European Spine Journal (2024) 33:1205–1212
https://doi.org/10.1007/s00586-023-08067-8

ORIGINAL ARTICLE

Comparison of the patient-derived modified Japanese Orthopaedic Association scale and the European myelopathy score

Eddie de Dios[1,2] · Håkan Löfgren[3,4] · Mats Laesser[1,2] · Lars Lindhagen[5] · Isabella M. Björkman-Burtscher[1,2] · Anna MacDowall[6,7]

Received: 1 August 2023 / Revised: 8 November 2023 / Accepted: 27 November 2023 / Published online: 19 December 2023

Abstract

Purpose To compare the patient-derived modified Japanese Orthopaedic Association (P-mJOA) scale with the European myelopathy score (EMS) for the assessment of patients with degenerative cervical myelopathy (DCM).

Methods In this register-based cohort study with prospectively collected data, included patients were surgically treated for DCM and had reported both P-mJOA and EMS scores at baseline, 1-year follow-up, and/or 2-year follow-up to the Swedish Spine Register. P-mJOA and EMS scores were defined as severe (P-mJOA 0–11 and EMS 5–8), moderate (P-mJOA 12–14 and EMS 9–12), or mild (P-mJOA 15–18 and EMS 13–18). P-mJOA and EMS mean scores were compared, and agreement was evaluated with Spearman's rank correlation coefficient (ρ), the intraclass correlation coefficient (ICC), and kappa (κ) statistics.

Results Included patients ($n=714$, mean age 63.2 years, 42.2% female) completed 937 pairs of the P-mJOA and the EMS. The mean P-mJOA and EMS scores were 13.9 ± 3.0 and 14.5 ± 2.7, respectively (mean difference –0.61 [95% CI –0.72 to –0.51; $p<0.001$]). Spearman's ρ was 0.84 ($p<0.001$), and intra-rater agreement measured with ICC was 0.83 ($p<0.001$). Agreement of severity level measured with unweighted and weighted κ was fair ($\kappa=0.22$ [$p<0.001$]; $\kappa=0.34$ [$p<0.001$], respectively). Severity levels were significantly higher using the P-mJOA ($p<0.001$).

Conclusion The P-mJOA and the EMS had similar mean scores, and intra-rater agreement was high, whereas severity levels only demonstrated fair agreement. The EMS has a lower sensitivity for detecting severe myelopathy but shows an increasing agreement with the P-mJOA for milder disease severity. A larger interval to define severe myelopathy with the EMS is recommended.

Keywords Patient-derived modified Japanese Orthopaedic Association · European myelopathy score · Degenerative cervical myelopathy · Intra-rater agreement · Severity grading

Isabella M. Björkman-Burtscher and Anna MacDowall these authors have equally contributed to this work.

✉ Eddie de Dios
eddie.dedios@gmail.com

[1] Department of Radiology, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[2] Department of Radiology, Sahlgrenska University Hospital, Region Västra Götaland, Bruna Stråket 11, 41345 Gothenburg, Sweden

[3] Neuro-Orthopedic Center, Region Jönköping County, Jönköping, Sweden

[4] Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden

[5] Uppsala Clinical Research Center, Uppsala University, Uppsala, Sweden

[6] Department of Surgical Sciences, Uppsala University, Uppsala, Sweden

[7] Department of Orthopedics, Uppsala University Hospital, Uppsala, Sweden

Introduction

Degenerative cervical myelopathy (DCM), previously referred to as cervical spondylotic myelopathy, comprises a range of age-related pathologies that affect the cervical spine and lead to spinal cord compression and neurological dysfunction [1, 2]. With an aging population, DCM contributes substantially to global disability, reduced quality of life, and healthcare costs [3]. The Japanese Orthopaedic Association (JOA) scale [4], the modified Japanese Orthopaedic Association (mJOA) scale [5], and the European myelopathy Score (EMS) [6] classify the severity of myelopathy [7–9]. Following adaptation for Western cultures, the mJOA is currently one of the most widely accepted assessment tools for myelopathy and functional status in patients with DCM [10, 11]. The mJOA is, however, physician-administered and not patient-reported, a drawback that can introduce classification bias. Further types of bias, such as recollection bias or reporting bias, might be added if data are collected retrospectively or based on patient records. Patient-reported outcome measures (PROMs) like the EMS overcome these problems, and the EMS has been used in the Swedish Spine Register (Swespine) since 2006 [12].

