Gait assessment tools for degenerative cervical myelopathy: a systematic review

Wen Jie Choy[1,2], Lingxiao Chen[3], Camila Quel De Oliveira[4], Arianne P. Verhagen[4], Omprakash Damodaran[2], David B. Anderson[5,6]

[1]NeuroSpine Surgery Research Group (NSURG), Sydney, Australia; [2]Department of Neurosurgery, Nepean Hospital, Sydney, Australia; [3]Faculty of Medicine and Health, The Back Pain Research Team, Sydney Musculoskeletal Health, The Kolling Institute, School of Health Sciences, University of Sydney, Sydney, Australia; [4]Discipline of Physiotherapy, Graduate School of Health, University of Technology Sydney, Sydney, Australia; [5]Faculty of Medicine and Health, School of Health Sciences, University of Sydney, New South Wales, Australia; [6]Sydney Spine Institute, Burwood, Sydney, Australia

Contributions: (I) Conception and design: DB Anderson, O Damodaran; (II) Administrative support: DB Anderson, WJ Choy; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: DB Anderson, WJ Choy, L Chen; (V) Data analysis and interpretation: DB Anderson, WJ Choy, L Chen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: David B. Anderson. Faculty of Medicine and Health, School of Health Sciences, University of Sydney, New South Wales, Australia. Email: david.anderson1@sydney.edu.au.

Background: Degenerative cervical myelopathy (DCM) is a common progressive neurological disorder which may affect one's activities of daily living or even result in paraplegia/tetraplegia if left untreated. Currently, there is lack of consensus of the gait assessment tools for DCM. This systematic review aims to (I) provide an appraisal of the psychometric properties of the available gait assessment tools for DCM, (II) to assess their methodological quality according to The Consensus-based Standards for the selection of health Measurement COSMIN risk of bias checklist and (III) to assess each measurement property result against externally validated criteria.

Methods: Six electronic full-text databases [PubMed (via NLM® database], Medline (via OvidSP), CINAHL (via Ebsco), EMBASE (via Ovid), PsycINFO (via CSA) and Web of Science (via Thomson Reuters)] were systematically searched from inception to June 2020. The methodological quality of each study was analysed using the COSMIN risk of bias checklist. The measurement property result and methodological quality of each study were evaluated.

Results: Twenty studies were included from 3,339 citations retrieved. Twelve assessment tools for assessing gait in DCM were identified. According to COSMIN criteria, only five studies (25%) included in this review were found to have "very good" methodological quality. For construct validity, five tools had "sufficient" quality. For reliability, two assessment tools [the Total modified Japanese Orthopaedic Association Score (Italian Translation) (mJOA-ITTotal) and the modified Japanese Orthopaedic Association (Italian Translation) Motor dysfunction of the Lower Extremity (mJOA-ITMDLE)] were rated as "sufficient" for interobserver reliability while six assessment tools (the 10 second step test (10 sec ST), 30 minute walk test (30MWT), foot tapping test, mJOA-ITTotal, mJOA-ITMDLE and the Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire) were rated as "sufficient" for test-retest reliability. The JOA (6 scores) received a "sufficient" rating for internal consistency. No assessment was available for responsiveness, as only the effect size was available.

Discussion: Based upon current evidence, the mJOA in combination with an objective functional test (i.e., 30MWT) is recommended for clinicians assessing gait in DCM, although this may change with an increase in the number of studies completed. Given the importance of assessment tools possessing adequate measurement properties, a focus on studies in this area is warranted.

Keywords: Measurement properties; cervical myelopathy; gait assessment; spinal stenosis; cervical spondylosis

Introduction

Degenerative cervical myelopathy (DCM) is a common progressive neurological disorder in adults over 65 years of age with a prevalence of surgical treatment for DCM estimated to be 1.6 in 100,000 (1,2). Patients suffering from DCM can present with varied symptoms, including: sensorimotor deficits, reduced manual dexterity and balance, lower extremity spasticity, neuropathic pain, and bowel/bladder dysfunction (3). Reduced walking ability is also a common sign of DCM, occurring even in the early stages of the condition (4). In DCM the rubrospinal, reticulospinal and vestibulospinal tracts within the spinal cord are thought to be affected (5). As these tracts are mainly involved in the control of postural stability and gait; compression can lead to impaired walking ability (6,7). If left untreated, patients with DCM may suffer from progressive neurological paraplegia/tetraplegia, functional decline, and significantly reduced quality of life (8).

The diagnosis of DCM requires careful history and clinical examination to identify the signs of myelopathy and their correlation with radiological findings. To date, the detection of DCM and monitoring for clinical progression relies upon a clinician's judgement and experience (8). Management of DCM also depends on clinical judgement, and the perceived severity of symptoms (8). In most cases, non-operative management is provided for patients with mild DCM [i.e., ≥15/17 Modified Japanese Orthopaedic (mJOA) score (9)], however, some spinal surgeons may operate in the earlier stages of the disease, based upon concern for a rapid progression of symptoms (10). The aim of surgical decompression is to prevent progression of symptoms, as pre-surgical damage to the spinal cord can be irreversible (11). Asymptomatic patients on the other hand, with signal changes on an MRI but no to minumum DCM symptoms, can be reviewed regularly and monitored for any functional decline associated with myelopathy. Assessment of DCM and its impairments are therefore, key for decision-making and appropriate treatment planning.

