2 test shows a significant agreement between the physiatrist initial diagnosis and final diagnosis ($\kappa = 0.396$, $P < .001$). Significant mean differences were not found, with respect to change in diagnosis (eg, asymptomatic = asymptomatic or asymptomatic → back pain or asymptomatic → stenosis) within each respective population. The very high scores in the 2 subjects with stenosis who were later asymptomatic and the higher scores in the 2 asymptomatic subjects who went on to present with stenosis are of interest. Other disorders that denervate the paraspinal muscles can mimic spinal stenosis.

Not shown here, among the persons who had stenosis, there was also no significant relationship between PM scores and numerous clinical factors, including the clinical severity (mild, moderate, or severe), ambulation velocity, Pain Disability Index, visual analog pain scale, the Quebec Back Pain Disability Scale, the McGill Pain Questionnaire, or the presence of pain below the knee on pain drawing.

Table 3. *Initial and follow-up paraspinal mapping (PM) scores in the stenosis, back pain, and asymptomatic groups*

	No Symptoms	Back Pain	Stenosis
No. of subjects followed	22	24	29
PM score initial (mean, SD)	1.68 (2.61)	2.83 (5.21)	3.93 (5.11)
PM score follow-up	1.36 (2.06)	4.42 (6.57)	3.62 (3.84)
Initial vs follow-up t value (paired-samples test)	−0.471	1.278	−0.299
Certainty (P value)	.642	.214	.767
PM abnormal (>4) initial (n, %)	3 (13.6%)	4 (16.7%)	11 (37.9%)
PM abnormal (>4) follow-up (n, %)	2 (9.1%)	7 (29.2%)	10 (34.5%)
Initial vs follow-up χ^2 (κ statistics)	−0.122	−0.038	0.329
Certainty (P value)	.556	.841	.076

There was no significant change in the PM scores over time (paired t test <1.3, $P > .2$ all cases) or in the category of normal vs abnormal over time, although there was a trend for the stenosis subjects to have similar scores at follow-up.

Table 4. *The relationship between initial PM scores and change in clinician diagnosis over 18 months*

	Initial Diagnosis					
	Asymptomatic		Back Pain		Stenosis	
Final Diagnosis	n	PM*	n	PM† (SD)	n	PM‡ (SD)
Asymptomatic	15	1.87 (2.67)	1	3.00 (—)	2	10.50 (0.71)
Back pain	5	0.00 (0.00)	14	2.21 (3.89)	11	4.09 (6.27)
Stenosis	2	4.50 (3.54)	9	3.78 (7.19)	16	3.00 (3.97)
Total	22	1.68 (2.61)	24	2.83 (5.21)	29	3.93 (5.11)

Values given as mean (SD).
Analysis of variance:
*$F = 2.585$, $P = .102$; †$F = 0.231$, $P = .796$; ‡$F = 2.075$, $P = .146$.

The relationship between the numerous radiologic measures and PM scores was explored among the persons with clinical spinal stenosis, as shown in Table 5. There were significant relationships between PM score and measures of the thecal sac at L4-5 (anteroposterior dimension, $r = -0.298$, $P = .038$; area of the sac, $r = -0.295$, $P = .040$). The measure of the "average of the 2 smallest anteroposterior sac measures" trended toward significance ($r = -0.261$, $P = .070$). Also, the area of the spinal canal at L1-2 trended toward a significant relationship with PM ($r = -0.257$, $P = .081$). All other measures—including anteroposterior measures of, and area of the spinal canal at each of the 5 levels, smallest sac and canal measurements, and averages of the 2 smallest sac and canal measures, a total of 24 of the 28 measures on the symptomatic side—showed no trend. There were no significant relationships between PM score on the asymptomatic side regarding any of these findings.

For persons with spinal stenosis, the change in diagnostic category did relate significantly to the change in PM score (analysis of variance, $F = 3.770$, $P = .037$) but not to change in spinal canal diameter (Table 6). The post hoc analysis using Tukey's honest significant difference test in the stenosis subjects identified a significant mean difference between the stenosis → asymptomatic and stenosis = stenosis groups (mean = −9.50 [2.12] vs mean = 1.00 [4.41], $P = .029$). No significant relationship between change in PM or MRI measures and change in clinical diagnosis was found for persons with back pain or no symptoms.

