CHAD COOK, PT, PhD, MBA, OCS, FAAOMPT[1] • **MATTHEW ROMAN,** PT, FAAOMPT[2] • **KATHLEEN M. STEWART,** PT, DPT[3]
LINDA GRAY LEITHE, MD[4] • **ROBERT ISAACS,** MD[5]

Reliability and Diagnostic Accuracy of Clinical Special Tests for Myelopathy in Patients Seen for Cervical Dysfunction

- **STUDY DESIGN:** Case control study.
- **BACKGROUND:** Myelopathy is a clinical diagnosis based largely on initial examination findings during a clinical screen, followed by imaging verification of cord injury or compression. At present, few studies have examined the reliability and diagnostic accuracy of clinical examination measures.
- **OBJECTIVES:** To determine the reliability and diagnostic accuracy of neurological tests associated with the diagnosis of myelopathy.
- **METHODS AND MEASURES:** Reliability and diagnostic accuracy of 7 frequently used tests and measures and subjective findings associated with myelopathy were examined on consecutive patients with cervical pain. Interrater reliability and diagnostic accuracy values, including posttest probability, based on a pretest probability of 40%, were calculated for each test and for combinations of tests and measures.
- **RESULTS:** Four of the 7 diagnostic tests were found to have a substantial interrater reliability. None of the single or clusters of tests yielded low negative likelihood ratios. Of the individual tests, the Babinski sign demonstrated the highest positive likelihood ratio (LR+, 4.0; 95% CI: 1.1-16.6) and posttest probability (73%) for diagnosis, but yielded only a moderate negative likelihood ratio (LR–, 0.7; 95% CI: 0.6-0.9). Combinations of tests did not yield improved accuracy values over single test results.
- **CONCLUSION:** This study demonstrated that 4 of 7 tests used to screen for myelopathy offered substantial levels of interrater agreement when used on individuals with cervical dysfunction. None of the tests when performed individually or in combinations are effective for screening; however, the Babinski sign did alter posttest probability more significantly than combinations of test findings.
- **LEVEL OF EVIDENCE:** Diagnosis, Level 2b.

J Orthop Sports Phys Ther 2009;39(3):172-178. doi:10.2519/jospt.2009.2938

- **KEY WORDS:** *cervical spine, diagnostic test, neck, neurological screen, validity*

Cervical spine myelopathy resulting from sagittal narrowing of the spinal canal and compression of the spinal cord is present in 90% of individuals by the seventh decade of life.[32] Although the exact prevalence is unknown,[32] cervical spine myelopathy is recognized as the most common form of spinal cord dysfunction in individuals over the age of 55.[39] Cord compression may occur from (1) osteophytes secondary to degeneration of intervertebral joints, (2) stiffening of connective tissues, such as the ligamentum flavum at the dorsal aspect of the spinal canal, which can impinge on the cord by "buckling" when the spine is extended, (3) degeneration of intervertebral discs together with subsequent bony changes, and (4) other connective tissue changes.[47]

Diagnosis of myelopathy is challenging, particularly in the early stages of the condition, as symptoms may present as hyperreflexia,[27,32,33,38] clumsiness in gait,[3,27,32,33,38] neck stiffness,[10,32,33,38] shoulder pain,[11] paresthesia in 1 or both arms or hands,[22] or radiculopathic signs.[24,32,33,38] When suspected after pertinent clinical examination findings, a diagnosis of myelopathy is confirmed or refuted by magnetic resonance imaging (MRI). Myelopathy may lead to anterior-posterior width reduction of the spinal cord, cross-sectional evidence of cord compression, or obliteration of the subarachnoid space.[6,19,34,35,46] At present, there are no definitive objective findings on MRI consistently described by radiologists that are reflective of myelopathy, with the exception of myelomalacia (identified through signal intensity changes to the cord). Signal intensity changes have been described as the most appropriate "gold standard" for confirmation of a spinal cord compression myelopathy.[1,9,21,25,28-30,37]

The clinical examination for myelopathy includes the use of Hoffmann's

[1]Associate Professor, Director of Outcomes Measurement, Center of Excellence in Surgical Outcomes, Department of Surgery, Duke University, Durham, NC. [2]Clinical Director of Physical Therapy and Occupational Therapy, Department of Physical Therapy and Occupational Therapy, Duke University, Durham, NC. [3]Clinical Resident, Cardiovascular and Pulmonary Physical Therapy Residency, Duke University, Durham, NC. [4]Associate Professor, Division of Radiology, Duke University, Durham, NC. [5]Director of Spine Surgery, Department of Neurosurgery, Duke University, Durham, NC. The protocol of this study was approved by The Institutional Review Board of Duke University Health System. Address correspondence to Dr Chad Cook, 2200 W Main Street, Suite A-210, Duke University Medical Center, Box 104002, Durham, NC 27708. E-mail: chad.cook@duke.edu