Currently multiple gait/walking assessment tools are available to grade mobility of patients suffering from DCM. The JOA (Japanese Orthopaedic Association), mJOA and Nurick Grade are common tools used by clinicians to assess function in people with DCM (12). Although, these tools provide an overall numerical scoring based on individual questionnaire responses, they do not provide specific information on parameters of gait (i.e., velocity,

cadence, step length, range of motion of joints) (13). Given gait analysis was found in one study to be associated with increased cord-signal intensity changes on MRI, focused assessments should be considered (14). However, accurate gait analysis has required specialised equipment and training, which is not feasible in most clinical settings. For example, accurate gait and postural analysis can require expensive 3-dimensional cameras, multiple sensors placed across the body, and a specialised analyst, all of which are barriers in measuring DCM patients who attend to a clinic (15). Before new or existing gait assessment tools can be used by clinicians to assess DCM, the assessment tools need to be assessed for their measurement properties.

Understanding a gait assessment tool's measurement properties (i.e., validity, reliability, responsiveness) allows the assessor to be confident in the results obtained (16). Validity refers to the degree to which an assessment instrument measures the construct its purports to measure (16). Reliability refers to the degree to which the measurement is free from measurement errors and therefore, inform if a tool is reproducible between clinicians as well as between trials. Responsiveness refers to the ability of a tool to detect changes over time in the construct measured (17). The Consensus-based Standards for the selection of health Measurement (COSMIN) has developed criteria on how the measurement properties of tools should be evaluated to determine the confidence in both the results and the measurement instruments (18). The COSMIN tool can help guide clinicians' decision in choosing an appropriate assessment tool. The aim of this review was to systematically review the literature to identify the measurement properties of all existing gait assessment tools for DCM using (I) the COSMIN criteria for assessing methodology and (II) externally validated criteria for assessing measurement properties results (18). We present the following article in accordance with the PRISMA reporting checklist (available at https://jss.amegroups.com/article/view/10.21037/jss-21-109/rc).

Methods

The protocol for this review was prospectively registered with the International Prospective Register of Systematic Reviews (PROSPERO, Identifier CRD42020198208) and published on the 11th of July 2020. This review followed The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) (19).

Study selection criteria

Studies were included in this systematic review if they met the following criteria: (I) study design: randomised controlled trials and observational studies (e.g., cohort, cross-sectional, case-control studies) published in English were included. Conference proceedings and trials with no full text available were excluded; (II) participants: studies whose participants were adults (age over 40 years) with a diagnosis of DCM of any duration were included. DCM was defined as compression of the spinal cord at the level of the cervical spine, caused by degenerative changes (i.e., osteophytes, disc degeneration). (III) outcome measures: studies that assessed at least one measurement property of a walking outcome measure for patients with DCM were included. Outcome measures could be either assessed by a clinician or self-reported.

Search strategy

The search strategy was developed and informed by previous systematic reviews of walking tests for spinal conditions and measurement properties, and in consultation with a librarian (12,20,21). Permutations of the following keywords were used for this search: Spinal cord compression, Atraumatic spinal cord injury, Cervical myelopathy, Cervical canal stenosis, Central cord syndrome, Gait, Walking, Assessment, Measurement, Reproducibility of results. We searched six electronic databases: PubMed (via NLM® database), Medline (via OvidSP), CINAHL (via Ebsco), EMBASE (via Ovid), PsycINFO (via CSA) and Web of Science (via Thomson Reuters) from inception to June 2020.

Study titles and abstracts were screened against the inclusion and exclusion criteria by two reviewers independently. Full texts of relevant studies were retrieved, and their reference lists were screened to identify further studies of relevance. Disagreement between reviewers were resolved by consensus, with a third reviewer consulted if agreement could not be reached.

Data extraction

We used a standardised data extraction form for each included article, the following information was extracted: study characteristics (year, authors, study type, sample size), participants (age, gender), outcomes, and data completeness (missing data). Measurement properties, defined according to the COSMIN (18), and their results were also extracted, including validity (content, construct, criterion, cross-cultural), reliability (test re-test, intrarater, interrater), responsiveness, floor and ceiling effects.

Methodological quality of each study

The COSMIN Risk of Bias Checklist (box 1., Patient-Reported Outcome Measures (PROMs) development) was completed by two independent assessors, to determine the methodology of individual studies. Each study's methodology were rates as "very good", "adequate", "doubtful" or "inadequate". The "worst score counts" principle was used in this analysis, as per COSMIN guidelines (22).