With the recent development and validation of the patient-derived modified Japanese Orthopaedic Association (P-mJOA) scale [13], it is now possible to use a myelopathy scale that is considered gold standard *and* patient-reported. Further, as the Swedish translation of the P-mJOA and the EMS has been used in parallel in Swespine since the fall of 2020, a comparison of these two PROMs is now appropriate. In case of concordance, it would also be possible to translate large amounts of long-term EMS-based follow-up data and findings into a more international setting. The comparison with the P-mJOA is therefore fundamental to legitimize and increase the transferability of previous findings, while also serving as a validation of the P-mJOA from a European point of view. For these purposes, the aim of this study was to compare the P-mJOA with the EMS for the assessment of DCM.

Methods

This study was approved by the Swedish Ethical Review Authority (2022-06976-01). Written informed consent was waived by the authority for this register study.

Study population and data collection

Data were extracted from Swespine, a national register where cervical spine surgeries have been registered since 2006. More than 95% of all spine units in Sweden, including all major units, are affiliated to the register, which covers around 85% of all spine surgeries in the country. The register is governed by the Swedish Society of Spinal Surgeons (www.4s.nu) with public financial support [14]. Patients complete baseline questionnaires and PROMs at baseline and 1, 2, 5, and 10 years after surgery [12]. Following a thorough translation process of the P-mJOA that consisted of six steps (Table 1)**,** Swespine registers since November 2020 the P-mJOA in parallel with the EMS. Details of the P-mJOA and the EMS are given in Table 2, and the Swedish translation of the P-mJOA is presented in Supplementary Table 1.

All patients who completed both the P-mJOA and the EMS since the implementation of P-mJOA in Swespine in November 2020 were eligible. Included patients had been surgically treated for DCM and had reported complete P-mJOA and EMS scores at baseline, 1-year follow-up, and/or 2-year follow-up to Swespine before the register extraction date for this study (March 15, 2023). Patients who did not have complete scores and subscores of both P-mJOA and EMS at the same timepoint were excluded. In order to minimize bias, no other exclusion criteria were used.

Table 1 The translation process of the patient-derived modified Japanese Orthopaedic Association (P-mJOA) to the Swedish version used in the Swedish Spine register

1	Translation of the form from English to Swedish by a professional translator
2	A debriefing meeting where versions, synonyms, and variants of the translation were discussed, and a preliminary version was created
3	Reverse translation of the Swedish version into English by a professional translator
4	New debriefing meeting where versions, synonyms, and variants of the translation were discussed based on the reverse translation, and an adjusted Swedish version was created
5	Qualitative evaluation by interviewing patients who could complete the form unassisted and communicate their impressions and suggestions for improvements and formulations
6	Compilation of the qualitative evaluation and posting of the final version of the form

The translation process was carried out by Anna MacDowall (Department of Surgical Sciences, Uppsala University, Uppsala, Sweden and Department of Orthopedics, Uppsala University Hospital, Uppsala, Sweden) and Håkan Löfgren (Neuro-Orthopedic Center, Jönköping, Region Jönköping County, Sweden and Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden)

Table 2 Comparison of the patient-derived modified Japanese Orthopaedic Association (P-mJOA) scale and the European myelopathy score (EMS)

P-mJOA [13]		EMS [6]	
Subscore item	Subscore	Subscore item	Subscore
Upper extremity motor function		Gait function	
Unable to move my hands	0	Unable to walk, wheelchair	1
Unable to eat with a spoon but am able to move my hands	1	Walking on flat ground only with cane or aid	2
Unable to button my shirt but able to eat with a spoon	2	Climbing stairs only with aid	3
Able to button my shirt with great difficulty	3	Gait clumsy, but no aid necessary	4
Able to button my shirt with slight difficulty	4	Normal walking and climbing stairs	5
Not having any trouble using my hands	5		
Lower extremity function		Bladder and bowel function	
Completely unable to move legs at all and have no feeling in legs	0	Retention, no control over bladder and/or bowel function	1
Having feeling in legs but not able to move them at all	1	Inadequate micturition and urinary frequency	2
Able to move my legs but am unable to walk	2	Normal bladder and bowel function	3
Able to walk on flat floor with a walking aid (cane or crutch)	3	Hand function	
Able to walk up- and/or downstairs with aid of a handrail	4	Handwriting and eating with knife and fork impossible	1
Able to walk up- and/or downstairs without handrail but I notice moderate to significant lack of stability/feeling of imbalance when I walk	5	Handwriting and eating with knife and fork impaired	2
Able to walk unaided (no crutches, canes, walker) with smooth reciprocation (i.e., legs move smoothly) but I still notice mild lack of stability/feeling of imbalance when walking	6	Handwriting, tying shoelaces or a tie clumsy	3
Able to walk without any problems of imbalance or instability	7	Normal handwriting	4
Upper extremity sensory function		Proprioception and coordination	
Complete loss of feeling in hands	0	Getting dressed only with aid	1
Severe loss of feeling, or have pain in my hands	1	Getting dressed clumsily and slowly	2
Mild loss of feeling in hands	2	Getting dressed	3
No loss of feeling in hands	3	Paresthesia/pain	
Urinary function		Invalidity due to pain	1
Am completely unable to control urination	0	Endurable paresthesia and pain	2
Have marked difficulty controlling urination	1	No paresthesia and pain	3
Have mild to moderate difficulty controlling urination	2		
No difficulty controlling urination	3		