TABLE 1 Clinical Special Tests Defined

Test Name	Operational Procedure	Positive Test
Hoffmann's sign	With the patient in standing or sitting, the clinician stabilizes the proximal interphalangeal joint of the middle finger and applies a stimulus to the middle finger by "flicking" the fingernail between his thumb and index finger into a flexed position.[23]	Adduction of the thumb and flexion of the fingers
Deep tendon reflex tests	In Biceps tendon testing, the patient assumes a sitting position while the clinician places the patient's slightly supinated forearm on his own forearm, assuring relaxation. The clinician's thumb is placed on the patient's biceps tendon and he strikes his own thumb with quick strikes of a reflex hammer. In Triceps tendon testing, the sitting patient's elbow is flexed passively via shoulder elevation to approximately 90°. The clinician then places his thumb over the distal aspect of the triceps tendon and applies a series of quick strikes of the reflex hammer to his own thumb.	Hyperreflexia
Inverted supinator sign	With the patient in a seated position, the clinician places the patient's slightly pronated forearm on his forearm to assure relaxation. The clinician applies a series of quick strikes near the styloid process of the radius at the attachment of the brachioradialis tendon. The test is performed in the same manner as a brachioradialis tendon reflex test.	Finger flexion or slight elbow extension
Suprapatellar quadriceps test	With the patient in sitting with his or her feet off the ground, the clinician applies quick strikes of the reflex hammer to the suprapatellar tendon.	Hyperreflexive knee extension
Hand withdrawal reflex	With the patient in sitting or standing, the clinician grasps the patient's palm and strikes the dorsum of the patient's hand with a reflex hammer.	Abnormal flexor response
Babinski sign	With the patient in supine, the clinician supports the patient's foot in neutral and applies stimulation to the plantar aspect of the foot (typically lateral to medial from heel to metatarsal) with the blunt end of a reflex hammer.	Great toe extension and fanning of the second through fifth toes
Clonus	With the patient in sitting with his or her feet off the ground, the clinician applies a quick stretch to the Achilles tendon via rapid passive dorsiflexion of the ankle.	Patient's ankle "beats" in and out of dorsiflexion for at least 3 beats

test,[16,24,44,47] deep tendon reflex testing,[15,47] inverted supinator sign,[17] suprapatellar quadriceps reflex testing,[14] hand withdrawal reflex testing,[15] Babinski sign,[4,14,20,23,41,47] and clonus.[47] To our knowledge, none of these tests have been assessed for reliability. Only the Hoffmann's test, the Babinski sign, and hand withdrawal reflex test have been investigated for diagnostic accuracy with studies that have exhibited methodological weaknesses and provided inconsistent results.[12,13] Further, most of the tests when performed in isolation fail to provide significant screening or diagnostic power, thus, it has been suggested that tests should be used in combinations or clusters during screening for myelopathy.[13]

While some evidence exists for the effective treatment of early myelopathic changes via conservative physical therapy interventions (ie, traction and thoracic manipulation),[8] conclusive evidence for the effectiveness of surgical intervention for myelopathy suggests surgery should be pursued when testing is positive.[18] Given that clinical findings of myelopathy are used during both screening and diagnosis, it is imperative that the clinical tests demonstrate strong reliability, adequate sensitivity for screening, and appropriate diagnostic value for identifying cervical myelopathy. Consequently, the purpose of this study was to assess the interrater reliability and diagnostic accuracy of clinical neurological tests (used singularly and in clusters) and subjective findings associated with a MRI-confirmed diagnosis (using signal intensity changes) of cervical spine myelopathy.

METHODS

Procedures for this study followed the Standards for the Reporting of Diagnostic (STARD) guidelines set forth by Bossuyt et al.[5] Briefly, the STARD standards are used to improve reporting processes for diagnostic accuracy studies and involve 25 items associated with topics germane in a typical case control design. Topics are oriented toward description of participant, statistical analysis, results, and conclusions of findings. Prior to the design of the study we used the STARD outlines to capture the majority of the procedural requirements suggested by the standards to improve the reporting of findings at completion of the study.

Subjects

The study protocol was approved by The Institutional Review Board of Duke University Health System. The 51 consecutive patients recruited for this study were identified at the neurosurgery practice at Duke University's Spine Clinics, with cervical pain as the patient's primary complaint. For inclusion into the study, patients were English-speaking, older than 18 years of age, had undergone MR imaging, and consented to participate in the study.

