Spine

SPINE Volume 38, Number 3, pp 245–252

DIAGNOSTICS

Reliability of Quantitative Magnetic Resonance Imaging Methods in the Assessment of Spinal Canal Stenosis and Cord Compression in Cervical Myelopathy

Alina Karpova, MSc,* Ranganathan Arun, DM, FRCSEd (Tr&Orth), PGDip (Orth Engin), MRCSEd,* Aileen M. Davis, PhD,‖ Abhaya V. Kulkarni, MD, PhD, FRCSC,¶ David J. Mikulis, MD, FRCPC,§ Chua Sooyong, MD,* Doron Rabin, MD,* Sorin Craciunas, MD, PhD,* Sean R. Smith, MD,* Mitchell A. Hansen, MD, PhD, FRACS,* Joshi George, MD, EDNS, FRCSGlasg,* and Michael G. Fehlings, MD, PhD, FRCSC*†‡

Study Design. Prospective, blinded reliability study of quantitative magnetic resonance imaging (MRI) measures in patients with cervical myelopathy.

Objective. To assess the intra- and interobserver reliability of commonly used quantitative MRI measures such as transverse area (TA) of spinal cord, compression ratio (CR), maximum canal compromise (MCC), and maximum spinal cord compression (MSCC).

Summary of Background Data. There is no consensus on an optimal quantitative MRI method(s) in assessing canal stenosis and cord compression.

Methods. Seven surgeons performed measurements on 17 digital MR images, on 4 separate occasions. The degree of stenosis was evaluated by measuring TA and CR on axial T2, MCC, and MSCC on midsagittal T1- and T2-weighted MRI sequences, respectively. Statistical analyses included repeated-measures analysis of variance and intraclass correlation coefficients (ICCs).

Results. The mean ± SD for intraobserver ICC was 0.88 ± 0.1 for MCC, 0.76 ± 0.08 for MSCC, 0.92 ± 0.07 for TA, and 0.82 ± 0.13 for CR. In addition, the interobserver ICC was 0.75 ± 0.04 for MCC, 0.79 ± 0.09 for MSCC, 0.80 ± 0.05 for CR, and 0.86 ± 0.03 for TA. Higher degree of canal compromise (MCC) was associated with lower modified version of Japanese Orthopaedic Association Scale score ($P = 0.05$). Also, a strong association was found between MSCC and lower modified version of Japanese Orthopaedic Association Scale score, greater number of steps, and longer walking time ($P < 0.05$).

Conclusion. All 4 measurement techniques demonstrated a good to moderately high degree of intra- and interobserver reliability. Highest reliability was noted in the assessment of T2-weighted sequences and axial MRI. Our results show that the measurements of MCC, MSCC, and CR are sufficiently reliable and correlate well with clinical severity of cervical myelopathy.

Key words: cervical canal stenosis, cord compression, inter- and intraobserver reliability, quantitative MRI. **Spine 2013;38:245–252**

Cervical myelopathy can broadly be defined as a symptomatic dysfunction of the cervical spinal cord caused by compressive etiologies.[1,2] It can occur in all adults because of cord compression resulting from 1 of several physiological factors including spondylosis/congenital stenosis, disc herniation, ossification of the posterior longitudinal ligament, hypertrophy of the ligamentum flavum, and degenerative subluxation. Previous studies have demonstrated inconsistencies in predicting surgical outcomes for patients with myelopathy.[3–6] Possible confounding factors include age, differing techniques for cervical decompression, and varying duration of symptoms. Lack of standardized imaging protocols to assess the severity of cord compression in cervical myelopathy may further contribute to difficulties in assessing severity and in predicting the outcome.

From the *Division of Neurosurgery, University of Toronto, Toronto, Canada; †Krembil Neuroscience Center, Toronto Western Hospital, Toronto, Canada; ‡Gerald and Tootsie Halbert Chair, Neural Repair and Regeneration, University Health Network, Toronto Western Hospital, Toronto, Canada; §Department of Radiology, Division of Neuroradiology, University of Toronto, Toronto Western Hospital, Toronto, Canada; ¶Division of Neurosurgery, Hospital for Sick Children, University of Toronto, Toronto, Canada; and ‖Health Care and Outcomes Research, Toronto Western Research Institute, University Health Network and Departments of Physical Therapy and Graduate Departments of Rehabilitation Science, Health Policy, Management and Evaluation and the Institute of Medical Science, University of Toronto, Toronto, Canada.

Acknowledgment date: September 14, 2011. First revision date: April 17, 2012. Second revision date: May 11, 2012. Acceptance date: June 26, 2012.

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

AOSpine North America and the Ontario Neurotrauma Foundation Scholarship funds were received in support of this work.