Extracted register variables are given in Table 3. P-mJOA and EMS scores are defined as severe (P-mJOA 0–11 and EMS 5–8), moderate (P-mJOA 12–14 and EMS 9–12), or mild (P-mJOA 15–18 and EMS 13–18) [6, 9, 13, 15–17].

Statistical analysis

Baseline data and P-mJOA and EMS scores at each time-point (baseline, 1-year follow-up, and 2-year follow-up) are presented descriptively. Categorical variables are reported with frequency and percentage. Continuous variables are reported with mean and SD.

The P-mJOA and EMS scores were compared using a two-sided paired *t*-test with a 95% confidence interval (CI) and a statistical significance level set to $p \leq 0.05$. The scales were also compared with Spearman's rank correlation coefficient (ρ), the intraclass correlation coefficient (ICC), and the kappa (κ) statistic. Spearman's ρ was used to estimate the strength and direction of association between the P-mJOA and the EMS. The ICC was used as a measure of intra-rater agreement. The κ statistic was used to evaluate agreement by severity level, with a κ statistic less than 0.00 indicating poor agreement, 0.00–0.20 indicating slight agreement, 0.21–0.40 indicating fair agreement, 0.41–0.60 indicating moderate agreement, 0.61–0.80 indicating substantial agreement, and 0.81–1.00 indicating excellent agreement [18]. The weighted κ was also calculated to assess similarity in answers by differentiating between partial agreement (severity level differing one step, e.g., moderate versus severe) and no agreement (severity level differing two steps, i.e., mild versus severe).

Table 3 Variable list extracted from the Swedish Spine Register, including total scores and subscores of the patient-derived modified Japanese Orthopaedic Association (P-mJOA) scale and the European myelopathy score (EMS)

Parameter	Score
Sex	Female/male
Age at surgery	Years
P-mJOA and EMS scores at baseline and follow-up at 1 year and 2 years after surgery:	
P-mJOA total score	0–18
Upper extremity motor function	0–5
Lower extremity function	0–7
Upper extremity sensory function	0–3
Urinary function	0–3
EMS total score	5–18
Gait function	1–5
Bladder and bowel	1–3
Hand function	1–4
Proprioception and coordination	1–3
Paresthesia/pain	1–3

In addition, a conversion table was created to compare the scales through a proportional division of the 14 steps (range 5–18) of myelopathy measured with the EMS into the 19 steps of the P-mJOA (range 0–18). Statistical analyses were performed as above.

Results

Data extraction resulted in 714 included patients (mean age 63.2 ± 11.1 years, range 18–91 years, 42.2% female) who in total had completed 937 pairs of the P-mJOA and the EMS. An additional 80 (7.9%) potential pairs in 37 (4.9%) patients had to be excluded as data were incomplete for either one or both scales.

The mean P-mJOA score was 13.9 ± 3.0 (range 1–18), and the mean EMS was 14.5 ± 2.7 (range 5–18). The mean difference between the P-mJOA and EMS was –0.61 (95% CI –0.72 to –0.51; $p < 0.001$), and the mean differences between the scales at the three measured timepoints (baseline, 1-year follow-up, and 2-year follow-up) were consistent with the total mean difference (Table 4).

Table 4 Mean scores of the patient-derived modified Japanese Orthopaedic Association (P-mJOA) and the European myelopathy score (EMS) at different timepoints (baseline, 1-year follow-up, and 2-year follow-up) and mean score differences (paired *t*-test)

Timepoint	*n*	P-mJOA	EMS	Difference EMS—P-mJOA	
		Mean ± SD (range)	Mean ± SD (range)	Mean (95% CI)	*p*-value
All	937	13.9 ± 3.0 (1–18)	14.5 ± 2.7 (5–18)	0.6 (0.5–0.7)	< 0.001
Baseline	282	13.5 ± 3.0 (1–18)	± 2.6 (5–18)	0.7 (0.5–0.9)	< 0.001
1-y follow-up	341	14.0 ± 3.1 (4–18)	14.6 ± 2.9 (5–18)	0.6 (0.5–0.8)	< 0.001
2-y follow-up	314	14.0 ± 2.8 (4–18)	14.5 ± 2.7 (7–18)	0.5 (0.3–0.7)	< 0.001