Procedures

Each patient provided the following information: age, race/ethnicity, and gender, in addition to answering yes/no subjective questions related to current neck pain, progressive experience of clumsiness or numbness in the hands, and progressive clumsiness during walking. Standard practice of the clinic involves independent examination from a neurosurgeon and physical therapist during every initial consultation. Prior

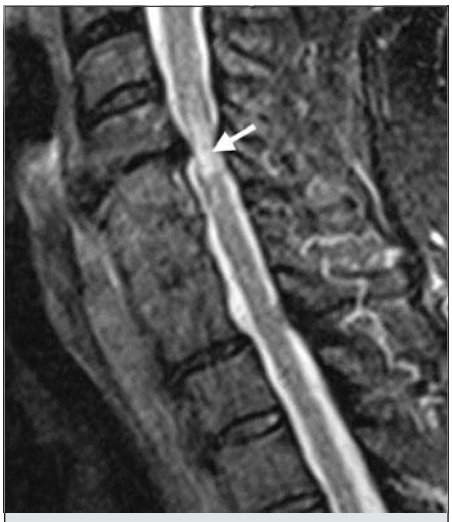

FIGURE 1. Magnetic resonance image (sagittal view) of patient with myelomalacia. Arrow identifies the signal intensity change associated with damage to the spinal cord.

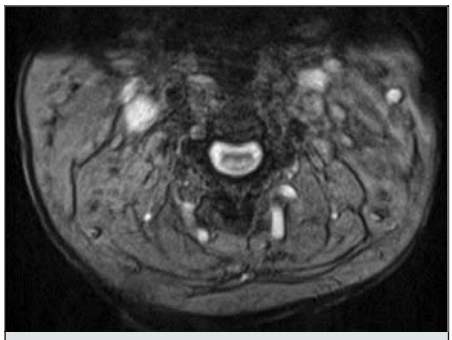

FIGURE 2. Magnetic resonance image (cross sectional view) of cervical spine at level of documented myelomalacia. Signal intensity changes (whitening of spinal cord) demonstrates presence of myelomalacia.

to the initiation of the study, the 2 clinicians standardized the following clinical special tests used in the study (**TABLE 1**): Hoffmann's test, deep tendon reflex testing, the inverted supinator sign, suprapatellar quadriceps reflex testing, hand withdrawal reflex testing, Babinski sign, and clonus, during a 2-hour educational session. Both clinicians agreed upon what constituted a positive test and how to perform and standardize each testing method. The order of clinician examination was not standardized (neither the physical therapist nor the neurosurgeon were routinely first to examine), and the clinical special tests were randomized for each patient.

For each patient, both clinicians separately completed a form indicating positive, negative, or inconclusive results for each of the aforementioned 7 clinical special tests. The clinicians were blinded to one another's results, the patients' MR images, and the true diagnosis of the patient. One tester was a neurosurgeon specializing in spine surgery with 15 years of surgical experience, whereas the second tester was a physical therapist with 18 years of experience as a clinician and extensive training in screening of the spine.

All MR images for each patient were read by a radiologist with over 19 years of experience, specializing in reading images of spinal cord-related orthopedic injuries. The radiologist was blinded to the clinical findings and diagnosis of the patient, and was responsible for documenting the presence of signal intensity changes associated with myelomalacia found on T2-weighted MRI (**FIGURES 1 AND 2**). These imaging-related findings have been previously used to confirm the presence of cord compression associated with myelopathy.[1,8,21,25,28-30,37]

Data Analysis

Descriptive statistics were calculated for the group of patients. Interrater reliability and diagnostic accuracy for each of the clinical diagnostic tests were determined using STATA 9.0 (StataCorp LP, College Station, TX). In instances of disagreement between testers, the results from the neurosurgeon were used for the diagnostic accuracy calculations. Kappa values, including level of significance and 95% confidence intervals (CIs), were used to assess interrater agreement for each clinical test. Values were categorized based on the benchmarked criteria articulated by Landis and Koch[26]: poor, <0.00; slight, 0.01-0.2; fair, 0.21-0.40; moderate, 0.41-0.60; substantial, 0.61-0.80; and almost perfect, 0.81-1.00. Sensitivity, specificity, and positive and negative likelihood ratios (LR+ and LR–), with 95% CIs, were calculated for the clinical examination and subjective response findings using a reference standard that required the presence of signal intensity changes confirming myelomalacia on MRI. Posttest probabilities of positive and negative findings were also calculated using the prevalence of myelopathy in this study (calculated as the frequency of patients who met the MRI criteria for diagnosis of myelopathy) as the pretest probability estimate.

Clusters of clinical examination and subjective response findings were also tabulated for combinations that would improve diagnostic accuracy. We examined all possible combinations to capture those with the strongest diagnostic accuracy measures and included posttest probability values.