No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and reprint requests to Michael G. Fehlings, MD, PhD, FRCSC, Krembil Chair in Neural Repair and Regeneration, The Toronto Western Hospital, University Health Network, Room 4W-449 399, Bathurst Street, Toronto, Ontario M5T 2S8, Canada; E-mail: Michael.Fehlings@uhn.on.ca

DOI: 10.1097/BRS.0b013e3182672307

The current radiological modality of choice to assess the severity of cervical myelopathy is magnetic resonance imaging (MRI). MRI also provides important information about the etiology of canal stenosis and degree of cord compression, and helps to assess the adequacy of decompression. It tells us about the pathological changes within the cord, thereby helping the clinician give a prognosis regarding the clinical outcomes after surgery.[7–14] The disadvantages of MRI in the assessment of spinal cord stenosis are truncation,[15,16] chemical shift, cerebrospinal fluid flow, and motion artefacts.[17,18] These mechanisms can lead to overestimation of the degree of cervical stenosis, especially in severe cases. In certain instances, cervical spinal canal dimensions can change with different positions of the neck. Dynamic MRI is a newer modality of imaging that was specifically developed to address this limitation.[19]

For the past 4 decades, there have been several attempts to correlate clinical severity with degree of spinal cord compression on MRI.[10,11,13,14,20–24] Initial attempts were performed using qualitative methods.[7–9,11,12] More recently, quantitative MRI measurements have been introduced. The most commonly used parameters are transverse area (TA)[4] and compression ratio (CR) of the spinal cord.[3,4,14,25,26] There have been several attempts at validating these parameters with clinical severity and postoperative outcomes. However, for a measurement instrument to be considered valid, it needs to be reliable and reproducible. Reliability, hence, represents the minimal requirement for a valid clinical measure. To date, no publications have examined the reliability of quantitative MRI methods such as measurements of TA, CR, maximum spinal cord compression (MSCC) and maximum canal compromise (MCC) to assess the extent of cord compression or canal stenosis in cases of cervical myelopathy. The latter 2 techniques have been widely applied in the setting of traumatic cervical spinal cord injury, which has allowed for greater precision in evaluating cord compression and predicting outcomes after surgery.[27,28]

Given the gaps in published literature, we sought to quantitatively examine the intra- and interobserver reliability of 4 published methods of examining cord compression and canal stenosis on axial (TA and CR) and sagittal (MSCC and MCC) MRI sections. It was proposed to evaluate the reliability of quantitative MRI in providing information on the severity of cervical canal stenosis and spinal cord compression using software-based image analysis.

This study was approved by the University Health Network research ethics board.

MATERIALS AND METHODS

By using a systematic approach, it was proposed to evaluate the degree of cervical canal stenosis and spinal cord compression with software-based tools applied to digitized MRI scans magnified 200%, using TA, CR, MSCC, and MCC. Written instructions detailing the use of different software programs and measurements were provided to each of the independent observers.

Study Participants

The study was approved by the local institutional ethics committee. Four female and 13 male patients (age, 37–82 yr; mean, 54.5 yr) were recruited in the study. There were 5 single-level and 11 multiple-level cervical spinal cord compression cases due to a variety of common pathologies occurring in clinical practice (Table 1). These included 8 cases of spondylosis occurring in a congenitally narrow canal, 3 cases of disc herniation, 3 cases of ossified posterior longitudinal ligament, 2 cases of hypertrophy of the ligamentum flavum, and 1 case of degenerative subluxation. The levels of cord compression were 3 at C3–C4, 5 at C4–C5, 7 at C5–C6, and 2 at C6–C7 levels. To eliminate subject selection bias and ensure adequate spread of data, patients with a wide range of symptom severity were selected—6 mild, 5 moderate, and 6 severe cases, on the basis of scores on the modified version of the Japanese Orthopaedic Association Scale (mJOA).[29] One researcher, who did not take part in MRI measurements, was responsible for selecting the cases. The patients were randomly selected from a cohort of patients recruited in to the AOSpine North America–cervical spondylotic myelopathy trial. The patients were numbered 1 to 17 and their corresponding MR scans were assigned a designated code. To avoid memory recall and ensure blinding, the observers enrolled in the study were informed that MRI coding was manually changed and presented in a random fashion throughout the period of the study. The cervical spine MRI examinations were evaluated by 7 observers—all fellowship-trained in spine surgery. Of these, 5 were neurosurgeons and 2 were orthopedic surgeons with a mean self-reported MRI experience of 8.57 years (range, 7–11 yr).

MRI Analysis

The majority of MRI scans (15/17) were performed using a 1.5-T General Electric MRI system (2 patients were imaged at 3.0 T on a General Electric system, Signa, GE Medical Systems, Milwaukee, WI). The preoperative midsagittal T1-weighted, axial, and midsagittal T2-weighted MR scans of all patients were included in a CD-ROM with eFilm Lite (2003) (version 1.9, Merge Technologies Inc., Milwaukee, WI) and Mango 2.0 software (Multi-Image Analysis GUI) (ric.uthscsa.edu/mango). Figures 1 and 2 illustrate examples of the measurement techniques used in this study. Research guidelines do not mention a predefined timeframe for test-retest cycles and generally 1 week has been accepted for the purposes of memory recall. All observers were instructed on a 1-to-1 basis to ensure that they were familiar with the software programs and to ensure consistency in measurements. The MR images were evaluated by the observers using methodological guidelines detailed in the original studies from Fehlings *et al*[27] for MSCC and MCC, from Okada *et al*[4] for TA, and Chung and Chung[14] for CR. Observers were asked to consistently magnify the images by 200% to reduce the procedural variability of the measurements of cervical canal stenosis and spinal cord compression.