Evaluation of the measurement property

Each measurement property was assessed using adapted criteria from Terwee *et al.* (18). For each measurement property, a score of either sufficient (positive), insufficient (negative), or indeterminate was awarded by two separate authors during data extraction, based upon the criteria. A third reviewer resolved any disagreements that occurred between the first two reviewers.

Data synthesis

The extracted data from the included studies were collected, presented in tables, and summarised in the manuscript. Data was presented on the assessment tools assessed in each study, the methods of diagnosis, and the measurement properties of each assessment tool (18).

Results

Search results

A total of 3,339 citations were retrieved across six databases and manual search. After duplicate removal and screening, 20 studies (5,23-41) were included (*Figure 1*). Assessment tools evaluated in the included studies were physical examination of functional impairment (n=7), clinician-administered (n=3) and self-administered patient-reported outcome measures (n=2) (*Table 1*). Sixteen studies analysed construct validity, through correlating the walking assessment tools with other gait assessment tools (*Table 2*). Nine studies analysed either inter-observer or test-retest
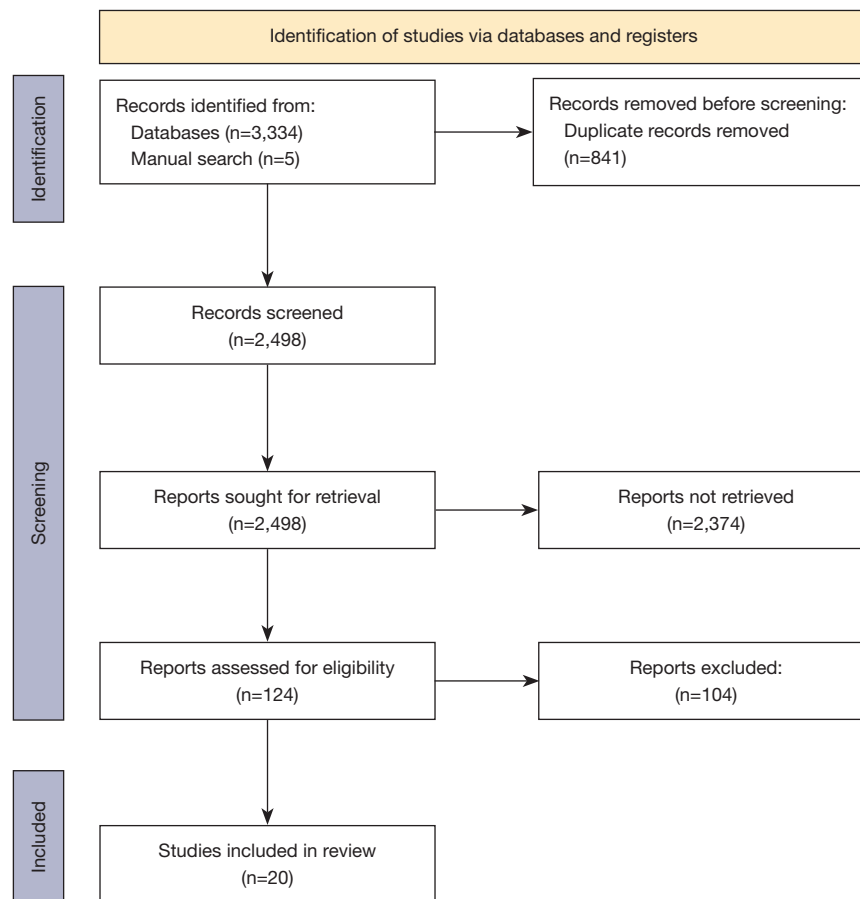
Figure 1 Flow chart outlining the process in this systematic review. Based upon the PRISMA preferred reporting for Systematic reviews.

reliability (*Table 3*). Internal consistency was measured in three studies, using Cronbach's alpha (*Table 4*). Three studies measured responsiveness through effect size (*Table 5*).

Our search revealed twelve different types of gait assessments, for which *Table 1* gives a comprehensive overview of each assessment tool. The most frequently reported assessment tools were the JOA score (n=4) with a total of 253 participants and the mJOA score (n=4) with 547 participants across studies. Other assessment tools were the 30-metre walk test (30MWT) (n=3) with a total of 743 participants, the 10-second step test (10 sec ST) (n=2, 1,468 participants), and the Nurick scale (n=2, 380 participants).

Methodological quality of studies

According to the COSMIN risk of bias tool, five studies were rated as "very good" (23,29-32), seven were rated as "adequate" (24,26,27,34,37,38,40), three were rated

"doubtful" (5,28,41) and five were rated as "inadequate" (25,33,35,36,39). These results are presented in the last column of *Tables 2-5*. Studies that employed the 30MWT (n=2) (31,32), 10 sec ST (n=2) (23,31), mJOA score (n=1) (29) and the JOACMEQ (Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire) (n=1) (30) were rated as "very good".