RESULTS

Descriptive Data

Fifty-one patients, from February 2007 through December 2007, satisfied the inclusion criteria (**TABLE 2**). Usable data (high-quality MRIs) were obtained on 45 of the 51 individuals. In all situations, the MRIs were taken within 1 year of the clinical examination for this study, with the majority of images dating between 2 weeks and 3 months. Ages ranged from 33 to 82 years, with a mean (SD) age of 52 (13.4) years. Eighty-two percent of patients identified themselves as white, 7% identified themselves as black, and 10% checked "other." Of the 45 patients, 25 (55.6%) were women and 20 were men. Final recorded diagnoses varied among the patients and included radiculopathy, myelopathy, and cervical strains.

At the time of recruitment, patients were asked to complete a subjective inventory of their signs and symptoms (**TABLE 2**). Thirty-three of 45 patients (73.3%) indicated the presence of neck pain during the clinical visit. Loss of hand dexterity was reported for 29 of 45 patients (64.4%). Report of medial numbness of both hands was present in 16 patients (35.6%). Clumsiness during walking was reported by 18 of the 45 patients

(40.0%). Duration of symptoms was not recorded.

Kappa values, 95% CIs, and strength-of-agreement ratings are reported in **TABLE 3**. The prevalence of positive findings varied by test but yielded percentages as low as 8% for clonus and as high as 61% for the suprapatellar reflex test. Four of the 7 diagnostic tests were found to have a substantial strength of agreement: Hoffmann's sign (κ = 0.76; 95% CI: 0.56-0.96), deep tendon reflexes (κ = 0.73; 95% CI: 0.5-0.95), suprapatellar tendon reflex (κ = 0.68; 95% CI: 0.46-0.89), and clonus (κ = 0.66; 95% CI: 0.03-0.99). The remaining 3 demonstrated moderate interrater agreement: inverted supinator sign (κ = 0.52; 95% CI: 0.26-0.78), hand withdrawal reflex (κ = 0.55; 95% CI: 0.34-0.82), and Babinski sign (κ = 0.56; 95% CI: 0.24-0.89).

Diagnostic accuracy of the tests and measures and subjective questions for identifying those with cervical myelopathy was calculated based on the data of 45 patients. Forty percent of the patients met the MR criteria of myelopathy of the cervical spine, most frequently involving the C5-C6 and C6-C7 levels. **TABLE 4** reflects the diagnostic accuracy values of the subjective questions. None of the questions demonstrated significant diagnostic accuracy, as indicated by the inclusion of 1.0 within the 95% CI for all the LR+ and LR– values (**TABLE 4**). During evaluation of clinical tests there were 2 equivocal findings (an inconclusive finding, neither

TABLE 2 Sample Demographics (n = 45)

Age (y)	
Mean (SD)	52 (13.4)
Range	33-82
Race/Ethnicity	
White	82%
Black	7%
Latino	0%
Other	10%
Gender	
Male	41%
Female	59%
Questions*	
1. Do you currently experience neck pain?	33, 6, 6
2. Have you experienced a loss of dexterity or, have you experienced clumsiness in your hands?	29, 10, 6
3. Do you experience numbness on the little finger-side (medial) of your hand?	16, 21, 8
4. Do you experience clumsiness during walking, more so than what you experienced a few years ago?	18, 20, 7

** Response values are number who answered yes, no, no response.*

TABLE 3 Clinician Percent Agreement, Kappa Values, and 95% Confidence Intervals (CIs)

Test	Clinician Agreement	Kappa Value	95% CI	SE	Positive Findings	Strength of Agreement
Hoffmann's sign	89%	0.76	0.56-0.96	0.10	45%	Substantial
Deep tendon reflexes	89%	0.73	0.50-0.95	0.10	31%	Substantial
Inverted supinator sign	78%	0.52	0.26-0.78	0.13	36%	Moderate
Suprapatellar tendon reflex	84%	0.68	0.46-0.89	0.10	61%	Substantial
Hand withdrawal reflex	80%	0.55	0.34-0.82	0.12	42%	Moderate
Babinski sign	89%	0.56	0.24-0.89	0.16	18%	Moderate
Clonus	98%	0.66	0.03-0.99	0.32	8%	Moderate

Abbreviation: SE, standard error.

TABLE 4 Diagnostic Accuracy Calculations for Subjective Response Findings*

Test	Sensitivity (%)	Specificity (%)	LR+	LR–
Report of current neck pain	93 (80-99)	18 (9-22)	1.1 (0.9-1.3)	0.4 (0.06-2.2)
Report of dexterity loss	73 (57-88)	27 (16-37)	1.1 (0.7-1.4)	0.9 (0.3-2.7)
Report of numbness in the hands	57 (37-74)	67 (53-78)	1.7 (0.8-3.4)	0.6 (0.3-1.2)
Report of clumsiness during gait	53 (35-71)	52 (39-65)	1.1 (0.6-2.0)	0.9 (0.4-1.7)

** Values in parentheses are 95% confidence intervals.*
Abbreviations: LR+, positive likelihood ratio; LR–, negative likelihood ratio.