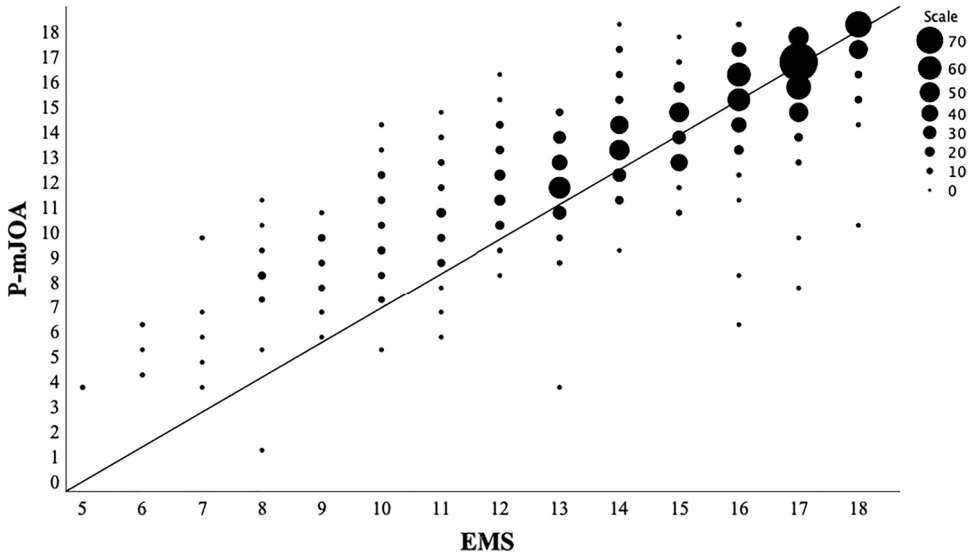

Fig. 1 Scatter plot of paired total scores of the patient-derived modified Japanese Orthopaedic Association (P-mJOA) scale and the European myelopathy score (EMS)

Figure 1 shows a scatter plot comparing P-mJOA versus EMS scores. A Spearman's ρ of 0.84 (95% CI 0.82–0.86; $p < 0.001$) demonstrated a strong monotonic association between the P-mJOA and the EMS. Intra-rater agreement measured with the ICC between the P-mJOA and the EMS was 0.83 (95% CI 0.76–0.87; $p < 0.001$).

The distribution of severity levels between the scales is given in Table 5. Severity levels in the population were significantly higher using the P-mJOA compared with the EMS (Pearson Chi-square; $p < 0.001$); however, if disease severity was classified as severe with the EMS ($n = 32$), it also classified as severe with P-mJOA. Further, moderate disease severity classified with the EMS ($n = 171$) was only four times classified as mild with P-mJOA. There was fair agreement between the P-mJOA and the EMS severity levels with both unweighted kappa ($\kappa = 0.22$; $p < 0.001$) and weighted kappa ($\kappa = 0.34$; 95% CI 0.30–0.37; $p < 0.001$).

Mean subscores for each evaluated item grouped by severity levels for the P-mJOA and the EMS are presented in Table 6. The largest mean difference in subscore points between measurement points classified as severe versus mild was seen in the lower extremity-function item for the P-mJOA (3.1 points) and in the gait-function item for the EMS (3.0 points).

In the supplementary material, a conversion table to compare the scales through a proportional division of the 14 steps (range 5–18) of myelopathy measured with the EMS into the 19 steps of the P-mJOA (range 0–18) is given

Table 5 Distribution of myelopathy severity levels between the patient-derived modified Japanese Orthopaedic Association (P-mJOA) and the European myelopathy score (EMS) at 937 measurement points in 714 patients