DISCUSSION

This is the first study to systematically look at a needle EMG diagnostic protocol for lumbar radiculopathy over time in an asymptomatic population or in a group with spinal disorders. This findings suggest that denervation as tested by EMG remains generally stable over 18 months. There are a number of clinical and scientific implications.

The study methodology has much to offer. This is also the first adequately masked study of needle EMG and the first to prospectively evaluate spinal stenosis alongside 2 different control groups—asymptomatic persons and those with a more clinically relevant alternative presentation of mechanical back pain. The codified, quantified procedures and the diversity of testing physicians suggest that the findings are both reproducible and representative of typical practice.

Some limitations exist, however. It is legitimate to question the ability of needle EMG to measure change over time. Neither the 0 to 4+ scale for abnormal insertional activity proposed by Daube nor the informal motor unit analysis used here (and in most clinical settings) has been subjected to interrater or intrarater reliability studies. With repair and regeneration, it is possible that the primary manifestation of denervation over time

Table 5. *The relationship between change in diagnosis over 18 months and changes in test results (MRI anteroposterior spinal canal diameter (AP canal) and paraspinal mapping scores)*

	Initial Diagnosis								
	Asymptomatic			Back Pain			Stenosis		
Final Diagnosis	n	PM* (SD)	AP Canal† (SD)	n	PM‡ (SD)	AP Canal§ (SD)	n	PM** (SD)	AP Canal†† (SD)
Asymptomatic	15	−0.20 (3.36)	−0.19 (1.87)	1	0.00 (—)	1.70 (—)	2	−9.50 (2.12)	2.65 (3.75)
Back pain	5	0.60 (0.89)	−0.62 (2.50)	14	2.93 (7.47)	−0.03 (2.80)	11	−0.55 (6.19)	−0.21 (2.31)
Stenosis	2	−3.50 (4.95)	−0.60 (0.28)	9	−0.33 (2.78)	0.17 (0.66)	16	1.00 (4.41)	0.69 (1.81)
Total	22	−0.32 (3.17	−0.33 (1.89)	24	1.58 (6.07)	0.11 (2.26)	29	−0.31 (5.59)	0.51 (2.16)

Values given as mean (SD).
Analysis of variance:
*$F = 1.259$, $P = .306$; †$F = 0.109$, $P = .898$; ‡$F = 0.813$, $P = .457$; §$F = 0.257$, $P = .776$; **$F = 3.770$, $P = .037$ (significant); ††$F = 1.676$, $P = .207$.

Table 6. *The relationship of change in MRI antero-posterior spinal canal diameter (AP canal) and PM measures to change in diagnosis over 18 months*

	Initial Diagnosis							
	Asymptomatic			Back Pain			Stenosis	
		PM* (SD)	AP Canal†		PM‡	AP Canal§	PM**	AP Canal††
Asymptomatic	15	−0.20 (3.36)	−0.19 (1.87)	1	0.00 (—)	1.70 (—)	2	−9.50 (2.12)	2.65 (3.75)
Back pain	5	0.60 (0.89)	−0.62 (2.50)	14	2.93 (7.47)	−0.03 (2.80)	11	−0.55 (6.19)	−0.21 (2.31)
Stenosis	2	−3.50 (4.95)	−0.60 (0.28)	9	−0.33 (2.78)	0.17 (0.66)	16	1.00 (4.41)	0.69 (1.81)
Total	22	−0.32 (3.17	−0.33 (1.89)	24	1.58 (6.07)	0.11 (2.26)	29	−0.31 (5.59)	0.51 (2.16)

Analysis of variance:
*$F = 1.259$, $P = .306$; †$F = 0.109$, $P = .898$; ‡$F = 0.813$, $P = .457$; §$F = 0.257$, $P = .776$; **$F = 3.770$, $P = .037$; ††$F = 1.676$, $P = .207$.

is motor unit amplitude or duration changes, which were not measured precisely in this study. This study points out the need to demonstrate the reliability of these core measures to validate the use of EDX in general.

