

Clinical and Radiological Characterization of Cervical Spinal Stenosis Using Muhle Grading on Flexion-Extension Dynamic MRI: Correlation With mJOA in an Institutional Cohort

Global Spine Journal
2026, Vol. 0(0) 1-11
© The Author(s) 2026
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/21925682261464620
journals.sagepub.com/home/gsj



José Ramón Rodríguez-Fonseca, MD¹, Daniel Alberto Reyes-Navarro, MD¹,
Armando Saúl Ruiz-Treviño, MD¹, Alejandro Miranda González, MD¹,
José Oscar Chávez-Bejar, MD¹, Alexis Agustín Altamirano Carrizoza, MD¹,
Christian Gabriel de Jesús Ordaz-Pérez, MD¹, and María Maldonado-Vega, PhD²

Abstract

Study Design: Observational retrospective cross-sectional study. **Objective:** To characterize patients with cervical spinal stenosis evaluated with flexion–extension dynamic MRI and assess the association between dynamic Muhle severity and modified Japanese Orthopaedic Association (mJOA) category. **Methods:** Adults undergoing sagittal T2 dynamic cervical MRI (neutral, extension, flexion) at a tertiary referral spine service (Mexico, 2024–2025) were included. Dynamic Muhle grade was the maximum across positions. We recorded spinal cord compression on neutral versus dynamic MRI, conversion to compression, increased compression, and an additional compressed level on dynamic sequences. T2 hyperintensity on neutral MRI was recorded as an additional imaging variable. Interobserver agreement was assessed across all readers, and intraobserver agreement in the principal readers. Exploratory diagnostic performance of neutral MRI relative to dynamic MRI was calculated. Ordinal logistic regression modeled mJOA category as the dependent outcome and dynamic Muhle grade as the principal predictor, adjusted for age and sex. **Results:** Fifty patients were analyzed (mean age 58.5±9.5 years; 62% female). Dynamic Muhle grades were 1: 10%, 2: 28%, and 3: 62%. Spinal cord compression was more frequent on dynamic than neutral MRI (90% vs 72%); conversion occurred in 18% and increased compression in 42%. Neutral MRI showed 80.0% sensitivity and 100% specificity relative to dynamic MRI for detecting compression. T2 hyperintensity on neutral MRI was present in 60% and correlated with higher dynamic Muhle grade ($\rho=0.536$, $p<0.001$). Interobserver agreement across all readers was moderate at the single-rater level and high at the averaged-rater level, whereas intraobserver agreement in the principal readers was excellent. In adjusted ordinal logistic regression, higher dynamic Muhle grade remained independently associated with worse mJOA category (OR 24.10; 95% CI 4.85–119.76; $p<0.001$). **Conclusions:** Flexion–extension dynamic MRI frequently reveals clinically relevant, posture-dependent cord compression underestimated on neutral imaging. Higher dynamic Muhle grades were strongly associated with worse clinical severity, supporting selective dynamic MRI when neutral findings are discordant with the clinical picture. Structured calibration in a teaching-hospital setting may improve reproducibility of Muhle grading.

¹Department of Neurosciences, Spine Service, Hospital Regional de Alta Especialidad del Bajío/ Universidad de Guanajuato, León, Mexico

²Research Department, Hospital Regional de Alta Especialidad del Bajío, León, Guanajuato, Mexico

Corresponding author:

José Ramón Rodríguez-Fonseca, MD, Department of Neurosciences, Spine Service, Hospital Regional de Alta Especialidad del Bajío / Universidad de Guanajuato, Blvd. Milenio #130, Col. San Carlos la Roncha, León, Guanajuato, CP 37544, Mexico.

Email: jose.ramon.rdz@gmail.com



Creative Commons Non Commercial No Derivs CC BY-NC-ND: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits non-commercial use, reproduction and distribution of the work as published without adaptation or alteration, without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Keywords

spinal cord compression, magnetic resonance imaging, dynamic magnetic resonance, spondylosis, myelopathy

Received: 5 March 2026; Revised: 2 June 2026; Accepted: 16 June 2026

Introduction

Cervical spinal stenosis is a common degenerative condition characterized by progressive narrowing of the cervical spinal canal, typically driven by disc degeneration, osteophyte formation, facet arthropathy, and ligamentous thickening (including ligamentum flavum hypertrophy)¹⁻³. This narrowing may be identified incidentally or in patients with neck pain and/or radicular symptoms, but it also represents the structural substrate for spinal cord compression and the development of degenerative cervical myelopathy (DCM).¹⁻³ Importantly, cervical canal narrowing and even radiographic cord compression can be present in patients with minimal symptoms, which contributes to under-recognition and delayed diagnosis in real-world practice.^{1,2}

DCM is the most common cause of adult spinal cord dysfunction and a major source of disability and health-care utilization.¹⁻³ The pathophysiology is not purely static: motion-dependent changes in canal geometry during flexion and extension can modify the subarachnoid space and functional reserve around the cord, plausibly contributing to dynamic compromise and symptom fluctuation.^{4,5} In this context, one of the main challenges is accurate severity stratification, particularly when neurologic findings suggest myelopathy but neutral MRI appears insufficient to explain the patient's clinical severity (clinico-radiological discordance).^{1-3,6}

Clinical severity in suspected or established DCM is commonly assessed using the modified Japanese Orthopaedic Association (mJOA) score, with validated approaches to interpret and categorize functional impairment.^{7,8} These categories are also used in contemporary clinical practice recommendations to guide management decisions.⁹ While neurologic examination and functional scoring define clinical impact, imaging remains essential for confirming central cord compression, defining involved levels, and supporting management planning. Standard cervical MRI is typically acquired in a neutral supine position; however, neutral imaging may underestimate clinically relevant compression in patients whose symptoms are posture-dependent or whose neurologic findings appear disproportionate to neutral MRI findings.^{1,2}

Flexion-extension dynamic MRI has therefore been proposed as a complementary tool in selected patients with cervical spinal stenosis to capture positional changes in cord compression.^{1,2} Prior work also suggests that dynamic imaging may reveal occult or missed dynamic canal stenosis compared with neutral MRI alone.³⁻⁵ Nevertheless, broader adoption has been limited by variability in indications and by the need for a reproducible framework to grade dynamic stenosis severity and support clinico-radiological correlation.