Construct validity

Of the ten assessment tools, eight studies were rated as "sufficient" in terms of correlation with other assessment tools (23-28,37,39). Both studies correlating the foot tapping test (FTT) to other assessment tools had a mixed rating of "sufficient" and "insufficient" (34,38). Mixed ratings of "sufficient" and "insufficient" were also given to two separate studies that correlated the enhanced gait variability index and the mJOA score to other assessment tools (5,29). One study that correlated the 30MWT to the Nurick

Table 1 Study characteristics of walking tools

Name of assessment	Studies (n=20)	Brief description of assessment	Methods of diagnosis	Number of cumulative participants	
				DCM	Controls
Physical examination of functional impairment					
10-second step test	Nakashima [2011] (31); Yukawa [2009] (23)	The patient is asked to perform stepping movements on the spot for 10 seconds with no external support, at maximal speed by lifting hips and knees to 90 degrees of flexion. If the patient is at risk of falling, they are asked to perform the test next to a handrail. The supervisor is positioned next to the patient. The number of steps is recorded and a decrease in steps per 10 seconds reflects a greater level of disability	Nakashima: neurological exam with MRI and CT myelogram findings; Yukawa: physical neurological exam with MRI and CT myelogram findings	264	1,204
30MWT	Bohm [2017] (32); Singh [1999] (28); Nakashima [2011] (31)	The 30MWT is conducted by recording the time it takes for a patient to walk 30 metres. Patients are instructed to walk "as quickly as possible", using any walking aid they might normally use. An increase in 30m walk time reflects a greater level of disability	Bohm: clinical sign(s) of myelopathy and MRI evidence of cord compression; Singh: confirmed by MRI; Nakashima: neurological exam with MRI and CT myelogram findings	743	0
3D Gait analysis	McDermott [2010] (33); Maezawa [2001] (39)	Gait analysis is conducted by having a patient walk along a specified track, whilst having their movements analysed using a camera motion analysis system. Reflective markers are placed on the patient's anatomical landmarks (usually shoulders, iliac crests, greater trochanters, lateral side of the femoral condyles, lateral malleoli, lateral calcaneus, head of the 5th metatarsals), which allow movement of the limbs to be picked up by the cameras. The camera system, track length and number of repetitions vary between studies	McDermott: clinical sign(s) of myelopathy and MRI evidence of cord compression; Maezawa: not specified	36	62
Enhanced Gait Variability Index	Kalsi-Ryan [2020] (5)	The eGVI considers step time and length, stance time, single-stance time and stride velocity to quantify the gait variability of an individual. These spatiotemporal gait parameters are measured with a gait analysis system	Kalsi-Ryan: clinical sign(s) of myelopathy and MRI evidence of cord compression	153	0
Foot tapping test	Numasawa [2012] (34); Enoki [2019] (38)	The FTT is completed with the patient sitting comfortably in a chair with hips and knees at 90 degrees of flexion. While keeping their heel on the floor, the patient is asked to tap their toes as fast as possible for ten seconds. After a few practice trials, each leg is assessed individually with an examiner counting the number of taps achieved. The value of each leg is combined for an overall mean (unless making side-to-side comparisons). A lower score reflects a greater level of lower limb disability	Numasawa: clinical sign(s) with MRI or CT myelogram findings; Enoki: not specified	329	792
Triangle step test	Mihara [2010] (40)	A triangular board with marked borders of 30cm is placed in front of a patient sitting in a chair with both the hip and knee joint flexed to 90 degrees and feet flat on the floor. The patient is then instructed to lift their foot of the floor and tap each mark after another as quickly as they can in 10 seconds. The number of steps is recorded. Each side should be completed individually	Mihara: not specified	270	60
Wearable EMG	Malone [2011] (36)	Patients are equipped with pre-amplified surface electrodes placed on common muscle groups bilaterally including the rectus femoris, biceps femoris, tibialis anterior and medial head of gastrocnemius. Patients are also equipped with a reference electrode on C7. The patient then undergoes a three-dimensional gait analysis using a motion analysis system	Malone: clinical sign(s) of myelopathy and MRI evidence of cord compression	12	0

Table 1 (*continued*)

Table 1 (*continued*)