TABLE 1. Characteristics of the Patient Cohort With Cervical Myelopathy

Sex	Age (yr)	Different Physiological Factors of Degenerative Cervical Myelopathy	No. Stenotic Motion Segments	Severity of Cervical Myelopathy by mJOA Grades*	Levels of Maximal Compression
M	50	Spondylosis + CS	1	14	C6–C7
M	53	Spondylosis + CS	2	17	C5–C6
M	52	Spondylosis + CS	3	15	C3–C4
M	43	OPLL + HLF	2	15	C4–C5
M	65	DH	4	16	C4–C5
M	60	Spondylosis	2	15	C5–C6
M	38	Spondylosis + CS	1	13	C4–C5
M	68	Spondylosis + SL	8	13	C3–C4
M	61	Spondylosis + HLF	3	14	C5–C6
M	37	DH	1	14	C4–C5
F	54	SL	2	12	C3–C4
M	82	OPLL	2	11	C5–C6
M	52	OPLL + CS	4	8	C6–C7
F	58	HLF	3	10	C5–C6
M	59	SL + CS + HLF	4	10	C4–C5
F	55	Spondylosis	1	11	C5–C6
F	40	DH	1	10	C5–C6

**Mild (mJOA score ≥ 15), moderate (mJOA score 12–14), and severe (mJOA score <12).*

mJOA scale assesses upper extremity function (5 points), lower extremity function (7 points), sensory function (3 points), and urinary bladder function (3 points). Scores range from 0–18 with higher scores indicating better function (Benzel et al, 1991[29]).

CS indicates congenital stenosis; DH, disc herniation; HLF, hypertrophic ligament flavum; mJOA, modified version of Japanese Orthopaedic Association Scale; OPLL, ossification of the posterior longitudinal ligament; SL, subluxation.

Defined Radiological Parameters

On the basis of 3 dimensions from MRI (Figure 1), the MSCC using sagittal T2-weighted MRI (Figure 1A) and canal compromise using sagittal T1-weighted MRI (Figure 1B) were calculated using the following formulae.[27]

$$\text{MSCC}\ (\%) = (1 - [d_i / \{d_a + d_b\}/2]) \times 100\%$$

$$\text{MCC}\ (\%) = (1 - [d_i / \{d_a + d_b\}/2]) \times 100\%$$

Where d_i is the anteroposterior spinal canal diameter at the level of MSCC, d_a is the anteroposterior spinal canal diameter at the first normal vertebral segment above, and d_b is the anteroposterior spinal canal diameter at the first normal vertebral segment below the level of injury.

Note: Measurements of the normal canal anteroposterior diameter should be taken at the midvertebral body level.

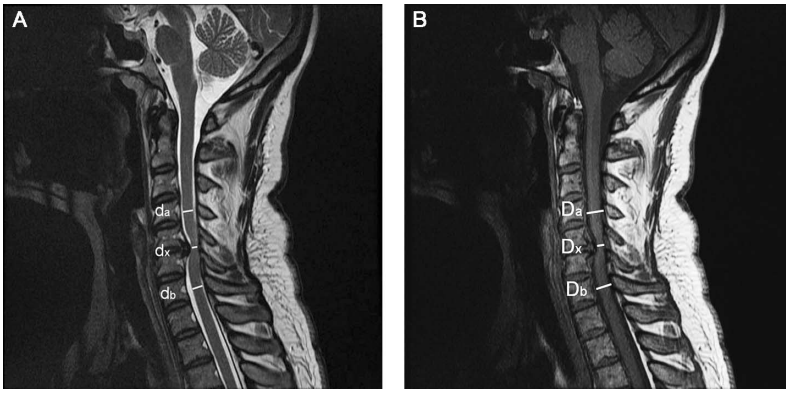

Figure 1. (**A**) Measurements for the maximum spinal cord compression using midsagittal T2-weighted magnetic resonance imaging (MRI). (**B**) Maximum canal compromise using midsagittal T1-weighted MRI.

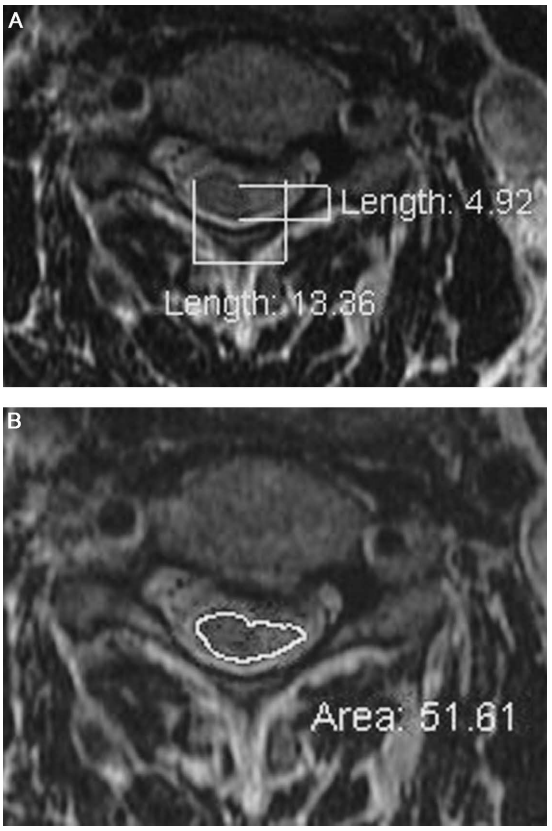

Figure 2. (**A**) Measurements for the compression ratio using axial T2-weighted MRI. (**B**) Method of ascertaining the transverse area of spinal cord using axial T2-weighted MRI.