The Muhle classification provides a practical approach to grade central canal compromise and can be applied to dynamic imaging to capture clinically meaningful changes across positions⁶. Beyond grading severity at a given level, dynamic MRI can reveal patterns with immediate clinical relevance, including conversion from no compression on neutral MRI to compression on dynamic sequences, increased compression relative to neutral imaging, and identification of an additional compressed level detected only on dynamic sequences.^{1-3,5}

Therefore, the objectives of this study were to describe the clinical and radiological characteristics of adults with cervical spinal stenosis evaluated with sagittal T2 flexion-extension dynamic MRI in a tertiary referral spine service, and to determine the association between dynamic Muhle severity and clinical severity measured by mJOA categories. We additionally characterized dynamic imaging patterns—including increased compression, conversion to compression, and an additional compressed level detected only on dynamic sequences—and explored their relationships with neurologic examination findings and radiographic instability criteria. We also quantified occult compression missed on neutral MRI and evaluated the reproducibility of Muhle grading in a teaching-hospital setting. We hypothesized that higher dynamic Muhle grades would correlate with worse mJOA categories and more frequent long-tract signs.

Materials and Methods*Study Design and Setting*

We conducted an observational, retrospective, cross-sectional study at a tertiary referral teaching-hospital spine service in Mexico. Adult patients evaluated for cervical spinal stenosis who had a cervical flexion-extension dynamic MRI performed between 2024 and 2025 were eligible for inclusion. This study is reported in accordance with the STROBE guideline for observational studies.

Ethics

The study was approved by the institutional Ethics Committee (CEI-080-2024) and the institutional Research Committee (CI/074/2024). Given the retrospective design and use of de-identified data, informed consent was waived in accordance with local regulations.

Participants

Patients were included if they were 18 years of age or older, had a clinical diagnosis of cervical spinal stenosis/suspected degenerative cervical myelopathy based on symptoms and neurologic examination, and underwent cervical dynamic MRI with sagittal T2-weighted sequences acquired in neutral, extension, and flexion positions within the study period. Patients were excluded if they had prior cervical spine surgery, non-degenerative causes of myelopathy (e.g., tumor, infection, inflammatory disease), acute trauma at presentation, or incomplete imaging preventing Muhle grading in any of the required positions.

Study size: No a priori sample-size calculation was performed; the study included all consecutive eligible patients meeting inclusion criteria during the prespecified period.

Bias: To reduce selection bias, we included consecutive eligible patients with cervical spinal stenosis who had flexion–extension dynamic MRI available during the study period. Variable definitions were prespecified, and data were abstracted using a standardized template to minimize information bias.

Clinical Assessment

Clinical assessment at the time of evaluation included functional severity using the modified Japanese Orthopaedic Association (mJOA) score and a standardized neurologic examination. The examination captured long-tract signs (Hoffmann and/or Trömmner), deep tendon reflexes, and sensory findings in the C5–T1 distribution. Cervicalgia was recorded as present/absent based on patient report at the clinical encounter. Grip strength was assessed by hand dynamometry and dichotomized as decreased (coded as 1) if below age- and sex-adjusted normative values based on published population reference tables⁹; otherwise it was coded as not decreased (0). The mJOA variable was analyzed as an ordinal clinical severity category (no myelopathy/mild/moderate/severe).^{10,11}.

Imaging Protocol: Flexion-Extension Dynamic MRI

All patients underwent conventional cervical MRI plus flexion-extension dynamic sagittal T2-weighted sequences acquired in neutral, extension (chin up), and flexion (chin down) positions using an in-house custom positioning device. Dynamic sequences were obtained using identical acquisition parameters across positions (TR/TE 3500/97 ms; slice thickness 3 mm; base resolution 320; phase resolution 80%; field of view 220 mm), with an approximate acquisition time of ~1:30 minutes per sequence. Positioning was performed to patient tolerance to capture physiologic range of motion while maintaining safety and image quality; sequences were interrupted if pain or new neurologic symptoms occurred. The institutional protocol is illustrated in Figure 1.^{1,2,12}.

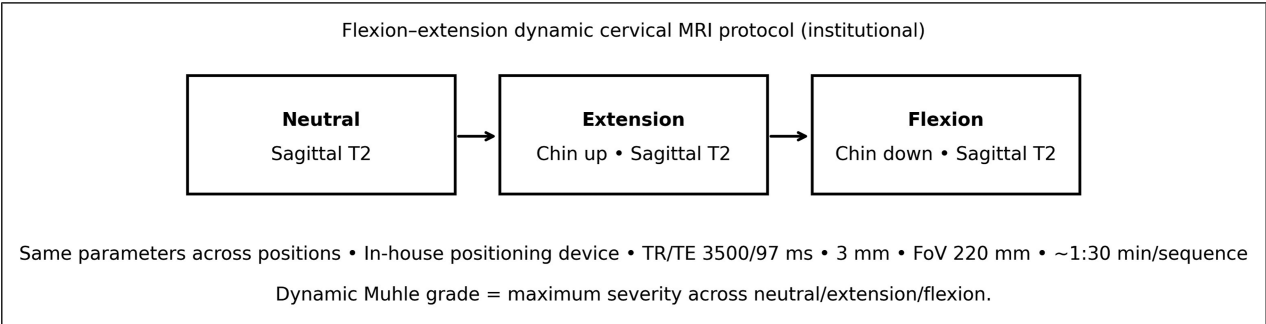


Figure 1. Dynamic cervical MRI protocol. Sagittal T2-weighted imaging obtained in neutral, extension, and flexion positions illustrating the dynamic acquisition used for Muhle grading and identification of positional compression patterns