Name of assessment	Studies (n=20)	Brief description of assessment	Methods of diagnosis	Number of cumulative participants DCM	Controls
Clinician-administered patient-reported outcome measures					
JOA Score	Yonenobu [2001] (35); Kato [2015] (25); Zheng [2016] (41); Singh [2001] (37)	The JOA score consists of six domains including motor dysfunction in the upper extremities, motor dysfunction in the lower extremities, sensory function in the upper and lower extremities and in the trunk and bladder function. The combined score for each domain totals from 0 to 17 with a lower score indicating greater severity of disease. The subsection "Motor dysfunction of the lower extremities" grades walking ability from "−1" being unable to walk to "5" equating to normal function	Yonenobu: not specified; Kato: not specified; Zheng: clinical and radiological evidence of cervical cord myelopathy; Singh: confirmed by MRI	253	0
mJOA Score	Kato [2015] (25); Whitmore [2013] (26); Kopjar [2015] (29); Longo [2016] (27)	The mJOA scoring system assesses the neurological function of patients presenting with degenerative cervical myelopathy. This consists of four domains differing to that of the JOA score by not including sensory function in the trunk and lower extremities. The mJOA scoring system assesses motor function in the upper and lower extremities, sensory function in the upper extremities and bladder function. The combined score from each domain totals from 0 to 18 with a lower score indicating greater severity of disease. The subsection "Motor dysfunction of the lower extremities", grades walking ability from "−7" (no dysfunction) to "0" (complete loss of motor and sensory function)	Kato: not specified; Whitmore: clinical signs and evidence of cervical cord compression; Kopjar: clinical sign(s) and MRI evidence of cord compression; Longo: clinical, electrophysiological and MRI findings	547	0
Nurick Grade	Kopjar [2015] (29); Whitmore [2013] (26)	This six-grade scale uses a combination of the patient's mobility status with or without an aid and its impact on their ability to participate in full time employment or activities of daily living to measure the severity of cervical myelopathy. A score of 0 equates to "signs or symptoms of root involvement but no evidence of spinal cord disease" and a score of 5 equates to being "chair bound or bedridden"	Kopjar: clinical sign(s) and MRI evidence of cord compression; Whitmore: clinical signs and evidence of cervical cord compression	380	0
Self-administered patient-reported outcome measures					
JOACMEQ	Fukui [2007] (30)	This is a seventy-seven item self-reported questionnaire used to measure function in Cervical Myelopathy patients. The questionnaire asks for information regarding the patient's motor function, continence, participation in daily activities, and emotional state. A higher score indicates higher severity of symptoms, thus a greater impact on the patient's quality of life. There are two items related to walking for which we were able to collect measurement properties including: "QOL04. Do you have difficulty walking more than 15 minutes?" and "C04. Can you walk on a flat surface?"	Fukui: not specified	201	0
PROMIS-Physical Function	Owen [2018] (24)	The PROMIS-PF is a self-reported questionnaire with one hundred twenty items based on four subdomains - activities of daily living, mobility of lower extremity function, back and neck and upper extremity function. Using computerised adaptive testing algorithms, the questions are adjusted based on the patient's prior answers. Each answer is correlated to a score of 1 to 5. Higher scores reflect better functioning. There are eight questions assessing walking in this tool	Owen: not specified	60	0

30MWT, 30-meter walk test; CM, cervical myelopathy; FTT, Foot tapping test; 3D, three dimensional; eGVI, Enhanced Gait Variability Index; EMG, Electromyography; mJOA score, modified Japanese Orthopaedic score; JOA score, Japanese Orthopaedic score; mJOA-IT, modified Japanese Orthopaedic score (Italian translation); JOACMEQ, Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire; PROMIS-PF, Patient-Reported Outcomes Measurement Information System-Physical Function.

Table 2 Construct Validity measured by correlations between assessment tools

Assessment tool	Studies first author (year)	Correlation with other assessment tools	Measurement property quality (using 0.5)	Methodological quality (COSMIN)
10 sec ST	Yukawa [2009] (23)	JOALEMF score (r=0.84)	+	Very good
	Nakashima [2011] (31)	JOA_{total} score [r^2=0.30 (pre-operative), 0.17 (post-operative]; JOAMEQ-L [r^2=0.40(pre-operative), 0.44 (post-operative)]	NA	Very good
30MWT	Bohm [2017] (32)	Nurick grade (r=0.49); $mJOA_{Total}$ score (r=−0.44)	−	Very good
	Nakashima [2011] (31)	JOA_{total} score [r^2=0.19 (pre-operative), 0.18 (post-operative)]; JOAMEQ-L [r^2=0.32 (pre-operative), 0.31 (post-operative)]	NA	Very good
	Singh [1999] (28)	Pre-operative, post-operative Nurick grade (r=0.61, 0.69)	+	Doubtful
Gait analysis (eGVI)	Kalsi-ryan [2020] (5)	$mJOA_{MDLE}$ score (r=0.57); Velocity & $mJOA_{MIDLE}$ (r=0.46)	+; −	Doubtful
Extension of knee in stance phase	Maezawa [2001] (39)	JOA_{Total} score (r=0.54)	+	Inadequate
Triangle ST	Mihara [2010] (40)	Pre-operative Nurick Score (significant correlation†); $mJOA_{MDLE}$ Score (significant correlation†)	NA	Adequate
FTT	Numasawa [2012] (34)	JOA_{LEMF} Score (r=0.70); JOA_{total} Score (r=0.66); postoperative gain & JOA_{LEMF} Score (r=0.43)	+; +; −	Adequate
	Enoki [2019] (38)	Nurick grade (r=−0.57); JOA_{LEMF} score (r=0.52); $30MWT_{walking\ time}$ (r=−0.51); $30MW_{number\ of\ steps}$ (r=−0.49)	+; +; +; −	Adequate
PROMIS-PF	Owen [2018] (24)	Baseline, 6-month follow-up; $mJOA_{Total}$ Score (r=0.61, 0.72)	+	Adequate
JOA Score	Kato [2015] (25)	JOA_{LEMF} & $mJOA_{MIDLE}$ (r=0.93)	+	Inadequate
	Zheng [2016] (41)	Gait parameters: double support duration (ms) (r^2=0.25), step duration (ms) (r^2=0.21)	NA	Doubtful
	Singh [2001] (37)	Nurick grade: pre-operative (r=0.59); post-operative (r=0.51)	+	Adequate
mJOA-IT Score	Longo [2016] (27)	$mJOA\text{-}IT_{Total}$ score correlated with Nurick (r=−0.62), $mJOA\text{-}IT_{MIDLE}$ score & Nurick grade (r=−0.65)	+	Adequate
mJOA Score	Kopjar [2015] (29)	$mJOA_{Total}$ Score & Nurick grade (r=−0.63); 30MWT (r=−0.38); $mJOA_{MIDLE}$ Score & Nurick grade (r=−0.68); 30MWT (r=−0.43)	+; −; +; −	Very good
	Whitemore [2013] (26)	Nurick grade (r=−0.73)	+	Adequate