TA was identified as the site of greatest compression using axial T2-weighted MRI (Figure 2B),[4] and CR was defined as the smallest sagittal diameter divided by the transverse diameter ratio of the spinal cord using axial T2-weighted MRI (Figure 2A).[14] The cross-sectional areas, and sagittal and transverse diameters of spinal cord were measured in the midline of vertebrae. In particular, the 3 measurements were limited to the spinal cord and not to the dural sac. The observers used a software-based algorithm to calculate the cross-sectional area of the spinal cord.

Statistical Analysis

Given that the primary objective of this study was to assess the reliability of 4 measures in the setting of myelopathy, we calculated a sample size of 17 patients based on 7 observers carrying out 4 separate observations on each subject to obtain results with a type 1 error of 5%, a minimal power of 80%, and a desired intraclass correlation coefficient (ICC) of 0.75 (expected level of ICC of 0.9).[30,31]

Intra- and Interobserver Reliability

All analyses were performed using constructed data sets in SAS version 9.3 (SAS Institute, Inc., Cary, NC) and Microsoft Excel 2003 software packages (Mississuaga, ON, Canada). Interobserver and intraobserver reliability were calculated from repeated-measures analysis of variance, according to Shrout-Fleiss models for random effects (model 2) using a 2-way random-effect model with absolute agreement (the raters were assumed to be randomly selected from the population).[32] Data were represented in terms of estimates of the true mean, SDs, standard error of the mean, and confidence intervals.[33] Correlation analysis was carried out using the Spearman correlation (nonparametric data) and the Pearson correlation (parametric data). $P < 0.05$ was considered significant.

The ratings of a single observer were treated as independent and therefore the following 3 factors were considered: observer, time (test-retest), and time × observer. A repeated-measures analysis of variance was used to calculate ICC values individually for each observer and retest.

RESULTS

Descriptive Statistics

All 17 patients underwent cervical axial and sagittal T1- and T2-weighted MRI on admission. For each MRI-based quantitative measure, 7 observers measured all 17 patients with cervical myelopathy on 4 separate occasions. The TA of the spinal cord varied from 59.6 to 80 mm^2 for T1-weighted MRI measurements. The CR and MCC ranged from 0.32% to 0.36% and from 82% to 85.7%, respectively. The MSCC was between 81% and 84.1% in our series (Table 2). The differences between the 7 observers for all 4 radiological parameters (MCC, MSCC, TA, and CR) were not statistically significant based on the repeated-measures analysis of variance (Table 2).

Assessment of Intraobserver Reliability

The intraobserver reliability ICCs for the 4 methods are shown in Table 3. The mean intraobserver ICC was 0.76 ± 0.08 (mean ± SD) for the T2-weighted MSCC, 0.82 ± 0.13 for the T2-weighted CR, 0.88 ± 0.1 for the T1-weighted MRI-MCC, and 0.92 ± 0.07 for the T2–weighted TA of spinal cord. All the 4 measurement methods had acceptable intraobserver reliability. (ICC values higher than 0.75 indicate acceptable reliability.[33])

Assessment of Interobserver Reliability

The interobserver reliability ICC values for the 4 methods of each session are shown in Table 4. The mean interobserver ICC was 0.75 ± 0.04 for T1-weighted MCC, 0.79 ± 0.09 for T2-weighted MSCC, 0.80 ± 0.05 for T2–weighted CR, and 0.86 ± 0.03 for T2–weighted TA. Again, all the 4 measurement methods had an acceptable interobserver reliability. (ICC values higher than 0.75 indicate acceptable reliability.[33])

Comparison of Axial and Midsagittal Section-Based Measurements

The intra- and interobserver ICCs were 0.92 ± 0.07 and 0.86 ± 0.03 for the T2-weighted TA, and 0.82 ± 0.13 and 0.86 ± 0.03 for the T2-weighted CR, respectively. In contrast, the intra- and interobserver ICCs were 0.88 ± 0.1 and 0.75 ± 0.04 for the T1-weighted MRI-MCC, 0.76 ± 0.08 and 0.79 ± 0.09 for the T2-weighted MSCC, respectively. TA

TABLE 2. Quantitative Assessment of the Magnetic Resonance Imaging Parameters of Canal Compromise and Cord Compression