Radiologic Assessment and Definitions

Cervical canal stenosis severity was graded using the Muhle classification on dynamic MRI. The dynamic Muhle grade was defined as the maximum severity observed across the three sagittal T2 positions (neutral, extension, flexion). Muhle grades were recorded as: grade 1, partial obliteration of the subarachnoid space without spinal cord compression; grade 2, spinal cord compression; and grade 3, severe spinal cord compression with marked cord deformation/flattening. Intramedullary T2 hyperintensity was not incorporated into the Muhle grading definition itself. The number of involved levels and the distribution of affected segments (C3-C4 through C6-C7) were recorded⁶.

Intramedullary T2 hyperintensity on neutral MRI was recorded as present or absent and analyzed as an additional imaging variable.

Reliability Analysis

Because Muhle grading on dynamic cervical MRI may be challenging and potentially subject to underestimation or overestimation, reliability was formally assessed. Image interpretation was performed in a tertiary referral teaching hospital by three spine surgeons and three fellows. Before image review, raters underwent repeated calibration sessions focused on the Muhle criteria and the study-specific grading approach used in this cohort. Interobserver agreement was assessed across all readers using intraclass correlation coefficients (ICC). Intraobserver agreement was evaluated in the principal readers using repeated grading rounds. We report both single-measure and average-measure ICC values, and exact agreement was also calculated for repeated intraobserver ratings.

Neutral vs Dynamic Compression Patterns

We defined: (1) Cord compression on neutral MRI; (2) cord compression on dynamic MRI (flexion and/or extension); (3) conversion to compression (no cord compression on neutral MRI with cord compression on dynamic sequences; neutral Muhle grade 0–1 to dynamic Muhle grade 2–3); (4) increased compression (worsening of compression severity on dynamic imaging compared with neutral imaging; dynamic Muhle grade > neutral Muhle grade); and (5) additional level (presence of an extra compressed level detected only on dynamic sequences and not identified as compressed on neutral MRI). These variables were recorded as binary (present/absent).

Radiographic Instability

Radiographic cervical instability was assessed using the White-Panjabi criteria based on available dynamic radiographs obtained as part of routine care. Instability was recorded as present/absent per these criteria¹³.

Data Management

Clinical and imaging variables were abstracted from medical records and imaging into a de-identified database. Data cleaning included range checks, consistency verification, and coding of categorical variables prior to analysis. Missing data were handled using complete-case analysis for each test/model.

Statistical Analysis

Continuous variables were summarized using mean $\pm$ standard deviation or median [interquartile range], as appropriate. Categorical variables were summarized as frequencies and percentages. Associations between dynamic Muhle grade and clinical/imaging variables were tested using Spearman correlation, Mann-Whitney U tests, and Fisher exact tests as appropriate. As an exploratory diagnostic performance analysis, neutral MRI was evaluated against dynamic MRI as the comparator reference for detecting spinal cord compression. Compression was defined as Muhle grade 2–3 and no compression as Muhle grade 0–1. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of neutral MRI were calculated with 95% confidence intervals. Ordinal logistic regression models were constructed with mJOA category as the dependent outcome and dynamic Muhle grade as the principal explanatory variable. The primary adjusted model included dynamic Muhle grade and was adjusted for age and sex. A secondary exploratory model additionally incorporated T2 hyperintensity on neutral MRI. Interobserver and intraobserver agreement were assessed using ICC with 95% confidence intervals. Results are reported as odds ratios with 95% confidence intervals. Statistical significance was set at $p < 0.05$ (two-sided). Analyses were performed using jamovi (version 2.6.44; The jamovi project).

Results

Fifty patients were included in the final analysis after screening 54 consecutive adults who underwent flexion–extension dynamic cervical MRI; 3 were excluded and 1 was removed after review (Figure 2). The mean age was 58.5 ± 9.5 years and 31 patients (62%) were female (Table 1). Clinical severity by mJOA category was most frequently mild myelopathy (42%), followed by no myelopathy (28%), moderate (18%), and severe (12%) (Table 1).

Dynamic MRI identified cord compression more frequently than neutral MRI (90% vs 72%). Conversion from no compression on neutral MRI to compression on dynamic MRI occurred in 18% (9/50), and increased compression severity on dynamic sequences was observed in 42% (21/50). Dynamic Muhle grade (maximum across neutral/extension/flexion) was grade 1 in 10% (n=5), grade 2 in 28% (n=14), and grade 3 in 62% (n=31) (Table 2; Figure 3; Figure 4).

Regarding segmental distribution, the most commonly affected level was C5–C6 (74%), followed by C4–C5 (62%), C6–C7 (50%), and C3–C4 (44%) (Table 2; Figure 5). An additional compressed level detected only on dynamic sequences (“additional level”) was present in 28% (14/50) and occurred exclusively among patients with dynamic Muhle grade 3 (14/14).