Some of these issues were circumvented by using PM. Tong et al [24] performed interrater reliability testing of PM showing good reproducibility ($r = 0.830$, $P = .041$) with monopolar needles. No valid method yet exists to quantify motor unit abnormalities in the paraspinal muscles. Because paraspinal muscle denervation was highly specific in earlier analyses [13], it is possible that the consequences of denervation—motor unit changes—would show a difference if we could only reproducibly measure these changes.

Because the purpose of the study was to follow changes over 18 months, people who were planning on undergoing surgery were excluded. Thus, the population does not include many persons with severe disease or disability.

The time between examinations was relatively short in the context of a lifetime. It is always possible that more obvious changes would occur over longer periods of time. The majority of subjects who were not tested at 18 months were excluded due to budgetary or scientific issues, not a decision that they made. Although the 9 subjects who underwent surgery did have more severe disease as measured by EMG and MRI (subgroup analysis not presented here), no statistically important bias was found between the dropouts as a group and the study completers. The dropout of 20 subjects due to funding did not likely bias the results but may have resulted in underpowering of the study to detect a change. At a minimum, the final population represents an important subgroup of persons with spinal stenosis, but the authors believe that it is representative of the natural history of the disease in general.

In the population of primarily younger persons including disc herniation and polyneuropathy in the study of Tong et al [24], a change of +7 or −9 was beyond the 95% confidence interval for repeatability. A tighter standard is likely more appropriate in the population presented here, who tended to have lower scores.

Signs of Denervation in Asymptomatic Persons

Aside from the foot, it is thought that limb fibrillations are not commonly found in asymptomatic persons [34,35]. This was generally true in the initial cohort in this study, but it is interesting to note that 2 subjects went on to have some fibrillations. It is accepted that a certain percentage of motor units in limb muscles are polyphasic [36,37]. Twenty-six percent of the asymptomatic group had more than 2/10 polys in a muscle—primarily the extensor hallicus longus and tensor fascia lata. These 2 muscles share the L5 nerve root, which is the most common root involved in radiculopathy. This pattern suggests that many asymptomatic older persons may have had an L5 radiculopathy (symptomatic or asymptomatic) in the past. Alternatively, an examiner bias due to expectation could be implicated. Regardless, because these people are without symptoms, clinicians are cautioned that the presence of polys in an L5 pattern does not necessarily mean that there is a clinically relevant disease.

It is now well established that asymptomatic persons—young and old—have "abnormal" spontaneous activity in the paraspinal muscles [29,34,38]. The current study confirms this. Three (13%) subjects had initial denervation sufficient to fall outside of the norms the authors had established a priori. The data suggest that a cutoff PM score of 6 is would more likely fit the traditional need for a 95% confidence interval. On the other hand, as reported elsewhere, the subgroup in which the physiatrist, radiologist, and spine surgeon all agreed independently were asymptomatic had PM scores of 0.67 ± 1.07 (SD), which compare well with the norms established for younger persons [30,33].

It has been hypothesized that the spontaneous paraspinal denervation that occurs in asymptomatic persons results in transient weakness and kinesthetic deficits, which in turn destabilize the spine, causing more ligamentous laxity and degenerative changes in the facet joints—perhaps leading to spinal stenosis, and also putting the posterior primary ramus at further risk [13]. No significant change was detected in the PM scores over time, and PM score did not relate to changes in the spinal measurements that were made. This study does not dismiss these theories. First, the radiologic measures did not include measures of facet joint hypertrophy or the nerve foramen. Second, the rate of change in bony structures may not allow for a difference in measurement in 18 months.

Because the multifidus muscle is only a few centimeters from the spinal canal, one would expect that individuals with posterior primary ramus damage would have re-

growth and repair resulting in a decrease in spontaneous activity over 18 months. The lack of a positive or negative trend in PM scores suggests an ongoing process of denervation and reinnervation, rather than a single injury in time. The authors can only speculate as to whether the substantial changes in a few individuals were related to an event or a symptom.

Is Mechanical Back Pain a Neurologic Disorder?