Measure (Mean ± SD, Min, Max)	Transverse Area	Compression Ratio	Maximum Canal Compromise	Maximum Spinal Cord Compression
Rater 1	74.8 ± 19.28	0.32 ± 0.08	82.0 ± 2.76	82.8 ± 3.45
	27.0, 119.3	0.18, 0.63	75.0, 90.0	76.9, 91.7
Rater 2	80.0 ± 23.2	0.36 ± 0.09	82.4 ± 3.82	82.4 ± 2.59
	34.7, 126.0	0.25, 0.67	75.0, 90.6	77.3, 91.7
Rater 3	59.6 ± 20.18	0.32 ± 0.09	85.7 ± 3.77	84.1 ± 2.88
	30.0, 107.4	0.18, 0.67	77.5, 95.5	75.0, 90.9
Rater 4	71.4 ± 18.99	0.35 ± 0.08	82.6 ± 3.41	82.1 ± 3.09
	7.87, 104.7	0.13, 0.53	73.7, 91.3	75.0, 90.9
Rater 5	69.3 ± 15.6	0.33 ± 0.09	83.0 ± 3.75	83.4 ± 3.33
	40.4, 110.7	0.19, 0.54	76.3, 93.8	76.9, 89.3
Rater 6	71.2 ± 17.23	0.36 ± 0.07	83.7 ± 4.10	81.0 ± 3.82
	39.4, 111.8	0.22, 0.53	76.5, 96.0	72.7, 89.3
Rater 7	70.9 ± 17.24	0.35 ± 0.08	83.1 ± 3.37	81.6 ± 3.42
	43.1, 107.2	0.21, 0.53	76.5, 92.5	70.8, 89.3

and CR as axial section–based parameters showed higher intra- and interobserver reliability than MCC and MSCC using midsagittal sections.

Comparison of T1- and T2-Weighted MRI Based Parameters

The intra- and interobserver ICCs were 0.88 ± 0.1 and 0.75 ± 0.04 for the T1–weighted MRI-MCC, respectively. In contrast, the intra- and interobserver ICCs were 0.76 ± 0.08 and 0.79 ± 0.09 for T2-weighted MSCC, respectively. T1-weighted sequences showed higher intraobserver but lower interobserver reliability than T2-weighted.

Comparison of TA and CR

The intra- and interobserver ICCs were 0.92 ± 0.07 and 0.86 ± 0.03 for T2-weighted TA, respectively. In contrast, the intra- and interobserver ICCs were 0.82 ± 0.13 and 0.86 ± 0.03 for the T2–weighted CR, respectively. TA showed higher intra- and interobserver reliability than CR.

Correlation of MRI Parameters With Clinical Severity

Each MRI parameter significantly correlated with at least 1 of the 4 clinical severity scales of cervical myelopathy (*i.e.*, mJOA, Nurick, and walking test including number of steps and time). TA tends to be lower with greater number of steps ($r = –0.42$; $P = 0.09$). Higher CR was associated with greater number of steps and longer walking time ($r = 0.48$; $P = 0.05$ and $r = 0.52$; $P = 0.03$). More substantial MCC was associated with lower mJOA score ($r = –0.48$; $P = 0.05$). Greater MSCC was associated with lower mJOA score ($r = –0.68$; $P = 0.005$), greater number of steps ($r = –0.62$; $P = 0.01$), and longer walking time ($r = –0.55$; $P = 0.02$). (Table 5).

DISCUSSION

Previous studies have not assessed the variability between observers' findings in evaluating the severity of cervical myelopathy using quantitative MRI-based approach.[3–6,13,14] Our study systematically analyzed the intra- and interobserver reliability of 4 commonly used quantitative methods to assess the severity of stenosis in cervical myelopathy patients. The present study suggests: (1) intraobserver reliability was high, whereas the interobserver ICCs were at a moderate level; (2) axial sections based on T2-weighted MRI parameters (CR and TA) were more reliable in assessing the degree of cervical spinal cord compression; (3) assessment of spinal canal stenosis was more reliable using T2- rather than T1-weighted sequences; and (4) TA was more reliable and versatile (in both symmetric and asymmetric cord compression) than CR in assessing spinal cord compression.

The current radiological modality of choice to assess the severity of cervical myelopathy is MRI because this modality provides important information about pathological changes within the cord, thereby helping the clinician to give a prognosis regarding improvement after surgery.[7–12] Several studies have developed MRI-based measurement instruments to quantify canal stenosis and cord compression.[3–6,13,14] Development of such a measurement tool involves multiple steps with a minimum requirement of appraisals for validity and reliability.

TABLE 3. Intrarater Reliability: ICC Values Using the Shrout-Fleiss Model for Random Effects Regarding Spinal Cord and Canal Deformities Evaluated by TA, CR, MCC, and MSCC

Measure	TA	CR	MCC	MSCC
Rater 1	0.82, 13.3 (61.4–89.9)	0.58, 0.07 (0.84–1.0)	0.82, 1.96 (80.8–84.9)	0.76, 2.56 (80.7–86.5)
Rater 2	0.99, 3.9 (76.5–85.5)	0.95, 0.04 (1.27–1.35)	0.97, 1.34 (82.1–84.7)	0.89, 1.53 (81.8–84.9)
Rater 3	0.98, 6.3 (55.1–66.1)	0.82, 0.06 (1.07–1.21)	0.80, 2.69 (83.5–89.5)	0.85, 1.86 (82.9–87.0)
Rater 4	0.88, 11.4 (60.4–84.2)	0.75, 0.06 (1.03–1.17)	0.72, 2.63 (80.3–86.3)	0.76, 2.36 (80.3–85.4)
Rater 5	0.97, 5.1 (65.1–75.5)	0.90, 0.05 (1.17–1.27)	0.92, 1.88 (81.9–85.9)	0.69, 2.50 (81.1–86.9)
Rater 6	0.85, 10.9 (60.0–84.2)	0.82, 0.05 (1.11–1.23)	0.96, 1.67 (83.1–86.3)	0.65, 2.74 (76.7–84.7)
Rater 7	0.96, 6.0 (65.2–78.6)	0.93, 0.04 (1.24–1.32)	0.96, 1.27 (82.7–85.3)	0.75, 2.61 (75.6–85.5)