Neutral MRI underestimated dynamic severity. Among patients classified as dynamic Muhle grade 2, 50% had no cord compression on neutral MRI; among dynamic Muhle grade 3, 6.5% had no compression on neutral MRI, corresponding to 18% (9/50) occult compressions on neutral imaging.

As an exploratory diagnostic performance analysis using dynamic MRI as the comparator reference, neutral MRI showed a sensitivity of 80.0% (95% CI 65.4–90.4), specificity of 100% (95% CI 47.8–100), positive predictive value of 100% (95% CI 90.3–100), and negative predictive value of 35.7% (95% CI 12.8–64.9) for detecting spinal cord compression.

T2 hyperintensity on neutral MRI was present in 60% of patients (30/50). Its frequency increased across dynamic Muhle categories, being present in 20.0% of patients with dynamic Muhle grade 1, 28.6% with grade 2, and 80.6% with grade 3. T2 hyperintensity on neutral MRI correlated with higher dynamic Muhle grade (Spearman $\rho=0.536$, $p<0.001$) and with worse mJOA category (Spearman $\rho=0.351$, $p=0.012$). When dynamic Muhle grade was dichotomized as grade 3 versus grades 1–2, T2 hyperintensity was associated with higher odds of severe dynamic compression (OR 11.67; 95% CI 3.01–45.24; $p<0.001$).

Interobserver agreement across all readers was moderate at the single-rater level and high at the averaged-rater level. For neutral Muhle grading, the single-measure ICC was 0.649 (95% CI 0.519–0.748) and the average-measure ICC was 0.917

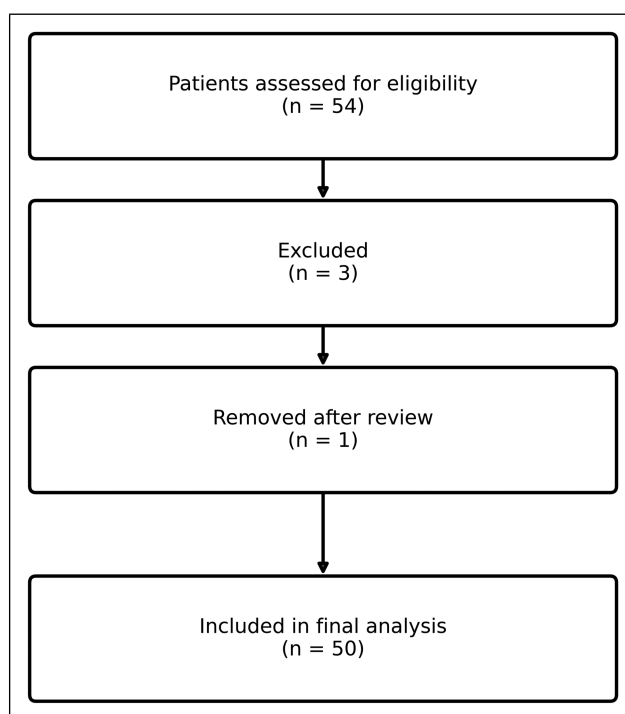


Figure 2. Study flow diagram. A total of 54 consecutive adult patients with cervical spinal stenosis who underwent flexion–extension dynamic cervical MRI were assessed for eligibility; 3 were excluded and 1 was removed after review, leaving 50 patients included in the final analysis

Table 1. Baseline Demographic and Clinical Characteristics (n = 50)

Variable	Overall cohort (n=50)
Age, mean $\pm$ SD (years)	58.5 $\pm$ 9.5
Female sex, n (%)	31 (62%)
mJOA category, n (%)	No myelopathy: 14 (28%); Mild: 21 (42%); Moderate: 9 (18%); Severe: 6 (12%)
Cervicalgia, n (%)	28 (56%)
Hoffmann and/or Trömner positive, n (%)	28 (56%)
Increased deep tendon reflexes, n (%)	19 (38%)
Decreased deep tendon reflexes, n (%)	3 (6%)
Sensory findings (C5-T1), n (%)	36 (72%)
Decreased grip strength (age- and sex-adjusted), n (%)	25 (50%)

(95% CI 0.866–0.947). For dynamic Muhle grading, the single-measure ICC was 0.618 (95% CI 0.449–0.730) and the average-measure ICC was 0.907 (95% CI 0.830–0.942). Intraobserver agreement in the principal readers was excellent. For JR, the single-measure ICC was 0.895 (95% CI 0.811–0.948) and the average-measure ICC was 0.977 (95% CI 0.955–0.989), with exact agreement in 93.5% of repeated ratings. For D, the single-measure ICC was 0.908 (95% CI 0.817–0.969) and the average-measure ICC was 0.980 (95% CI 0.957–0.994), with exact agreement in 93.5% of repeated ratings.

In bivariate analyses, dynamic Muhle grade was associated with the number of involved levels (Spearman $\rho=0.259$, $p=0.069$) and strongly associated with the presence of an additional level (Fisher exact $p<0.001$). Radiographic instability was not significantly associated with dynamic Muhle grade (Fisher exact $p=0.533$).

Dynamic Muhle grade correlated strongly with clinical severity by mJOA category (Spearman $\rho=0.684$, $p<0.001$). Higher Muhle grades were associated with long-tract signs (Hoffmann/Trömner, $p<0.001$; increased reflexes, $p=0.004$) and cervicalgia ($p=0.043$), while sensory findings ($p=0.280$), decreased reflexes ($p=0.183$), and decreased grip strength by dynamometry ($p=0.061$) were not significantly associated (Table 3).