Severity	P-mJOA			EMS		
	Range	*n* (%)	Mean ± SD	Range	*n* (%)	Mean ± SD
Scores grouped by severity level at baseline						
Severe	1–11	65 (23.0%)	9.2 ± 2.2	5–8	11 (3.9%)	7.4 ± 1.1
Moderate	12–14	102 (36.2%)	13.1 ± 0.8	9–12	53 (18.8%)	11.1 ± 0.9
Mild	15–18	115 (40.8%)	16.3 ± 1.0	13–18	218 (77.3%)	15.3 ± 1.6
Scores grouped by severity level at the 1-year follow-up						
Severe	4–11	70 (20.5%)	9.2 ± 1.8	5–8	10 (2.9%)	6.9 ± 1.1
Moderate	12–14	106 (31.1%)	13.1 ± 0.8	9–12	65 (19.1%)	10.7 ± 1.1
Mild	15–18	165 (48.4%)	16.6 ± 1.1	13–18	266 (78.0%)	15.9 ± 1.7
Scores grouped by severity level at the 2-year follow-up						
Severe	4–11	58 (18.5%)	9.5 ± 1.7	7–8	11 (3.5%)	7.6 ± 0.5
Moderate	12–14	97 (30.1%)	13.1 ± 0.8	9–12	53 (16.9%)	10.7 ± 1.1
Mild	15–18	159 (50.6%)	16.2 ± 1.1	13–18	250 (79.6%)	15.6 ± 1.6
Scores grouped by severity level at all timepoints						
Severe	1–11	193 (20.6%)	9.3 ± 1.9	5–8	32 (3.4%)	7.3 ± 1.0
Moderate	12–14	305 (32.6%)	13.1 ± 0.8	9–12	171 (18.2%)	10.8 ± 1.1
Mild	15–18	439 (46.9%)	16.4 ± 1.1	13–18	734 (78.3%)	15.6 ± 1.6

Table 6 Mean scores for each severity level and in total, grouped by each item using the patient-derived modified Japanese Orthopaedic Association (P-mJOA) and the European myelopathy score (EMS)

Item (range)	P-mJOA, mean score (range) difference from total mean score			
Disease severity Observations	Severe *n* = 193	Moderate *n* = 305	Mild *n* = 439	Total *n* = 937
Upper extremity motor function (0–5)	2.8 (0–5) *−1.3*	4.0 (1–5) *−0.1*	4.7 (3–5) *+0.6*	4.1 (0–5) *n/a*
Lower extremity function (0–7)	3.3 (0–6) *−1.9*	4.8 (2–7) *−0.4*	6.4 (4–7) *+1.2*	5.2 (0–7) *n/a*
Upper extremity sensory function (0–3)	1.4 (0–3) *−0.7*	1.9 (0–3) *−0.2*	2.5 (0–3) *+0.4*	2.1 (0–3) *n/a*
Urinary function (0–3)	1.8 (0–3) *−0.7*	2.4 (0–3) *−0.1*	2.8 (0–3) *+0.3*	2.5 (0–3) *n/a*
Item (range)EMS, mean scores (range) difference from total mean score				
Disease severity *Observations*	*Severe* *n* = 32	*Moderate* *n* = 171	*Mild* *n* = 734	*Total* *n* = 937
Gait function (1–5)	1.4 (1–3) *−2.5*	2.5 (1–5) −1.4	4.4 (1–5) *+0.5*	3.9 (1–5) *n/a*
Bladder and bowel function (1–3)	1.8 (1–3) *−0.9*	2.3 (1–3) *−0.4*	2.8 (1–3) *+0.1*	2.7 (1–3) *n/a*
Hand function (1–4)	1.5 (1–3) *−1.8*	2.6 (1–4) *−0.7*	3.5 (1–4) *+0.2*	3.3 (1–4) *n/a*
Proprioception and coordination (1–3)	1.2 (1–3) *−1.4*	2.0 (1–3) *−0.6*	2.8 (1–3) *+0.2*	2.6 (1–3) *n/a*
Paresthesia/pain (1–3)	1.4 (1–3) *−0.6*	1.5 (1–3) *−0.5*	2.1 (1–3) *+0.1*	2.0 (1–3) *n/a*

(Supplementary Table 2). The mean score of this converted version of the EMS (cEMS) was 13.0 ± 3.9. The mean difference between the P-mJOA and the cEMS was at 0.85 (95% CI 0.72–0.98; $p < 0.001$) slightly higher than that between the P-mJOA and the EMS. The number of equal values between the P-mJOA and the cEMS was 224 (23.4%). A Spearman's ρ of 0.86 ($p < 0.001$) demonstrated a strong monotonic association between the P-mJOA and the cEMS. Intra-rater agreement with the ICC between the P-mJOA and the cEMS was 0.80 (95% CI 0.71–0.83; $p < 0.001$). By severity level, there was moderate agreement with the unweighted κ statistic ($\kappa = 0.53$; $p < 0.001$) and substantial agreement with the weighted κ statistic ($\kappa = 0.64$; 95% CI 0.60–0.68; $p < 0.001$).