(ICC, SEM, 95% CI**); *SEM = square root of MSE; **Mean 95% CI: Mean + ICC ± 1.96 *SD *squared root of [ICC (1 − ICC)], where SD = square root of (SST/n − 1). (Weir et al, 2005[34].)*

CI indicates confidence interval; CR, compression ratio; ICC, intraclass correlation coefficient; MCC, maximum canal compromise; MSCC, maximum spinal cord compression; SEM, standard error of mean; SST, Total Sum of Squares; TA, transverse area.

Previous validity studies correlating the relationship between quantitative MRI measures and clinical outcomes showed either a good or poor association.[4,5] One of the potential reasons for this could be the lack of reliability studies assessing quantitative MRI measurement techniques. This study was an attempt to assess the intra- and interobserver reliability of 4 most frequently used quantitative MRI measures in the assessment of spinal canal stenosis and cord compression.

Optimal visualization of the entire spinal cord is typically obtained with sagittal slices in cervical myelopathy. Important additional localizing information may be obtained from axial sections such as asymmetry of spinal cord compression. In addition, spinal cord compression in cervical myelopathy tends to have multilevel involvement in most cases. The disagreement on the most compressed site of the spinal cord in multilevel cases on the midsagittal section can be resolved by measuring the cross-sectional area of the spinal cord on axial section. Therefore, the lack of reliability among observers noticed on the mid-sagittal sections can be explained by the lack of consistency in agreeing on the site of maximal spinal canal stenosis.

The spinal cord has a better visual contrast on T2-weighted MRI against the background of bright cerebrospinal fluid; this facilitates more reliable measurements. This may explain the better intra- and interobserver reliability for MSCC measurements on mid-sagittal T2-weighted when compared with MCC measurements that were performed on T1-weighted MR images. In addition to higher reliability, measuring the degree of spinal cord compression, quantified by MSCC, correlates with clinical severity to a greater degree than the degree of canal compromise. In the past, a number of studies frequently used TA of the spinal cord as a severity measure of cervical myelopathy based solely on axial T1-weighted sequences.[4–6] However, the findings from this study suggest that axial T2-weighted MRI should be preferred to assess the severity of spinal cord compression in cervical myelopathy.

It is well known that clinical scoring systems such as Nurick and mJOA scales for cervical myelopathy have low sensitivities ($r = 0.21$, $r = 0.42$, respectively).[37] In our study, CR measurement had a lower reliability index than TA measurement. Therefore, it was not surprising that previous studies that compared clinical scores with MRI using these parameters showed a poor correlation.[3,4,14] CR may misrepresent the degree of spinal cord deformation when both sagittal and transverse diameters are compromised equally. CR is also limited in its application when there is asymmetrical compression. Most studies have reported consistent correlation between TA

TABLE 4. Interrater Reliability: ICC Values Using the Shrout-Fleiss Model for Random Effects Regarding Spinal Cord and Canal Deformities Evaluated by TA, CR, MCC, and MSCC

Measure	TA	CR	MCC	MSCC
Time 1	0.82, 14.3 (58.0–86.4)	0.82, 0.06 (1.16–1.22)	0.74, 3.0 (80.6–87.0)	0.82, 2.6 (81.2–86.2)
Time 2	0.86, 15.2 (57.2–86.7)	0.75, 0.08 (0.95–1.11)	0.81, 2.9 (81.2–86.9)	0.66, 2.6 (82.3–86.0)
Time 3	0.87, 13.4 (60.1–86.7)	0.75, 0.07 (1.03–1.17)	0.72, 3.3 (80.4–87.2)	0.84, 2.4 (80.7–85.3)
Time 4	0.89, 12.2 (59.0–81.2)	0.86, 0.06 (1.14–1.26)	0.73, 3.1 (81.0–87.6)	0.85, 2.3 (81.3–86.0)

(ICC, SEM, 95% CI**); *SEM = square root of MSE; **Mean 95% CI: Mean + ICC ± 1.96 *SD *squared root of [ICC (1 − ICC)], where SD = square root of (SST/n − 1). (Weir et al, 2005[34].)*

CI indicates confidence interval; CR, compression ratio; ICC, intraclass correlation coefficient; MCC, maximum canal compromise; MSCC, maximum spinal cord compression; SEM, standard error of mean; SST, Total Sum of Squares; TA, transverse area.