In the primary ordinal logistic regression model, with mJOA category as the dependent outcome and dynamic Muhle grade as the principal predictor, higher dynamic Muhle grade remained independently associated with worse clinical severity after adjustment for age and sex (adjusted OR 24.10; 95% CI 4.85–119.76; $p<0.001$). Age and sex were not significantly associated with mJOA category in the adjusted model. In a secondary exploratory model that additionally included T2 hyperintensity on neutral MRI, dynamic Muhle grade remained independently associated with worse mJOA category

Table 2. Dynamic MRI Findings and Segmental Distribution

Dynamic Muhle grade	n (%)
Grade 1	5 (10%)
Grade 2	14 (28%)
Grade 3	31 (62%)
Neutral vs dynamic compression patterns	n (%)
Cord compression on neutral MRI	36 (72%)
Cord compression on dynamic MRI	45 (90%)
Conversion to compression (neutral no $\rightarrow$ dynamic yes)	9 (18%)
Increased compression on dynamic sequences	21 (42%)
Additional compressed level (dynamic only)	14 (28%)
Level distribution	n (%)
C3-C4	22 (44%)
C4-C5	31 (62%)
C5-C6	37 (74%)
C6-C7	25 (50%)

Dynamic MRI includes flexion and/or extension sagittal T2 sequences.

"Additional level" indicates a compressed segment detected only on dynamic sequences and not identified as compressed on neutral MRI.

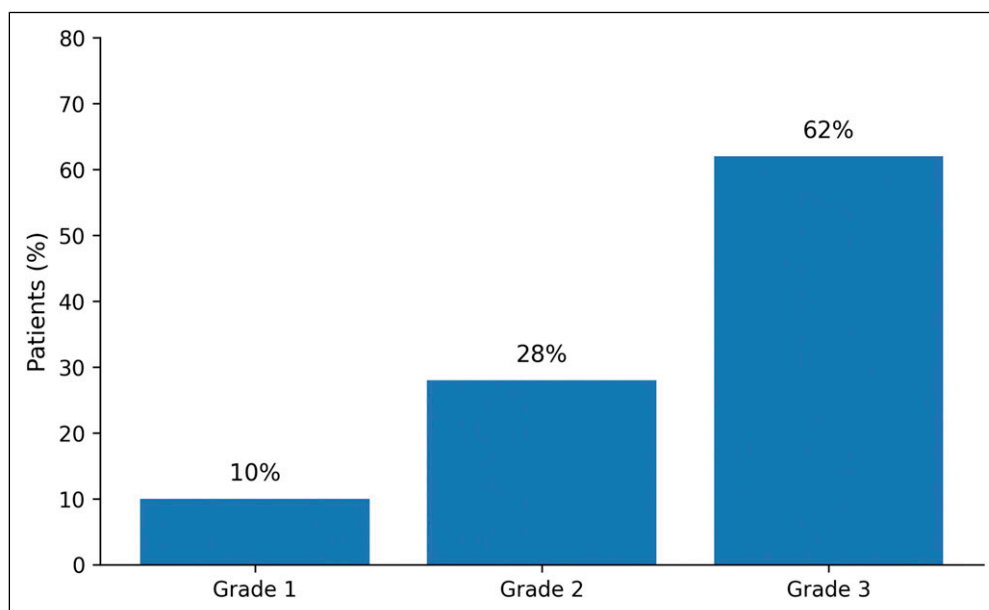


Figure 3. Distribution of dynamic Muhle grades (n = 50). Bar chart showing the proportion of patients classified as dynamic Muhle grades 1, 2, and 3 (maximum grade across neutral/extension/flexion)

(adjusted OR 26.44; 95% CI 4.84–144.27; $p < 0.001$), whereas T2 hyperintensity did not retain an independent association after adjustment (adjusted OR 0.79; 95% CI 0.20–3.05; $p = 0.730$) (Table 4).

Discussion

In this institutional cohort, flexion–extension dynamic MRI frequently revealed positional cervical cord compression that was underestimated on neutral imaging. The main findings were: (1) dynamic MRI increased detection of cord compression and identified occult compression in 9/50 (18%) of patients; (2) higher dynamic Muhle grade was strongly associated with worse mJOA category in the revised primary model; and (3) formal reliability analysis showed moderate interobserver agreement across the whole reading group and excellent intraobserver agreement in the principal readers.

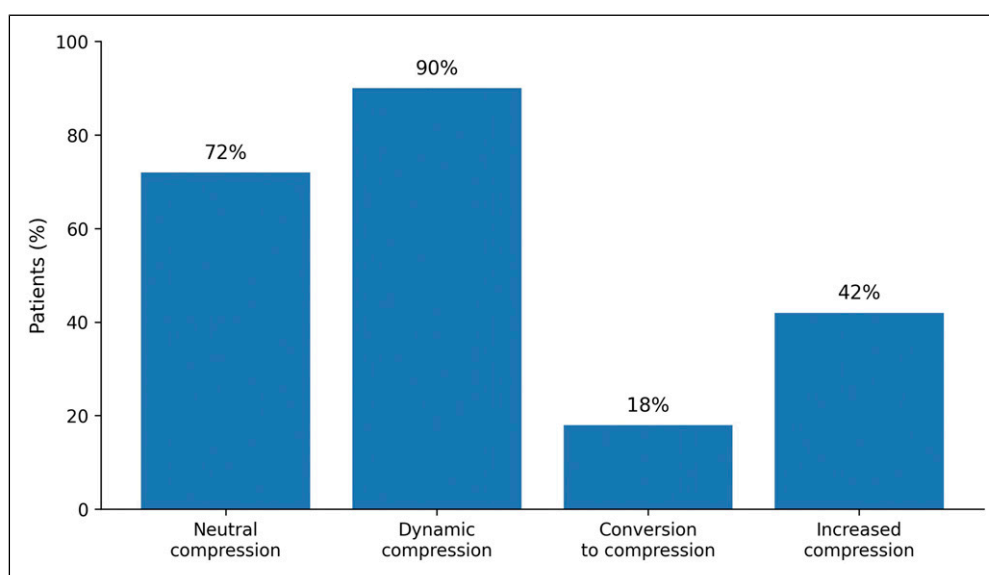


Figure 4. Neutral versus dynamic MRI detection of cord compression. Comparison of compression frequency on neutral MRI and dynamic MRI, highlighting conversion to compression and increased compression detected only on dynamic sequences