Discussion

This is the first study to compare the newly developed P-mJOA with the established EMS, and our comparison shows that the scales are highly correlated and have high intra-rater agreement in terms of total scores, but only show a fair agreement when assessed by severity level. Naturally, the comparison of the P-mJOA and the EMS was complicated by the difference in number of items, item content, and score intervals between the scales. Despite this, mean scores only differed slightly more than 0.5 points, which is less than the previously reported minimum clinically important difference of the mJOA estimated to between 1 and 2 points depending on severity [19].

When comparing the scales by their classification of patients into three previously used severity levels for each scale [9, 13, 17], the P-mJOA scale classified significantly more patients as severe or moderate. This difference might be important as surgical decision making in accordance with current international guidelines suggests surgical intervention for this patient group, whereas initial non-operative treatment can be considered for patients with mild myelopathy [15]. However, our study also demonstrated that a large proportion of patients classified as having mild myelopathy (41 and 77% for P-mJOA and EMS, respectively) still undergo surgical treatment. As our study population represents a national setting along with previous studies presenting comparable mean preoperative scores [9, 13, 20–22], these results are likely to reflect a common management strategy for patients with mild myelopathy. This is also in agreement with several studies showing that less severe baseline myelopathy is a predictor of improved functional outcome [17, 23–25]. Another important aspect is that neither scale is particularly influenced by the degree of pain that the patient experiences, which in clinical practice might be a determining indication for surgery due to radiculopathy present in combination with myelopathy [15, 26–28], nor the main symptom in other conditions that may be associated with myelopathy [29].

The number of patients classified as severe with the EMS was unexpectedly low at 3.4% but can be explained by the narrow range (EMS 5–8) used to define severe myelopathy in previous literature [6, 9, 17]. In contrast, the P-mJOA uses 12 steps (range 0–11) to define severe myelopathy. In theory, the four steps of severe myelopathy using the EMS could have corresponded to the 12 steps of severe myelopathy using P-mJOA, but this study clearly contradicts that. For both scales, the largest differences at item level between patients classified as severe versus mild were seen for the corresponding lower extremity- and gait-function items. This was somewhat expected as these items have the largest point ranges in both scales. Still, this finding suggests that the degree of walking disability has the largest impact on the severity grading, and it could also be indicative of a more advanced disease progression. The importance of walking disability is further supported by the widespread use of the Nurick scale that only reports walking disability [2, 30]. There was, however, no obvious explanation for the much lower proportion of patients classified as severe with the EMS, which suggests that the interval to classify a patient as severe with the EMS is too narrow.

The proportional conversion table, converting the EMS into the cEMS, proved to be of limited value. Although the kappa statistics increased from fair agreement to substantial agreement with the cEMS, the number of patients classified as severe was instead overestimated when the 0–11 range of P-mJOA was divided into a proportionally equivalent 0–13 EMS range.

Our results indicate that the original EMS has a low sensitivity for detecting severe myelopathy and lacks the ability to subcategorize severe myelopathy compared with the P-mJOA. This limitation could partially be reduced by increasing the upper interval limit of severe myelopathy to an extent where an equal number of patients are classified as having severe myelopathy, but to fully account for the difference, the EMS would need more steps in the lower end of the scale.

Despite the large mismatch in the classification of myelopathy severity between the scales, the mean scores as well as Spearman's ρ and the ICC demonstrated strong agreement levels. This indicates that the upper half of the scales have a higher degree of agreement and possibly that the scales can be used interchangeably to define mild myelopathy and to some extent moderate myelopathy.

Several limitations with this study are related to the comparison of structurally different scales. Although the difference in ranges and intervals might be addressed through a proportional conversion, the difference in content within the items, further complicated by an additional item in the EMS, cannot easily be adjusted for. This also means that

the similarity in mean scores and correlations between the scales should be interpreted with great caution. Strengths and limitations of this study might relate to the register-based study design, consent bias, and potential selection bias related to patients not providing complete PROMs. However, less than 10% of the timepoints from the data extraction did not have complete P-mJOA and EMS data and had to be excluded. Selection bias might further occur related to compliance with completing all PROMs, which might be individually influenced positively or negatively by disease severity, motivation, or ability. Completion rates at the different timepoints were, however, evenly distributed, and the large sample size with high nationwide representation and non-exclusive selection criteria employed in this study should at least partially counteract these biases and yield generalizable findings. A register-based study design further always introduces a risk of collection and confounding biases; however, the latter should be limited as we only used paired observations. Finally, comparative subscore analyses of the different items between the scales would have been interesting but were not deemed feasible due to the structural differences.