TABLE 5 Diagnostic Accuracy Calculations for the Clinical Tests*

Test	Sensitivity (%)	Specificity (%)	LR+	LR–	Posttest Probability LR+	Posttest Probability LR–
Hoffmann's sign	44 (28-58)	75 (63-86)	1.8 (0.8-4.1)	0.7 (0.5-1.1)	55%	32%
Deep tendon reflexes	44 (28-59)	71 (59-82)	1.5 (0.7-3.4)	0.8 (0.5-1.2)	50%	35%
Inverted supinator sign	61 (44-74)	78 (65-88)	2.8 (1.2-6.4)	0.5 (0.3-0.9)	65%	25%
Suprapatellar tendon reflex	56 (39-72)	33 (22-46)	0.8 (0.5-1.3)	1.3 (0.6-2.8)	35%	46%
Hand withdrawal reflex	41 (25-58)	63 (51-75)	1.1 (0.5-2.3)	0.9 (0.6-1.5)	42%	38%
Babinski sign	33 (19-41)	92 (81-98)	4.0 (1.1-16.6)	0.7 (0.6-0.9)	73%	32%
Clonus	11 (3-16)	96 (90-99)	2.7 (0.4-20.1)	0.9 (0.8-1.1)	64%	38%

** Values in parentheses for sensitivity, specificity, and positive and negative likelihood ratios are 95% confidence intervals.*
Abbreviations: LR+, positive likelihood ratio; LR–, negative likelihood ratio.

TABLE 6 Diagnostic Accuracy Calculations for the Best Combinations of Clinical Tests (Clusters)*

Test	Sensitivity (%)	Specificity (%)	LR+	LR–	Posttest Probability LR+	Posttest Probability LR–
Any 1 of 2 positive IVS and suprapatellar reflex	67 (51-82)	33 (21-45)	1.0 (0.6-1.4)	1.0 (0.4-2.2)	40%	40%
2 of 2 positive Babinski sign and IVS	50 (34-63)	75 (62-86)	2.0 (0.9-4.5)	0.7 (0.4-1.1)	57%	32%
2 of 3 positive Babinski sign, Hoffmann's reflex, and IVS	50 (34-62)	83 (71-92)	3.0 (1.2-8.1)	0.6 (0.4-0.9)	67%	29%
3 of 4 Hoffmann's, suprapatellar reflex, Babinski, and IVS	44 (29-57)	83 (72-92)	2.7 (1.1-7.4)	0.7 (0.5-0.9)	64%	32%

** Values in parentheses for sensitivity, specificity, and positive and negative likelihood ratios are 95% confidence intervals.*
Abbreviations: IVS, inverted supinator sign; LR+, positive likelihood ratio; LR–, negative likelihood ratio.

positive nor negative); 1 for the inverted supinator sign and 1 for the hand withdrawal reflex. In both cases, the equivocal finding was tabulated as a negative finding. Only the inverted supinator sign and the Babinski sign demonstrated significant diagnostic accuracy. The Babinski sign demonstrated the highest LR+, with a value of 4.0 (95% CI: 1.1-16.6). Using the pretest probability of myelopathy of 40%, the posttest probability for a positive finding was most significantly altered using the Babinski test (posttest probability, 73%). The inverted supinator sign had the smallest LR– value of 0.5 (95% CI: 0.3-0.9); therefore, a negative inverted supinator sign altered the posttest probability of myelopathy to the greatest extent (posttest probability, 25%). The accuracy statistics for the individual clinical test findings are outlined in **TABLE 5**.

Combining clinical examination tests (**TABLE 6**) provided marginal improvements in the diagnostic accuracy of the clinical tests. Similarly, any combination of patient self-report of numbness, dexterity loss, and clumsiness provided very little value. Any positive finding during testing of the suprapatellar quadriceps reflex testing or inverted supinator sign only slightly improved the sensitivity, but weakened the posttest probability of diagnosis. The combination of a positive Babinski and inverted supinator sign resulted in a LR+ of 2.0 and a posttest probability of 57%. A combination of 3 of 4 positive tests among Hoffmann's, suprapatellar reflex, Babinski, and the inverted supinator sign resulted in improved diagnostic values but was not superior to the combination of any 2 positive findings among Babinski sign, inverted supinator sign, and Hoffmann's reflex (posttest probability, 67%). None of the clusters provided posttest probabilities for negative findings that were superior to individual findings. All other combinations resulted in lower diagnostic accuracy values than the Babinski sign used in isolation.