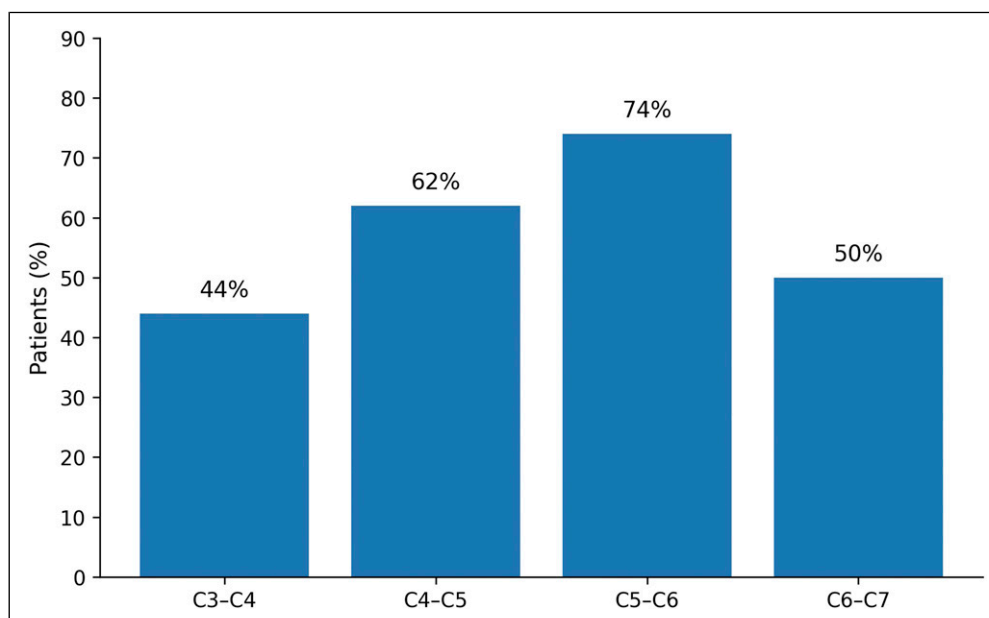


Figure 5. Segmental distribution of involved levels. Frequency of affected cervical segments (C3-C4 through C6-C7) according to the dynamic assessment

Prior studies using flexion–extension MRI have described meaningful grade shifts with motion. Zeitoun et al reported that 22.5% of patients classified as Muhle grade 1 on neutral MRI were reclassified as grade 3 on dynamic imaging. In our cohort, the most frequent transition was grade 2→3 (12/50, 24%), while grade 1→3 occurred in 2/50 (4%). Notably, neutral MRI underestimated cord compression in 18% of patients (9/50), supporting targeted use of dynamic MRI when clinical severity appears discordant with neutral imaging.³

Dynamic MRI has been proposed as a complementary study when symptoms and neutral MRI are discordant. Our results reinforce the concept that cervical stenosis is not purely static: extension and flexion can alter canal dimensions, ligamentum flavum configuration, and the effective reserve around the spinal cord, leading to clinically meaningful changes in compression severity. The higher frequency of occult or exacerbated compression observed in this cohort may reflect referral patterns typical of tertiary spine services, where patients often present with more advanced degenerative disease^{1,2,4,14}.

The “additional level” finding has practical implications. Detecting a compressed level only in dynamic positions may influence characterization of disease extent, surgical planning (e.g., level selection), and counseling, particularly in patients with severe dynamic compression. In our cohort, all additional levels were observed in dynamic Muhle grade 3, suggesting that this variable captures a phenotype of more extensive, posture-dependent disease rather than incidental imaging variation.^{2,5}

Clinically, dynamic Muhle severity aligned with established myelopathic markers. The strong association between dynamic Muhle grade and mJOA category persisted after adjustment for age and sex, indicating that dynamic Muhle grading carries functional relevance beyond descriptive imaging. Associations with Hoffmann/Trömmner and hyperreflexia further support biological plausibility, as these signs reflect corticospinal tract involvement typical of degenerative cervical

Table 3. Associations Between Dynamic Muhle Grade and Clinical/Radiologic Variables

Variable	Test	Effect/summary	p-value
mJOA category vs dynamic Muhle grade	Spearman	$\rho = 0.684$	<0.001
Hoffmann/Trömmner (present) vs Muhle grade	Fisher exact		<0.001
Increased reflexes (present) vs Muhle grade	Fisher exact		0.004
Cervicalgia (present) vs Muhle grade	Fisher exact		0.043
Number of involved levels vs Muhle grade	Spearman	$\rho = 0.259$	0.069
Radiographic instability vs Muhle grade	Fisher exact	OR ≈ 1.50	0.533
T2 hyperintensity on neutral MRI vs dynamic Muhle grade	Spearman	$\rho = 0.536$	<0.001

OR indicates odds ratio.