"Mechanical" back pain without leg pain is generally thought to spare the nerve roots. But this study shows evidence of denervation outside of the range of normal established in asymptomatic populations in a large minority of the subjects with back pain. Limb fibrillations were found in 21.4%, 39.3% had more than 2/10 polys, and 16.7% had PM scores greater than 4. In fact, there is scant literature on needle EMG findings in persons with back pain alone, and past research has been hampered by the lack of masking and the lack of established norms [39].

These findings suggest that some people thought clinically to have mechanical back pain actually have radicular involvement. However the relationships between paraspinal denervation, radiculopathy, back pain, and spinal stenosis are not straightforward. In addition to radiculopathy, myopathies, polyneuropathies, and a difficult-to-prove disorder involving entrapment of the posterior primary ramus [40] are among the many causes of isolated paraspinal denervation. Because 8 persons in the initial cohort of 150 were found to have unsuspected neuromuscular diseases and the EDX protocol here did not exhaustively examine for neuromuscular disease, an alternative explanation is the presence of a coincidentally occurring subclinical polyneuropathy or myopathy. Clinically, the authors have noted that the diffuse neuromuscular diseases have diffuse paraspinal denervation, but this remains to be proved.

Even in the absence of diffuse neuromuscular disease, the presence of denervation does not always imply pain. Even when the radiculopathy is from spinal stenosis, the pain and disability that a patient complains of may come from some other source, such as the facet joint. Research from this study presented elsewhere shows a lack of any good relationship between the clinical syndrome recognized as spinal stenosis and radiologic measures of spinal canal size [5,15]. At this point, the authors do not believe there is a "gold standard" for the diagnosis of the clinical syndrome commonly called spinal stenosis.

On the other hand, the clinical syndrome did relate to PM scores, which were 100% specific, although only moderately sensitive. Thus, the authors have proposed that the term "spinal stenosis" be reserved for anatomists until such time as a clinically relevant anatomical definition can be found [15]. However, this proposal challenges a century of medical tradition and textbooks, as well as the billing codes that justify interventional procedures.

The Neurophysiological Natural History of the Paraspinal Denervation Syndrome (aka, Clinical Spinal Stenosis)

As expected, the people with clinical evidence for spinal stenosis (a term the authors will use reluctantly for the purpose of clarity) had more denervation than the others—21.9% had limb fibrillations, 43.8% had more than 2 polys in 1 or more muscles, and 37.9% had paraspinal denervation outside of the range of normal. There was no significant progression of disease in this group over time. Instead, some subjects improved and some declined neurophysiologically. In fact, prior to this work there has never been, to the authors' knowledge, a single case report that prospectively follows electrodiagnostic findings in a person with any spinal disorder over time. The dogma regarding other spinal disorders can be questioned as well.

Clinical Implications

Previous work from the Michigan Spinal Stenosis Study has shown that MRI measures and the radiologist's clinical impression do not differentiate persons with symptoms of clinical stenosis from age-matched persons who have no symptoms. The electrodiagnostic examination is highly specific and moderately sensitive for the disorder, but only if PM is used and norms established in this study are used. The authors have shown elsewhere that the EMG is not predictive of outcome. The current report provides specific details of that conclusion, especially the relationships between paraspinal denervation and outcome. Previous limited information available on natural history also suggests that some subjects get better or worse over time clinically, with a general trend toward improvement [8,23,32,41-45].

This slow or absent progression means that there need not be any rush to surgical intervention. Delays related to numerous conservative treatments, functional rehabilitation, and lifestyle modifications are not going to put most patients at neurological risk. Electrodiagnosticians and clinicians who interpret their findings should be slow to use a positive EMG as evidence that the patient's problem is progressing. Instead, the factors that help a clinician best decide whether to leave symptoms alone, treat conservatively, or invoke surgical intervention remain poorly understood.

CONCLUSIONS

There is no significant trend toward neurological worsening or in persons without symptoms, with back pain, or with clinical spinal stenosis over 18 months. The presence of paraspinal denervation does not predict clinical, radiological, or electrodiagnostic decline in people with spinal stenosis over this period of time.