DISCUSSION

There are a number of clinically relevant findings in this study. First, to our knowledge, this is the only study that has examined the interrater reliability of clinical special tests performed with the purpose of diagnosis of cervical myelopathy. Our results indicate moderate to substantial strength of agreement between 2 experienced clinicians when evaluating 7 commonly used clinical special tests. The high level of clinician agreement revealed in this study provides initial support for use of these tests and measures. For 4 of 7 tests we identified substantial agreement and for the 3 remaining tests moderate agreement. Kappa statistic provides a chance-corrected agreement, thus accounting for "guesses," and has been questioned in the past for penalizing selected test values although these tests yield value in clinical practice.[7] This overcorrection of Kappa is commonly referred to as Kappa's "base rate prob-

lem"[6,42,43,45] and results in artificially low values despite high levels of percentage of agreement. The prevalence of positive findings was moderately dispersed among our clinical tests, thus reducing the need for a prevalence-adjusted kappa, a technique occasionally used to compensate for high or low prevalence.[40]

Some of the clinical tests used to diagnose myelopathy demonstrated statistically significant diagnostic accuracy (namely, the Babinski sign and inverted supinator sign), but none substantially modified the posttest probability of myelopathy. Our findings support previous reports that no tests for myelopathy are inherently sensitive, as evidenced by the insufficiently small LR– values,[13,21] thus call into question the usefulness of these tests in screening for myelopathy. In addition, our results did not support the ability of these tests to confirm a diagnosis of myelopathy as evidenced by the insufficient LR+ values.

In traditional clinical practice, examination tests are used initially to identify the presence of myelopathy and an MRI is used to confirm the presence of the clinical finding. For the initial stages, this practice requires tests and measures with low LR– values. Our findings indicate that no single test or subjective finding demonstrated sufficiently low LR– values for screening purposes and that there were no combinations of testing which improved these values. Recently, Cook et al[13] suggested the use of combinations of tests to adjust for the inherently poor LR– values of myelopathic screening tests. Our findings indicate that combinations of tests do not substantially improve LR– values and may not be useful in clinical practice.

In addition to identifying the initial presence of myelopathy, selected tests are used in concert with appropriate MRI findings to justify a clinical diagnosis of cervical myelopathy. We elected to use myelomalacia as the reference standard defining the presence of myelopathy on MRI because it is a definitive measure of cord damage.[1,9,21,25,28-30,37] Consequently, poor association of a clinical finding in the presence of myelomalacia would suggest a weakness of the clinical measure for diagnosis of cervical myelopathy. Our findings suggest that the most accurate finding to confirm the presence of myelopathy on MRI was the Babinski sign in isolation. Combinations of findings did not improve the diagnostic accuracy of the tests at a rate greater than the standalone test of the Babinski sign.

The selection of myelomalacia (signal intensity changes on the cord) as a reference standard served to provide a highly specific confirmation of myelopathy.[1,7,21,25,28-30,37] An MRI is the suggested reference standard with reference to imaging, because it expresses the amount of compression placed on the spinal cord,[2,37] and demonstrates relatively high levels of sensitivity (79%-95%) and specificity (82%-88%) in identifying selected abnormalities such as space occupying tumors,[34] disc herniation,[36,48] and ligamentous ossification.[31] Presence of signal intensity changes are prominent in patients with chronic myelopathic changes.[37] Thus, the clinical examination findings of the tests and measures used in this study may actually exhibit lower LR– if a less specific parameter of the MRI was used. Further, it is worth noting that patients examined in this study were referred to a neurosurgeon for surgical consult, suggesting a higher degree of severity of symptoms than that of patients screened in a typical clinical practice. In a typical clinical setting, these tests may actually yield lower sensitivities.

Limitations

There are a number of limitations to our study. Firstly, we calculated reliability between two very experienced clinicians, thus our moderate to substantial levels of agreement may be reflective of their skill level, more so than the intrinsic functions of the tests and measures. Secondly, as stated, the patients of this study were older and had been referred to a neurosurgeon for a surgical consult. This suggests that many patients had more severe signs and symptoms than those expected in a standard population of a typical clinical practice. Although the actual prevalence of myelopathy is currently unknown, our pretest probability was likely higher than what would be the case in a traditional clinic and may actually inflate our posttest probability findings. Lastly, there was inconsistent time between the clinical examination and the imaging examination, which could have also influenced the clinical examination findings.

CONCLUSION

Our findings suggest that standard clinical tests for myelopathy may exhibit moderate to substantial reliability among skilled clinicians. However, routine tests and measures used in screening for myelopathy exhibit low to moderate sensitivity and LR– values, indicating that the tests may fail to adequately rule out the presence of myelopathy during the clinical screen. ◉

KEY POINTS

FINDINGS: Seven commonly used clinical tests to screen or diagnose myelopathy demonstrate moderate to substantial reliability between 2 highly skilled clinicians. No single test, cluster of clinical tests, or subjective findings demonstrated high sensitivity and low LR– for appropriate screening utility of patients with MR substantiated myelopathy.

IMPLICATIONS: While reliable, the tests and subjective findings used in this study were not helpful in the diagnosis of myelopathy.