Table 4. Ordinal Logistic Regression Models for mJOA Category (Outcome)

Predictor	Adjusted OR	95% CI	p-value
Model 1: Primary Adjusted Model			
Dynamic Muhle grade	24.10	4.85–119.76	<0.001
Model 2: Exploratory adjusted model including T2 hyperintensity on neutral MRI.			
Dynamic Muhle grade	26.44	4.84–144.27	<0.001
T2 hyperintensity on neutral MRI	0.79	0.20–3.05	0.730

Outcome is mJOA category. Models adjusted for age and sex.

myelopathy. Conversely, sensory symptoms and grip strength did not show statistically significant associations in the primary analysis, consistent with the heterogeneity of clinical phenotypes and the potential contribution of concomitant radiculopathy or non-myelopathic pain generators.^{11,15-17}

A key revision in the present study was the reframing of the primary regression analysis so that clinical severity (mJOA category) was modeled as the outcome and dynamic Muhle grade as the explanatory variable. This approach is statistically more consistent with the intended clinico-radiological interpretation and confirmed that greater dynamic compression severity was strongly associated with worse functional status, even after adjustment for age and sex.

We also incorporated T2 hyperintensity on neutral MRI as an additional imaging variable. T2 signal change was more frequent in patients with higher dynamic Muhle grades and correlated with worse mJOA category in unadjusted analyses, supporting its relevance as a complementary radiologic marker. However, T2 hyperintensity did not retain an independent association in the exploratory adjusted model, which may reflect the limited sample size and the close relationship between cord signal change and dynamic compression severity.

Muhle grading on dynamic cervical MRI can be challenging, particularly in borderline cases in which compression severity may be underestimated or overestimated. This is especially relevant in flexion-extension imaging, where positional changes and subtle differences in cord contour may complicate classification. In the present study, reliability was formally evaluated in the context of a teaching hospital, where image interpretation involved both attending spine surgeons and fellows. Interobserver agreement across the whole group was moderate at the individual-reader level and high at the averaged-reader level, which likely reflects the variability expected among observers with different levels of experience. In contrast, the principal readers showed excellent intraobserver agreement. We believe that the repeated calibration sessions and continuous case-based teaching focused on the Muhle system may have contributed to the reproducibility observed. These findings suggest that, although the classification is not free from subjectivity, structured training and standardized review may improve its consistency in clinical and academic practice.

Radiographic instability was not associated with dynamic Muhle grade in this cohort. This may indicate that posture-dependent compression can occur even in the absence of translational instability detectable on radiographs, potentially driven by degenerative changes such as disc-osteophyte complex, alignment, and ligamentum flavum buckling. Therefore, dynamic MRI may provide complementary information that is not captured by standard instability criteria.^{13,18}

These findings support a targeted approach to dynamic cervical MRI. Candidates may include patients with suspected degenerative cervical myelopathy or long-tract signs when neutral MRI underestimates clinical severity, when the level(s) responsible are uncertain, or when additional-level disease could alter management decisions. Standardizing acquisition and structured reporting with dynamic Muhle grading may improve reproducibility and facilitate clinico-radiological correlation in selected patients.^{6,19,20}

This study has limitations. The retrospective, cross-sectional design limits causal inference and precludes prognostic conclusions. The sample size was modest and derived from a single institutional cohort, which limits statistical precision and generalizability. Although T2 hyperintensity on neutral MRI was incorporated as an additional imaging variable, it was not included within the Muhle grading definition itself. Interobserver agreement was assessed across the full reading group, whereas intraobserver agreement was evaluated only in the principal readers. The excellent intraobserver agreement observed under structured training conditions should be interpreted cautiously, as repeated calibration sessions and reader familiarity may have contributed to the observed consistency^{21,22}. In addition, the effective range of flexion–extension achieved during imaging was not quantified, and the Muhle system primarily captures central canal compression rather than dynamic foraminal stenosis. Larger prospective studies incorporating longitudinal outcomes, quantitative motion metrics, and broader imaging biomarkers are needed to better define the incremental clinical value of dynamic MRI.

Generalisability: Because this was a single-center tertiary referral cohort and dynamic MRI was performed in selected patients, generalizability to broader cervical stenosis populations may be limited. Nevertheless, the teaching-hospital setting

allowed evaluation of both group-level reproducibility and principal-reader consistency under structured calibration conditions.

Conclusions

Flexion–extension dynamic MRI identified clinically relevant, posture-dependent cervical cord compression in patients with cervical spinal stenosis that was frequently underestimated on neutral imaging. Higher dynamic Muhle grades were strongly associated with worse clinical severity as measured by mJOA, even after adjustment for age and sex. Neutral T2 hyperintensity provided additional radiologic characterization of the cohort but did not remain independently associated with mJOA in the exploratory adjusted model. These findings support selective use of dynamic MRI to improve characterization and clinico-radiological correlation in selected patients with cervical spinal stenosis, particularly when degenerative cervical myelopathy is suspected and neutral MRI findings appear discordant with the clinical picture.

ORCID iDs

José Ramón Rodríguez-Fonseca  <https://orcid.org/0009-0004-7287-4687>

Daniel Alberto Reyes-Navarro  <https://orcid.org/0000-0001-9471-1177>

Armando Saúl Ruiz-Treviño  <https://orcid.org/0000-0003-1895-6863>

Alejandro Miranda González  <https://orcid.org/0000-0003-1538-3633>

José Oscar Chávez-Bejar  <https://orcid.org/0009-0005-3187-0097>

Alexis Agustín Altamirano Carrizoza  <https://orcid.org/0009-0002-3985-9317>

Christian Gabriel de Jesús Ordaz-Pérez  <https://orcid.org/0009-0006-5137-1436>

Ethical Considerations

Approved by CEI-080-2024 and CI/074/2024.

Consent to Participate

Informed consent was waived due to the retrospective use of de-identified data.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability Statement

De-identified data that support the findings of this study are available from the corresponding author upon reasonable request.