ACKNOWLEDGMENT

Study examiners and staff also included April M. Fetzer, DO, Marcus J. Harris, BS, Janis Huff, Richard W. Kendall, DO,

Allan Rowley, MD, Matthew J. Smith, MD, André Taylor, MD, and John A. Yarjanian, MD.

REFERENCES

1. Aronow WS. Management of peripheral arterial disease. Cardiol Rev 2005;13:61-68.
2. Goh KJ, Khalifa W, Anslow P, Cadoux-Hudson T, Donaghy M. The clinical syndrome associated with lumbar spinal stenosis. Eur Neurol 2004;52:242-249.
3. Haig AJ. Tong HC. Kendall R. The bent spine syndrome: myopathy + biomechanics = symptoms. Spine J 2006;6:190-194.
4. Boden S, McCowin PR, Davis DO, Dina TS, Mark AS, Wiesel S. Abnormal magnetic resonance scans of the lumbar spine in asymptomatic subjects. J Bone Joint Surg Am 1990;72:1178-1184.
5. Haig AJ, Tong HC, Yamakawa KSJ, et al. Spinal stenosis, back pain, or no symptoms at all? A masked study comparing radiologic and electrodiagnostic diagnoses to the clinical impression. Arch Phys Med Rehabil 2006;87:897-903.
6. Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic resonance imaging of the lumbar spine in people without back pain. N Engl J Med 1994;14:69-73.
7. Ertekin C, Sirin H, Koyuncuoglu HR, et al. Diagnostic value of electrical stimulation of lumbosacral roots in radiculopathies. Acta Neurol Scand 1994;90:26-33.
8. Jacobson RE. Lumbar stenosis. An electromyographic evaluation. Clin Orthop Relat Res 1976;115:68-71.
9. Johnsson KE, Rosen I, Uden A. Neurophysiologic investigation of patients with spinal stenosis. Spine 1987;12:483-487.
10. Kondo M, Matsuda H, Kureya S, Shimazu A. Electrophysiological studies of intermittent claudication in lumbar stenosis. Spine 1989;14:862-866.
11. Plastaras CT. Electrodiagnostic challenges in the evaluation of lumbar spinal stenosis. Phys Med Rehabil Clin North Am 2003;14:57-69.
12. Seppalainen AM, Alaranta H, Soini J. Electromyography in the diagnosis of lumbar spinal stenosis. Electromyogr Clin Neurophysiol 1981;21:55-66.
13. Haig AJ. Paraspinal denervation and the spinal degenerative cascade. Spine J 2002;2:372-380.
14. 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.
15. Haig AJ, Geisser ME, Tong HC, et al. Electromyographic and magnetic resonance imaging to predict lumbar stenosis, low-back pain, and no back symptoms. J Bone Joint Surg Am 2007;89:358-366.
16. Haig AJ, Tong HC, Yamakawa KS, et al. Predictors of pain and function in persons with spinal stenosis, low back pain, and no back pain. Spine 2006;31(25):2950-2957.
17. Bijur PE, Silver W, Gallagher EJ. Reliability of the visual analog scale for measurement of acute pain. Acad Emerg Med 2001;8:1153-1157.
18. Kopec JA, Esdaile JM, Abrahamowics M, et al. The Quebec Back Pain Disability Scale: conceptualization and development. J Clin Epidemiol 1996;49:151-161.
19. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. Pain 1975;1:277-279.
20. Ohlund C, Eek C, Palmblad S, Areskous B, Nachemson A. Quantified pain drawing in subacute low back pain. Validation in nonselected outpatient industrial sample. Spine 1996;21:1021-1031.
21. Tait RC, Chibnall JT, Krause S. The Pain Disability Index: psychometric properties. Pain 1990;40:171-182.
22. Haig AJ, Weiner JB, Tew J, Quint D, Yamakawa K. The relation among spinal geometry on MRI, paraspinal electromyographic abnormalities, and age in person referred for electrodiagnostic testing of low back symptoms. Spine 2002;27:1918-1925.
23. Porter RW. Spinal stenosis and neurogenic claudication. Spine 1996; 21:2046-2052.