Conclusions

In this study, the P-mJOA and the EMS had similar mean scores, and intra-rater agreement measured with the ICC was high. However, agreement by severity level measured with kappa statistics only demonstrated fair agreement. Compared with the P-mJOA, the EMS has a low sensitivity for detecting severe myelopathy but shows an increasing agreement with the P-mJOA for milder disease severity. For improved comparison between the P-mJOA and the EMS, a larger interval to define severe myelopathy with the EMS is suggested.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00586-023-08067-8.

Funding Open access funding provided by University of Gothenburg.

Declarations

Conflict of interest All authors declare that they have no conflict of interest.

References

1. Davies BM, Mowforth OD, Smith EK, Kotter MR (2018) Degenerative cervical myelopathy. BMJ 360:k186. https://doi.org/10.1136/bmj.k186
2. Badhiwala JH, Ahuja CS, Akbar MA, Witiw CD, Nassiri F, Furlan JC, Curt A, Wilson JR, Fehlings MG (2020) Degenerative cervical myelopathy - update and future directions. Nat Rev Neurol 16:108–124. https://doi.org/10.1038/s41582-019-0303-0
3. Nouri A, Cheng JS, Davies B, Kotter M, Schaller K, Tessitore E (2020) Degenerative cervical myelopathy: a brief review of past perspectives, present developments, and future directions. J Clin Med. https://doi.org/10.3390/jcm9020535
4. Hirabayashi K, Miyakawa J, Satomi K, Maruyama T, Wakano K (1981) Operative results and postoperative progression of ossification among patients with ossification of cervical posterior longitudinal ligament. Spine 6:354–364. https://doi.org/10.1097/00007632-198107000-00005
5. Benzel EC, Lancon J, Kesterson L, Hadden T (1991) Cervical laminectomy and dentate ligament section for cervical spondylotic myelopathy. J Spinal Disord 4:286–295. https://doi.org/10.1097/00002517-199109000-00005
6. Herdmann J, Linzbach M, Krzan M, Dvorak J, Bock W (1994) The European myelopathy score. In: Cerebellar Infarct Midline Tumors Minimally Invasive Endoscopic Neurosurgery (MIEN). Springer. pp 266–268
7. Furlan JC, Catharine Craven B (2016) Psychometric analysis and critical appraisal of the original, revised, and modified versions of the Japanese Orthopaedic Association score in the assessment of patients with cervical spondylotic myelopathy. Neurosurg Focus 40:E6. https://doi.org/10.3171/2016.3.Focus1648
8. Singh A, Crockard HA (2001) Comparison of seven different scales used to quantify severity of cervical spondylotic myelopathy and post-operative improvement. J Outcome Meas 5:798–818
9. Vitzthum HE, Dalitz K (2007) Analysis of five specific scores for cervical spondylogenic myelopathy. Eur Spine J 16:2096–2103. https://doi.org/10.1007/s00586-007-0512-x
10. Kato S, Oshima Y, Oka H, Chikuda H, Takeshita Y, Miyoshi K, Kawamura N, Masuda K, Kunogi J, Okazaki R, Azuma S, Hara N, Tanaka S, Takeshita K (2015) 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 10:e0123022. https://doi.org/10.1371/journal.pone.0123022
11. Kopjar B, Tetreault L, Kalsi-Ryan S, Fehlings M (2015) Psychometric properties of the modified Japanese Orthopaedic Association scale in patients with cervical spondylotic myelopathy. Spine 40:E23-28. https://doi.org/10.1097/brs.0000000000000648
12. Strömqvist B, Fritzell P, Hägg O, Jönsson B (2009) The Swedish Spine register: development, design and utility. Eur Spine J 18(Suppl 3):294–304. https://doi.org/10.1007/s00586-009-1043-4
13. Rhee JM, Shi WJ, Cyriac M, Kim JY, Zhou F, Easley KA, Patel A (2018) The P-mJOA: a patient-derived, self-reported outcome instrument for evaluating cervical myelopathy: comparison with the mJOA. Clin Spine Surg 31:E115-e120. https://doi.org/10.1097/bsd.0000000000000591
14. Svensk Ryggkirurgisk Förening. http://www.4s.nu/. Accessed July 28 2023