Use of AI tools

Generative AI tools (ChatGPT) were used to assist with language editing and drafting of the manuscript. All content was reviewed, edited, and verified by the authors, who take full responsibility for the accuracy and integrity of the work. No AI tools were used to generate or alter study data, statistical analyses, or original clinical images.

Supplemental Material

Supplemental material for this article is available online.

References

1. Mahdavi A, Rasti S. Dynamic Flexion-Extension Magnetic Resonance Imaging of the Cervical Spine: An Evolutionary Tool for Diagnosis and Management of Cervical Spondylotic Myelopathy. *World Neurosurg.* 2024;184:138-147.
2. Shin JJ, Yoo SJ, Kim TW, et al. Radiological and Clinical Significance of Cervical Dynamic Magnetic Resonance Imaging for Cervical Spondylotic Myelopathy. *Neurospine.* 2024;21(2):443-454.
3. Zeitoun D, Hajj FE, Sariali E, Catonné Y, Pascal-Moussellard H. Evaluation of spinal cord compression and T2 hyperintense lesions in CSM using flexion-extension MRI. *Spine J.* 2015;15(4):668-674.

4. Martinez SS, Krecke KN, Syc-Mazurek SB, et al. Flexion-Extension MRI to Reveal Occult Spondylotic Compression in Undifferentiated Cervical Myelopathies With Cord T2 Hyperintensity. *Neurol Clin Pract.* 2025;15(6):e200555.
5. Hayashi T, Wang JC, Suzuki A, et al. Risk Factors for Missed Dynamic Canal Stenosis in the Cervical Spine. *Spine.* 2014;39(10):812-819.
6. Muhle C, Metzner J, Weinert D, et al. Classification system based on kinematic MR imaging in cervical spondylitic myelopathy. *AJNR Am J Neuroradiol.* 1998;19(9):1763-1771.
7. Nouri A, Tetreault L, Singh A, Karadimas SK, Fehlings MG. Degenerative Cervical Myelopathy. *Spine.* 2015;40(12):E675-E693.
8. Gibson J, Nouri A, Krueger B, et al. Degenerative Cervical Myelopathy: A Clinical Review. *Yale J Biol Med.* 2018;91(1):43-48.
9. Lázaro MLM, Lázaro MAP, Losantos FB, Bayo AP, Jarabo MLMLS de NHOPAR. Nuevas tablas de fuerza de la mano para población adulta de Teruel NEW TABLES ON HAND STRENGTH IN THE ADULT POPULATION FROM TERUEL. *Nutr Hosp.* 2008;23(1):35-40.
10. Kato S, Oshima Y, Oka H, et al. Comparison of the Japanese Orthopaedic Association (JOA) Score and Modified JOA (mJOA) Score for the Assessment of Cervical Myelopathy: A Multicenter Observational Study. *PLoS One.* 2015;10(4):e0123022.
11. Tetreault L, Kopjar B, Nouri A, et al. The modified Japanese Orthopaedic Association scale: establishing criteria for mild, moderate and severe impairment in patients with degenerative cervical myelopathy. *Eur Spine J.* 2017;26(1):78-84.
12. Peña OMP. Medición dinámica y comparación de los cambios de diámetro del conducto espinal cervical mediante RM T2 en neutra. In: *flexión y extensión en estenosis cervical [Thesis]*; 2024.
13. Panjabi MM, White AA. Basic Biomechanics of the Spine. *Neurosurgery.* 1980;7(1):76-93.
14. Baram A, Choueiri JE, Brembilla C, et al. Dynamic MRI in Degenerative Cervical Myelopathy: A Systematic Review of Radiological Markers, Correlations, and Outcomes. *J Clin Med.* 2025;15(1):265.
15. Tetreault L, Goldstein CL, Arnold P, et al. Degenerative Cervical Myelopathy. *Neurosurgery.* 2015;77(Suppl 1):S51-S67.
16. Tejus MN, Singh V, Ramesh A, Kumar VRR, Maurya VP, Madhugiri VS. An evaluation of the finger flexion, Hoffmann's and plantar reflexes as markers of cervical spinal cord compression – A comparative clinical study. *Clin Neurol Neurosurg.* 2015;134:12-16.
17. Chang CW, Chang KY, Lin SM. Quantification of the Trömner signs: a sensitive marker for cervical spondylotic myelopathy. *Eur Spine J.* 2011;20(6):923-927.
18. Kalidindi KKV, Sharma JK, Vishwakarma G, Duggi S, Chhabra HS. Ligamentum Flavum Buckling vs Hypertrophy in Cervical Myelopathy Using Dynamic Imaging: Are Instability and Loss of Disc Height the Culprits? *Int J Spine Surg.* 2022;16(2):300-308.
19. Fehlings MG, Evaniew N, Weng PV, et al. AO Spine Clinical Practice Recommendations for Diagnosis and Management of Degenerative Cervical Myelopathy: Evidence Based Decision Making. *Global Spine J.* 2025;15(5):2585-2593.
20. Rodrigues-Pinto R, Montenegro TS, Davies BM, et al. Optimizing the Application of Surgery for DCM (AO Spine RECODE-DCM Priority #10). *Global Spine J.* 2022;12(1_suppl):147S-158S.
21. Kong X, Liu Z, Song K, et al. T2-weighted MRI high signal in CSM is associated with dynamic change. *J Orthop Surg Res.* 2025;20(1):360.
22. Morgado A, Berthiller J, Subtil F, et al. A New Multi-Parametric MRI-Based Scoring System for DCM: SIMS. *Brain Sci.* 2025;15(6):557.