Measurement property quality was rated based on Terwee et al.'s quality criteria assessment. Correlation (r) ≥0.50 was rated as "+" whereas correlation (r) <0.5 was rated as "−". "+" = sufficient; "−" = insufficient; "r" = Spearman's or Pearson's correlation coefficient; "r^{2}" = squared correlation coefficient from linear regression. †, correlation value not specified in article. NA, not applicable; 10 sec ST, 10 second Step test; JOALEMF, Japanese Orthopaedic Association Lower Extremity Motor function; JOA_{Total}, Total Japanese Orthopaedic Association; JOAMEQ-L, Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire Lower Limb Function 30MWT, 30-Meter Walk Test; $mJOA_{Total}$, total Modified Japanese Orthopaedic; eGVI, Enhanced Gait Variability Index; $mJOA_{MIDLE}$, modified Japanese Orthopaedic Association Motor dysfunction of the Lower Extremity; Triangle ST, Triangle Step test; FTT, Foot tapping test; PROMIS-PF, Patient Reported Outcome Measurement Information System-Physical Function; JOA, Japanese Orthopaedic association; ms, Millisecond; mJOA-IT, modified Japanese Orthopaedic Association (Italian Translation); $mJOA\text{-}IT_{Total}$, Total modified Japanese Orthopaedic Association Score (Italian Translation); $mJOA\text{-}IT_{MIDLE}$, modified Japanese Orthopaedic Association (Italian Translation) Motor dysfunction of the Lower Extremity.

Table 3 Interobserver and test-retest reliability

Assessment tool	Studies first author (year)	Sample size of CM population (n)	Interobserver reliability	Measurement property for interobserver reliability	Test-retest reliability	Measurement property quality for test-retest reliability	Methodological quality (COSMIN)
10 sec ST	Yukawa [2009] (23)	163	Not reported	NA	r=0.89	+	Very good
	Nakashima [2011] (31)	168	Not reported	NA	Pre-operative: r^2=−0.28 (SD 1.04); post-operative: r^2=−0.55 (SD 0.87)	NA	Very good
30MWT	Bohm [2017] (32)	757	Not reported	NA	r=0.97	+	Very good
	Nakashima [2011] (31)	168	Not reported	NA	Pre-operative: r^2=0.95 (SD 1.66); post-operative: r^2=0.89 (SD 1.33)	NA	Very good
Gait Parameters	McDermott [2010] (33)	12	Not reported	NA	Cadence (Steps. min): ICC =0.99; double support (s): ICC =0.99; foot off (%): ICC =0.97; opposite foot contact: ICC =0.61; single support(s): ICC =0.96; step length (m): ICC =0.99; step width (m): ICC =0.91; stride length (m): ICC =0.99; stride time (s): ICC =0.99	*_	Inadequate
FTT	Numasawa [2012] (34)	77	Not reported	NA	Right foot: r=0.93; left foot: r=0.90	+	Adequate
JOA	Yonenobu [2001] (35)	29	ICC =0.81	*_	ICC =0.83	*_	Inadequate
mJOA-IT_Total Score	Longo [2016] (27)	75	KI =0.80	+	r=0.91	+	Adequate
mJOA-IT_MDLE Score	Longo [2016] (27)	75	KI =0.73	+	r=0.93	+	Adequate
JOACMEQ	Fukui [2007] (30)	201	Not reported	NA	KI =0.74 (95% CI: 0.69–0.79); KI =0.73 (95% CI: 0.67–0.78)	+	Very good
Wearable EMG during gait	Malone [2011] (36)	12	Not reported	NA	Rectus Femoris: ICC =0.59; biceps femoris: ICC =0.56; tibialis anterior: ICC =0.81; gastrocnemius: ICC =0.55	*_	Inadequate