24. Tong HC, Young IA, Koch J, Haig AJ, Yamakawa KS, Wallbom A. Paraspinal electromyography that compares concentric with monopolar needles: a blinded study. Am J Phys Med Rehabil 2003;82:917-924.
25. Haig AJ, Yamakawa KSJ, Kendall R, Miner JA, Parres C, Harris M. masking in electrodiagnostic medicine research. Am J Phys Med Rehabil 2006;85:475-481.
26. Dillingham TR, Lauder TD, Andary M, Shannon S. Identifying lumbosacral radiculopathies: an optimal electromyographic screen. Am J Phys Med Rehabil 2000;79:496-503.
27. Daube JR. AAEM Minimonograph #11: Needle examination in clinical electromyography. Muscle Nerve 1991;4:685-700.
28. Haig AJ, LeBreck DB, Powley SG. Paraspinal mapping: quantified needle electromyography of the paraspinal muscles in persons without low back pain. Spine 1995;20:715-721.
29. Haig AJ. clinical experience with paraspinal mapping, I: neurophysiology of the paraspinal muscles in various spinal disorders. Arch Phys Med Rehabil 1997;78:1177-1184.
30. Haig AJ. Clinical experience with paraspinal mapping. II: a simplified technique that eliminates three fourths of needle insertions. Arch Phys Med Rehabil 1997;78:1185-1190.
31. Haig AJ. Paraspinal mapping. American Association of Neuromuscular and Electrodiagnostic Medicine course handout. Revised 2005. Rochester, MN: AANEM. Available at: www.aanem.org.
32. Haig AJ, Yamakawa K, Hudson DM. Paraspinal electromyography in high lumbar and thoracic lesions. Am J Phys Med Rehabil 2000;79(336-342.
33. Tong HC, Haig AJ, Yamakawa KSJ, Miner JA. Paraspinal electromyography: age-correlated normative valves in asymptotic subjects. Spine 2005;30(17):E499-E502.
34. Dumitru D, Diaz CA, King JC. Prevalence of denervation in paraspinal and foot intrinsic musculature. Am J Phys Med Rehabil 2001;80:482-490.
35. Gatens PF, Saeed MA. Electromyographic findings in the intrinsic muscles of normal feet. Arch Phys Med Rehabil 1982;63:317-318.
36. Dumitru D, Amato A, Zwarts MJ. Electrodiagnostic Medicine. 2nd ed. Philadelphia, PA: Hanley and Belfus; 2001.
37. Kimura J. Electrodiagnosis in Diseases of Nerve and Muscle: Principles and Practice. New York, NY: Oxford University Press; 2001.
38. Date ES, Mar EY, Bugola MR, Teraoka JK. The prevalence of lumbar paraspinal spontaneous activity in asymptomatic subjects. Muscle Nerve 1996;19:350-354.
39. Durette MR, Rodriquez AA, Agre JC, Silverman JL. Needle electromyographic evaluation of patients with myofascial or fibromyalgic pain. Am J. Phys Med Rehabil 1991;70:154-156.
40. Atlas S, Chang Y, Wu Y, Keller R, Singer D. Predictors of longitudinal outcomes for patients with lumbar spinal stenosis followed over a 10 year period. International Society for the Study of the Lumbar Spine; New York, NY; May 12, 2005; poster 106.
41. Fisher MA, Kacr D, Houchin J. Electrodiagnostic examination, back pain, and entrapment of posterior rami. Electromyogr Clin Neurophysiol 1985;25:187.
42. Benoist M. The natural history of lumbar degenerative spinal stenosis. J Bone Spine Rev Rhumat 2002;69:450-457.
43. Johnsson KE, Rosen I, Uden A. The effect of decompression on the natural course of spinal stenosis. A comparison of surgically treated and untreated patients. Spine 1991;16:615-619.
44. Johnsson KE, Rosen I, Uden A. The natural course of lumbar spinal stenosis. Clin Orthop 1992; 279:82-86.
45. Porter RW, Hibbert C, Evans C. The natural history of root entrapment syndrome. Spine 1984;9:418-421.