CAUTION: Reliability data were obtained from 2 experienced clinicians and diagnostic data were obtained on older patients seen for a surgical consult, which may not reflect the typical clinicians and patient population.

ACKNOWLEDGEMENTS: *This study was partially funded by the Orthopaedic Physical Therapy Products Grant from the American Academy of Orthopaedic Manual Physical Therapists, awarded 2005.*

REFERENCES

1. Al-Mefty O, Harkey LH, Middleton TH, Smith RR, Fox JL. Myelopathic cervical spondylotic lesions demonstrated by magnetic resonance imaging. *J Neurosurg.* 1988;68:217-222.
2. Batzdorf U, Flannigan BD. Surgical decompressive procedures for cervical spondylotic myelopathy. A study using magnetic resonance imaging. *Spine*. 1991;16:123-127.
3. Bednarik J, Kadanka Z, Dusek L, et al. Presymptomatic spondylotic cervical cord compression. *Spine*. 2004;29:2260-2269.
4. Berger JR, Fannin M. The "bedsheet" Babinski. *South Med J.* 2002;95:1178-1179.
5. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med.* 2003;138:40-44.
6. Braga-Baiak A, Shah A, Pietrobon R, Braga L, Neto AC, Cook C. Intra- and inter-observer reliability of MRI examination of intervertebral disc abnormalities in patients with cervical myelopathy. *Eur J Radiol.* 2008;65:91-98. http://dx.doi.org/10.1016/j.ejrad.2007.04.014
7. Brennan PF, Hays BJ. The kappa statistic for establishing interrater reliability in the secondary analysis of qualitative clinical data. *Res Nurs Health.* 1992;15:153-158.
8. Browder DA, Erhard RE, Piva SR. Intermittent cervical traction and thoracic manipulation for management of mild cervical compressive myelopathy attributed to cervical herniated disc: a case series. *J Orthop Sports Phys Ther.* 2004;34:701-712. http://dx.doi.org/10.2519/jospt.2004.1519
9. Chen CJ, Lyu RK, Lee ST, Wong YC, Wang LJ. Intramedullary high signal intensity on T2-weighted MR images in cervical spondylotic myelopathy: prediction of prognosis with type of intensity. *Radiology*. 2001;221:789-794.
10. Chiles BW, 3rd, Leonard MA, Choudhri HF, Cooper PR. Cervical spondylotic myelopathy: patterns of neurological deficit and recovery after anterior cervical decompression. *Neurosurgery*. 1999;44:762-769; discussion 769-770.
11. Clark CA, Barker GJ, Tofts PS. Magnetic resonance diffusion imaging of the human cervical spinal cord in vivo. *Magn Reson Med.* 1999;41:1269-1273.
12. Cook CE, Hegedus E. *Orthopedic Physical Examination Tests: An Evidence-Based Approach.* Upper Saddle River, NJ: Prentice-Hall; 2008.
13. Cook CE, Hegedus E, Pietrobon R, Goode A. A pragmatic neurological screen for patients with suspected cord compressive myelopathy. *Phys Ther.* 2007;87:1233-1242. http://dx.doi.org/10.2522/ptj.20060150
14. de Freitas GR, Andre C. Absence of the Babinski sign in brain death: a prospective study of 144 cases. *J Neurol.* 2005;252:106-107. http://dx.doi.org/10.1007/s00415-005-0605-6
15. Denno JJ, Meadows GR. Early diagnosis of cervical spondylotic myelopathy. A useful clinical sign. *Spine*. 1991;16:1353-1355.
16. Emery SE, Bohlman HH, Bolesta MJ, Jones PK. Anterior cervical decompression and arthrodesis for the treatment of cervical spondylotic myelopathy. Two to seventeen-year follow-up. *J Bone Joint Surg Am.* 1998;80:941-951.
17. Estanol BV, Marin OS. Mechanism of the inverted supinator reflex. A clinical and neurophysiological study. *J Neurol Neurosurg Psychiatry.* 1976;39:905-908.
18. Fujiwara K, Yonenobu K, Ebara S, Yamashita K, Ono K. The prognosis of surgery for cervical compression myelopathy. An analysis of the factors involved. *J Bone Joint Surg Br.* 1989;71:393-398.
19. Fukushima T, Ikata T, Taoka Y, Takata S. Magnetic resonance imaging study on spinal cord plasticity in patients with cervical compression myelopathy. *Spine*. 1991;16:S534-538.
20. Ghosh D, Pradhan S. "Extensor toe sign" by various methods in spastic children with cerebral palsy. *J Child Neurol.* 1998;13:216-220.
21. Glaser JA, Cure JK, Bailey KL, Morrow DL. Cervical spinal cord compression and the Hoffmann sign. *Iowa Orthop J.* 2001;21:49-52.
22. Good DC, Couch JR, Wacaser L. "Numb, clumsy hands" and high cervical spondylosis. *Surg Neurol.* 1984;22:285-291.