Measurement property quality was rated based on Terwee *et al.*'s quality criteria assessment. ICC/KI ≥0.70 or Spearman's/Pearson's r≥0.80 was rated as "+" whereas ICC/ KI <0.70 OR Spearman's/Pearson's r<0.80 was rated as "_". "_", "_" = insufficient; "+" = sufficient; "r" = use of Spearman's or Pearson's correlation coefficient; "r^2" = squared correlation coefficient from linear regression. *, use of ICC. NA, not applicable; COSMIN, Consensus-based Standards for the selection of health Measurement Instruments; 10 sec ST, 10 second step test; NA, not applicable; 30MWT, 30-meter walk test; ICC, Intraclass correlation coefficient; FTT, Foot tapping test; JOA, Japanese Orthopaedic Association; JOALEMF, Japanese Orthopaedic Association Lower Extremity Motor function; KI, weighted Kappa coefficient; mJOA-IT_Total, Total modified Japanese Orthopaedic Association Score (Italian Translation); mJOA-IT_MDLE, modified Japanese Orthopaedic Association (Italian Translation) Motor dysfunction of the Lower Extremity; JOACMEQ, Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire; EMG, electromyography.

Table 4 Internal consistency by Cronbach's Alpha

Assessment tool	Studies first author (year)	Cronbach's Alpha	Measurement property quality	Methodological quality (COSMIN)
mJOA	Kopjar [2015] (29)	Total score =0.63	–	Very good
mJOA	Longo [2016] (27)	Total score =0.6	–	Adequate
JOA	Singh [2001] (37)	6 scores: [0.72 (pre-operative), 0.73 (post-operative)]; 4 categories [0.66 (pre-operative), 0.65 [post-operative)]	+; –	Adequate

Measurement property quality was rated based on Terwee *et al.*'s quality criteria assessment. Cronbach's alpha ≥0.70 was rated as "+" whereas Cronbach's alpha <0.70 was rated as "–". "–" = insufficient; "+" = sufficient. COSMIN, Consensus-based Standards for the selection of health Measurement Instruments; mJOA, modified Japanese Orthopaedic Association; JOA, Japanese Orthopaedic Association; mJOA-IT, modified Japanese Orthopaedic Association (Italian Translation).

Table 5 Responsiveness represented through effect size

Measurement tool	Studies first author (year)	Responsiveness (Cohen's d)	Measurement property quality	Methodological quality (COSMIN)
30MWT	Bohm [2017] (32)	Test performed at baseline level score =0.30; for patients who had a baseline 30MWT time above the median value =0.45	NA	Very good
mJOA-IT score	Longo [2016] (27)	0.87	NA	Adequate
mJOA	Kopjar [2015] (29)	1	NA	Very good

NA, not applicable; COSMIN, Consensus-based Standards for the selection of health Measurement Instruments; 30MWT, 30 metre walk test; mJOA-IT, modified Japanese Orthopaedic Association (Italian Translation); mJOA, modified Japanese Orthopaedic Association.

grade and mJOA$_{Total}$ score was rated as "insufficient" (32). We were unable to rate three studies [two studies used linear regression for correlation with other assessment tools (31,41) and one study did not provide a correlation score (40)]. *Table 2* summarises the correlations between different assessment tools. The 10 sec ST test had a correlation (r=0.84) with the JOA$_{LEMF}$ (JOA Lower Extremity Motor Function) sub score in a group of 163 DCM patients from a single study (23). In a separate study, the 6-month follow-up PROMIS-PF (Patient-Reported Outcomes Measurement Information System-Physical Function) score adequately correlated (r=0.61, 0.72) with the mJOA$_{Total}$ score in a group of 60 DCM patients (24). The JOA$_{LEMF}$ and mJOA$_{MDLE}$ (mJOA Motor dysfunction of the lower extremity) scores adequately correlated (r=0.93) in a group of 92 participants in another study (25). Lastly, the mJOA$_{Total}$ score had an adequate correlation with the Nurick grade (r=–0.73) in another study with 103 DCM patients (26).

Notably, multiple assessments across three studies had adequate correlations (r=0.60–0.69) with the Nurick grade, including the 30MWT (preoperatively and postoperatively,

r=0.61, 0.69), the mJOA-IT (mJOA Italian Translation) (total and MDLE scores, r=–0.62, –0.65) and the mJOA Score (total and MDLE scores, r=–0.63, –0.68) (27-29).

Reliability

Nine studies assessed the reliability of various walking tools in DCM populations and are shown in *Table 3* (23,27,30-36). Two studies reported on interobserver reliability (27,35). The mJOA-IT$_{Total}$ and mJOA-IT$_{MDLE}$ were rated as having "sufficient" interobserver reliability a weighted kappa coefficient of 0.80 and 0.73 respectively (27). The JOA was also found to have adequate interobserver reliability, with a correlation of 0.81, but the methodological quality of the study was 'inadequate' (35).