TABLE 5. Relationship Between Spinal Cord Compression, Canal Compromise, and Clinical Severity of Cervical Myelopathy

Measure	Transverse Area	Compression Ratio	Maximum Canal Compromise	Maximum Spinal Cord Compression
mJOA	r = –0.07	r = 0.45	r = –0.48	r = –0.65
	P = 0.80	*P* = 0.07	*P* = 0.05	*P* = 0.005
Nurick*	r = –0.39	r = 0.34	r = 0.23	r = –0.24
	P = 0.12	*P* = 0.18	*P* = 0.37	*P* = 0.36
Walking time (s)†	r = –0.41	r = 0.48	r = –0.05	r = –0.62
	P = 0.10	*P* = 0.05	*P* = 0.85	*P* = 0.01
Number of steps†‡	r = –0.42	r = 0.52	r = –0.10	r = –0.55
	P = 0.09	*P* = 0.03	*P* = 0.72	*P* = 0.02

Primary outcome measures included results of disease-specific instruments (Nurick grade[35], mJOA,[29] and time/cadence of the 30 m walking test (30MWT).[36]

**Nurick grade: score ranges from 0 (least disability) to 4 (most disability). Grades 0 and 1, no gait abnormalities; 2–4, increasing mobility issues.*

†Walking test: time in seconds and number of steps (cadence) in a 30-m walk. Improvements of several seconds and steps would translate to significant improvements in the patient's ability and confidence to walk independently.

‡Number of steps (cadence) in a 30-m walk.

mJOA indicates modified Japanese Orthopaedic Association scale.

of the spinal cord and the clinical severity of cervical myelopathy.[4–6] Data from our study support the fact that TA is also a "reliable" quantitative MRI parameter. However, our data do not support TA as a valid measurement instrument of cervical myelopathy possibly due to small sample size and low sensitivity of currently used clinical scales. Subsequently, for further validation of this parameter, TA measurements were completed on 65 patients with cervical myelopathy.[38] The results showed a significant correlation between TA of the spinal cord and mJOA score (r = 0.02; *P* = 0.02).

Previous studies from our center have shown the intra- and interobserver ICCs of 0.68 and 0.70, and 0.55 and 0.61 for MSCC and MCC measures, respectively, in acute traumatic cervical spinal cord injuries.[28] On the basis of the present study, MSCC and MCC appear to be more reliable in the setting of patients with nontraumatic spinal cord injury. The higher ICCs measured in nontraumatic spinal cord injury might result from a lesser disruption of normal anatomy. These findings also suggest that parameters such as MSCC used to measure severity of cord compression correlate better with clinical severity than measurements of spinal canal dimensions (MCC).

CONCLUSION

The intraobserver reliability of the frequently used quantitative MRI measures such as CR, MCC, MSCC, and TA, in cases of cervical myelopathy is high. The interobserver reliability for the same measures was moderately reliable. This study supports the use of T2-weighted axial MRI sequence based parameters such as TA and CR in the clinical and research settings. Parameters used to measure severity of cord compression such as MSCC correlate better with clinical severity than measurements of spinal canal dimensions (MCC).

➢ Key Points

- ❑ All 4 measurement techniques demonstrated a good to moderately high degree of intra- and interobserver reliability.
- ❑ Highest reliability was noted in the assessment of T2-weighted sequences and axial MRI.
- ❑ MSCC measurement is a more reliable measure and correlates with clinical severity to a greater degree than MCC.
- ❑ Also, TA measurement of the spinal cord was found to be the most reliable parameter; however, a weak correlation with clinical severity in our study suggests future research is necessary using a larger sample size.

Acknowledgments

The authors are grateful for technical support of Yuriy Petrenko, MD, and Amy Lem.

References

1. Montgomery DM, Brower RS. Cervical spondylotic myelopathy. Clinical syndrome and natural history. [Review] [54 refs]. *Orthop Clin North Am* 1992;23:487–93.
2. Cadotte DW, Karpova A, Fehlings MG. Cervical spondylotic myelopathy: surgical outcomes in the elderly. *Int J Clin Rheumatol* 2010;5:327–37.
3. Chen CJ, Lyu RK, Lee ST, et al. Intramedullary high signal intensity on T2-weighted MR images in cervical spondylotic myelopathy: prediction of prognosis with type of intensity. *Radiol* 2001;221:789–94.
4. Okada Y, Ikata T, Yamada H, et al. Magnetic resonance imaging study on the results of surgery for cervical compression myelopathy. *Spine* 1993;18:2024–9.