23. Hindfelt B, Rosen I, Hanko J. The significance of a crossed extensor hallucis response in neurologic disorders: a comparison with the Babinski sign. *Acta Neurol Scand.* 1976;53:241-250.
24. Kadanka Z, Bednarik J, Vohanka S, et al. Conservative treatment versus surgery in spondylotic cervical myelopathy: a prospective randomised study. *Eur Spine J.* 2000;9:538-544.
25. Kumar VG, Rea GL, Mervis LJ, McGregor JM. Cervical spondylotic myelopathy: functional and radiographic long-term outcome after laminectomy and posterior fusion. *Neurosurgery*. 1999;44:771-777; discussion 777-778.
26. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
27. MacFadyen DJ. Posterior column dysfunction in cervical spondylotic myelopathy. *Can J Neurol Sci.* 1984;11:365-370.
28. Matsuda Y, Miyazaki K, Tada K, et al. Increased MR signal intensity due to cervical myelopathy. Analysis of 29 surgical cases. *J Neurosurg.* 1991;74:887-892.
29. Matsumoto M, Toyama Y, Ishikawa M, Chiba K, Suzuki N, Fujimura Y. Increased signal intensity of the spinal cord on magnetic resonance images in cervical compressive myelopathy. Does it predict the outcome of conservative treatment? *Spine*. 2000;25:677-682.
30. McCormick WE, Steinmetz MP, Benzel EC. Cervical spondylotic myelopathy: make the difficult diagnosis, then refer for surgery. *Cleve Clin J Med.* 2003;70:899-904.
31. Mizuno J, Nakagawa H, Hashizume Y. Analysis of hypertrophy of the posterior longitudinal ligament of the cervical spine, on the basis of clinical and experimental studies. *Neurosurgery*. 2001;49:1091-1097; discussion 1097-1098.
32. Montgomery DM, Brower RS. Cervical spondylotic myelopathy. Clinical syndrome and natural history. *Orthop Clin North Am.* 1992;23:487-493.
33. Nurick S. The natural history and the results of surgical treatment of the spinal cord disorder associated with cervical spondylosis. *Brain*. 1972;95:101-108.
34. Ono K, et al. Cervical myelopathy secondary to multiple spondylotic protrusions. A clinicopathologic study. *Spine*. 1977;2:125.
35. Penning L, Wilmink JT, van Woerden HH, Knol E. CT myelographic findings in degenerative disorders of the cervical spine: clinical significance. *AJR Am J Roentgenol.* 1986;146:793-801.
36. Pui MH, Husen YA. Value of magnetic resonance myelography in the diagnosis of disc herniation and spinal stenosis. *Australas Radiol.* 2000;44:281-284.
37. Rao SC, Fehlings MG. The optimal radiologic method for assessing spinal canal compromise and cord compression in patients with cervical spinal cord injury. Part I: An evidence-based analysis of the published literature. *Spine*. 1999;24:598-604.
38. Rowland LP. Surgical treatment of cervical spondylotic myelopathy: time for a controlled trial. *Neurology*. 1992;42:5-13.
39. Seichi A, Takeshita K, Kawaguchi H, et al. Neurologic level diagnosis of cervical stenotic myelopathy. *Spine*. 2006;31:1338-1343. http://dx.doi.org/10.1097/01.brs.0000219475.21126.6b
40. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sam ple size requirements. *Phys Ther.* 2005;85:257-268.
41. Smith MS. Babinski's sign--abduction also counts. *JAMA*. 1979;242:1849-1850.
42. Spitznagel EL, Helzer JE. A proposed solution to the base rate problem in the kappa statistic. *Arch Gen Psychiatry.* 1985;42:725-728.
43. Stewart GW, Rey JM. A partial solution to the base rate problem of the kappa statistic. *Arch Gen Psychiatry.* 1988;45:504-505.
44. Sung RD, Wang JC. Correlation between a positive Hoffmann's reflex and cervical pathology in asymptomatic individuals. *Spine*. 2001;26:67-70.
45. Uebersax JS. Diversity of decision-making models and the measurement of interrater agreement. *Psychol Bull.* 1987;101:140-146.
46. Wong TM, Leung HB, Wong WC. Correlation between magnetic resonance imaging and radiographic measurement of cervical spine in cervical myelopathic patients. *J Orthop Surg (Hong Kong).* 2004;12:239-242.
47. Young WF. Cervical spondylotic myelopathy: a common cause of spinal cord dysfunction in older persons. *Am Fam Physician.* 2000;62:1064-1070, 1073.
48. Yousem DM, Atlas SW, Hackney DB. Cervical spine disk herniation: comparison of CT and 3DFT gradient echo MR scans. *J Comput Assist Tomogr.* 1992;16:345-351