15. Fehlings MG, Tetreault LA, Riew KD, Middleton JW, Aarabi B, Arnold PM, Brodke DS, Burns AS, Carette S, Chen R, Chiba K, Dettori JR, Furlan JC, Harrop JS, Holly LT, Kalsi-Ryan S, Kotter M, Kwon BK, Martin AR, Milligan J, Nakashima H, Nagoshi N, Rhee J, Singh A, Skelly AC, Sodhi S, Wilson JR, Yee A, Wang JC (2017) A clinical practice guideline for the management of patients with degenerative cervical myelopathy: recommendations for patients with mild, moderate, and severe disease and Non-myelopathic patients with evidence of cord compression. Global Spine J 7:70s–83s. https://doi.org/10.1177/2192568217701914
16. Tetreault L, Kopjar B, Nouri A, Arnold P, Barbagallo G, Bartels R, Qiang Z, Singh A, Zileli M, Vaccaro A, Fehlings MG (2017) The modified Japanese Orthopaedic Association scale: establishing criteria for mild, moderate and severe impairment in patients with degenerative cervical myelopathy. Eur Spine J 26:78–84. https://doi.org/10.1007/s00586-016-4660-8
17. de Dios E, Laesser M, Björkman-Burtscher IM, Lindhagen L, MacDowall A (2022) Improvement rates, adverse events and predictors of clinical outcome following surgery for degenerative cervical myelopathy. Eur Spine J 31:3433–3442. https://doi.org/10.1007/s00586-022-07359-9
18. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174
19. Tetreault L, Nouri A, Kopjar B, Côté P, Fehlings MG (2015) The minimum clinically important difference of the modified Japanese Orthopaedic Association scale in patients with degenerative cervical myelopathy. Spine 40:1653–1659. https://doi.org/10.1097/brs.0000000000001127
20. Mjåset C, Zwart JA, Kolstad F, Solberg T, Grotle M (2022) Clinical improvement after surgery for degenerative cervical myelopathy; A comparison of patient-reported outcome measures during 12-month follow-up. PLoS ONE 17:e0264954. https://doi.org/10.1371/journal.pone.0264954
21. Gulati S, Vangen-Lønne V, Nygaard ØP, Gulati AM, Hammer TA, Johansen TO, Peul WC, Salvesen ØO, Solberg TK (2021) Surgery for degenerative cervical myelopathy: a nationwide registry-based observational study with patient-reported outcomes. Neurosurgery 89:704–711. https://doi.org/10.1093/neuros/nyab259
22. de Dios E, Heary RF, Lindhagen L, MacDowall A (2022) Laminectomy alone versus laminectomy with fusion for degenerative cervical myelopathy: a long-term study of a national cohort. Eur Spine J 31:334–345. https://doi.org/10.1007/s00586-021-07067-w
23. Tetreault LA, Karpova A, Fehlings MG (2015) Predictors of outcome in patients with degenerative cervical spondylotic myelopathy undergoing surgical treatment: results of a systematic review. Eur Spine J 24(Suppl 2):236–251. https://doi.org/10.1007/s00586-013-2658-z
24. Gerdhem L, Charalampidis A, Gerdhem P (2023) Patient-reported Data as Predictors of surgical outcome in patients with degenerative cervical myelopathy: analysis of a national multicenter dataset. Spine 48:113–119. https://doi.org/10.1097/brs.0000000000004469
25. Karpova A, Arun R, Davis AM, Kulkarni AV, Massicotte EM, Mikulis DJ, Lubina ZI, Fehlings MG (2013) Predictors of surgical outcome in cervical spondylotic myelopathy. Spine 38:392–400. https://doi.org/10.1097/BRS.0b013e3182715bc3
26. Boerger T, Alsouhibani A, Mowforth O, Hamilton J, Lalkhen A, Davies BM, Kotter MRN (2022) Moving beyond the neck and arm: the pain experience of people with degenerative cervical myelopathy who have pain. Global Spine J 12:1434–1442. https://doi.org/10.1177/21925682220986143
27. Hesni S, Baxter D, Saifuddin A (2023) The imaging of cervical spondylotic myeloradiculopathy. Skeletal Radiol. https://doi.org/10.1007/s00256-023-04329-0
28. Jiang SD, Jiang LS, Dai LY (2011) Degenerative cervical spondylolisthesis: a systematic review. Int Orthop 35:869–875. https://doi.org/10.1007/s00264-010-1203-5
29. Holmberg ST, Gulati AM, Johansen TO, Salvesen ØO, Lønne VV, Solberg TK, Tronvik EA, Nygaard ØP, Gulati S (2022) Surgery for degenerative cervical myelopathy in patients with rheumatoid arthritis and ankylosing spondylitis: a nationwide registry-based study with patient-reported outcomes. Acta Neurochir (Wien) 164:3165–3171. https://doi.org/10.1007/s00701-022-05382-9
30. Nurick S (1972) The pathogenesis of the spinal cord disorder associated with cervical spondylosis. Brain 95:87–100. https://doi.org/10.1093/brain/95.1.87