5. Morio Y, Teshima R, Nagashima H, et al. Correlation between operative outcomes of cervical compression myelopathy and MRI of the spinal cord. *Spine* 2001;26:1238–45.
6. Fukushima T, Ikata T, Taoka Y, et al. Magnetic resonance imaging study on spinal cord plasticity in patients with cervical compression myelopathy. *Spine* 1991;16(suppl):S534–8.
7. Nagata K, Kiyonaga K, Ohashi MS, et al. Clinical value of magnetic resonance imaging for cervical myelopathy. *Spine* 1990;15:1089–96.
8. Nagata K, Ohashi T, Abe J, et al. Cervical myelopathy in elderly patients: clinical results and MRI findings before and after decompression surgery. *Spinal Cord* 1996;34:220–6.
9. Uchida K, Nakajima H, Sato R, et al., Multivariate analysis of the neurological outcome of surgery for cervical compressive myelopathy. *J Orthop Sci* 2005;10:564–73.
10. Houten JK, Cooper PR. Laminectomy and posterior cervical plating for multilevel cervical spondylotic myelopathy and ossification of the posterior longitudinal ligament: effects on cervical alignment, spinal cord compression, and neurological outcome. *Neurosurg* 2003;52:1081–7; discussion 1087–8.
11. Singh A, Crockard HA, Platts A, et al. Clinical and radiological correlates of severity and surgery-related outcome in cervical spondylosis. *J Neurosurg* 2001;94(suppl):189–98.
12. Matsuyama Y, Kawakami N, Mimatsu K. Spinal cord expansion after decompression in cervical myelopathy. Investigation by computed tomography, myelography, and ultrasonography. *Spine* 1995;20:1657–63.
13. Yone K, Sakou T, Yanase M, et al. Preoperative and postoperative magnetic resonance image evaluations of the spinal cord in cervical myelopathy. *Spine* 1992;17(suppl):S388–92.
14. Chung S, Chung KH. Factors affecting the surgical results of expansive laminoplasty for cervical spondylotic myelopathy. *Int Orthop* 2002;26:334–8.
15. Czervionke LF, Czervionke JM, Daniels DL, et al. Characteristic features of MR truncation artifacts. *AJNR Am J Neuroradiol* 1988;9:815–24.
16. Breger RK, Kass CL, Yu EG, et al. Truncation artefact in MR images of the intervertebral disk. *AJNR Am J Neuroradiol* 1988;9:825–8.
17. Sherman JL, Citrin CM, Gangarosa RE, et al. The MR appearance of CSF pulsation in the spinal canal. *AJNR Am J Neuroradiol* 1986;7:879–84.
18. Rubin JB, Enzmann DR. Imaging of spinal CFS pulsation by 2DFT MR: significance during clinical imaging. *AJNR Am J Neuroradiol* 1987;8:297–306.
19. Miura J, Doita M, Miyata K, et al. Dynamic evaluation of the spinal cord in patients with cervical spondylotic myelopathy using a kinematic magnetic resonance imaging technique. *J Spinal Disord Tech* 2009;22:8–13.
20. Matsuda Y, Shibata T, Oki S, et al. Outcomes of surgical treatment for cervical myelopathy in patients more than 75 years of age. *Spine* 1999;24:529–34.
21. Yukawa Y, Kato F, Yoshihara H, et al. MR T2 image classification in cervical compression myelopathy: predictor of surgical outcomes. *Spine* 2007;32:1675–8; discussion 1679.
22. Yamazaki T, Yanaka K, Sato H, et al. Cervical spondylotic myelopathy: surgical results and factors affecting outcome with special reference to age differences. *Neurosurg*. 2003;52:122–6; discussion 126.
23. Wada E, Ohmura M, Yonenobu K. Intramedullary changes of the spinal cord in cervical spondylotic myelopathy. *Spine* 1995;20:2226–32.
24. Fernández de Rota JJ, Meschian S, Fernández de Rota A, et al. Cervical spondylotic myelopathy due to chronic compression: the role of signal intensity changes in magnetic resonance images. *J Neurosurg Spine* 2007;6:17–22.
25. Shin JJ, Jin BH, Kim KS, et al. Intramedullary high signal intensity and neurological status as prognostic factors in cervical spondylotic myelopathy. *Acta Neurochir (Wien)* 2010;10: 1687–94.
26. Ahn JS, Lee JK, Kim BK. Prognostic factors that affect the surgical outcome of the laminoplasty in cervical spondylotic myelopathy. *Clin Orthop Surg* 2010;2:98–104.
27. Fehlings MG, Furlan JC, Massicotte EM. Interobserver and intraobserver reliability of maximum canal compromise and spinal cord compression for evaluation of acute traumatic cervical spinal cord injury. *Spine* 2006;31:1719–25.
28. Furlan JC, Fehlings MG, Massicotte EM, et al. A quantitative and reproducible method to assess cord compression and canal stenosis after cervical spine trauma *Spine* 2007;32:2083–91.
29. Benzel EC, Lancon J, Kesterson L, et al. Cervical laminectomy and dentate ligament section for cervical spondylotic myelopathy. *J Spinal Disorders*. 1991;4:286–95
30. Kraemer HC, Komera AF. Statistical alternatives in assessing reliability, consistency and individual differences for quantitative measures: application to behavioral measures of neonates. *Psychol Bull* 1976;83:914–21.
31. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Medi* 1998;17:101–10.
32. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychological Bull* 1979;86:420–8.
33. Burdock EI, Fleiss JL, Hardesty AS. A new view of interobserver agreement. *Perspect Psychol* 1963;16:373–84.
34. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Conditioning Res* 2005;19:231–40.
35. Nurick S. The pathogenesis of the spinal cord disorder associated with cervical spondylosis. *Brain* 1972;95:87–100.
36. Singh A, Crockard HA. Quantitative assessment of cervical spondylotic myelopathy by a simple walking test. *Lancet* 1999;354: 370–3.
37. Singh A, Crockard HA. Comparison of seven different scales used to quantify severity of cervical spondylotic myelopathy and post-operative improvement. *J Outcome Measures* 2001;5: 798–818.
38. Karpova A, Ranganathan A, Davis AM, et al. Predictors of outcomes in patients undergoing surgery for cervical spondylotic myelopathy *Spine In press*.