Nine studies reported on test-retest reliability and intraobserver reliability (23,27,30-36). The JOACMEQ was rated "sufficient" (30). The JOA, gait parameters via three-dimensional gait analysis and wearable electromyography (EMG) all had adequate reliability, but the methodological quality of the studies was inadequate, according to

COSMIN, due to the small sample size of less than 50 participants (33,35,36). The 10 sec ST, 30MWT, FTT, mJOA-IT$_{Total}$ and mJOA-IT$_{MDLE}$ all reported adequate reliability, with a correlation coefficient ≥0.80 (23,27,32,34). One study that assessed the 30MWT had very good methodological quality, but presented an r^2, which although high (range, 0.89 to 0.95) (31), could not be assessed with the externally validated quality criteria employed in this review (18).

Internal consistency

Three studies reported on internal consistency and are shown in *Table 4*. The JOA score (37) received a "sufficient" rating when participants performed the six questions, with a Cronbach's alpha of 0.72 preoperatively and 0.73 postoperatively. The mJOA (29) and mJOA-IT (27) scores were both rated as "insufficient" due to a Cronbach's Alpha value <0.70.

Responsiveness

The three studies that reported on responsiveness measured it using effect sizes, which is not a method recognised by the validated quality criteria employed in this review (18). Thus, we could not analyse responsiveness as a measurement property based on these guidelines.

As outlined in *Table 5*, the mJOA-IT (27) and mJOA (29) recorded large effect sizes (0.87, 1 respectively), while the 30MWT demonstrated a small effect size of 0.26 (32).

Discussion

This review included twenty studies that assessed the measurement properties of walking tests for DCM. Twelve tools were evaluated to assess walking in people with DCM, with three (25.0%) being clinician-administered assessments, two (16.7%) being self-administered patient-reported outcomes and seven (58.3%) being physical examination of functional impairments. The most commonly assessed tests were JOA and mJOA scores across four studies each.

The clinician-administered outcome measures identified in our study included the commonly used JOA and mJOA scoring systems, and the Nurick grade (8). The main difference between the JOA and mJOA scores is that the mJOA does not measure sensory function (25). Each of these tools have a lower limb subscore for classifying a patient's walking ability. The reliability of the JOA score was inadequate due to the small sample sizes, while the reliability of the mJOA score has not yet been assessed. This leads to the lack of knowledge on the measurement errors and potential inaccuracy when monitoring changes overtime in the same patient and when measuring differences between patients. The JOA$_{total}$ score demonstrated adequate internal consistency across six points both pre- and post-operatively, while the mJOA's internal consistency was inadequate. The mJOA-IT score demonstrated adequate interobserver reliability. Despite the uncertainty of the measurement properties of these tools, they remain useful in classifying the severity of the patient's functional impairments.

Similar to the lower limb subscore of the mJOA and the mJOA scoring system, the Nurick grade measures ambulatory status. The Nurick grade correlated sufficiently with both the lower limb subscores of the mJOA and the overall mJOA scores (27,29). Thus, it has been argued that this tool may be redundant if it correlates sufficiently with the lower limb subscore of the mJOA (13). Despite this, its overall validity quality was downgraded due to a lack of pre-defined hypotheses in the study (27).

Self-administered PROMs are important in identifying an individual's perception of their own health, function and walking ability. Two PROMs, the Patient-Reported Outcomes Information System (PROMIS) Physical Function (PF) and the JOACMEQ were identified in our literature search. Both are used to assess self-reported functional ability in DCM and contain walking items (24,30). The PROMIS-PF measures self-reported function, with eight of the total one hundred twenty items being specific to walking ability. Validity was the only measurement property collected for this tool with a sufficient measurement property quality. However, individual items of the questionnaire were not assessed in the DCM population (24). Therefore, caution should be taken when employing solely the PROMIS-PF to assess walking function in patients with DCM. Compared to the PROMIS-PF, the JOACMEQ targets symptoms specific to DCM, making it more relevant to patients with DCM. The JOACMEQ is a seventy-seven-item questionnaire, however only twenty-four items had their measurement properties assessed and of these, only two items address walking ability (30). Although both items have adequate test-retest reliability, we were unable to conclude if these items provide an accurate measurement of a patient's walking ability.

The last group of tools identified in our search were quantitative physical measures. Enoki *et al.* (38), recommended the use of the 30MWT to assess walking in this population. However, based on Terwee *et al.*'s quality criteria assessment (18), the 30MWT had inadequate validity (31,32) and responsiveness (32). Furthermore, this test has been criticised by Nakashima, *et al.* for its use in patients with severe myelopathy, where an individual may not be able to mobilise for thirty metres (31). Other tools such as the 10 sec ST, FTT and triangle step test were all developed to address the concern that the 30MWT cannot be used in patients with severe myelopathy (23,31,34,38,40). Despite this criticism, this test is a cost-effective and feasible test for clinicians to use as it requires minimal training and equipment.

3D gait analysis was used to measure walking function in three studies. The reliability of gait parameters, assessed by McDermott *et al.*, seemed favourable, however was deemed inadequate due to the study's small sample size (33). Extension of the knee during the stance phase was a specific gait parameter measured by Maezawa *et al.*, with adequate validity (39). The enhanced Gait Variability Index (eGVI) assesses the quality of gait with accuracy, through measuring spatiotemporal parameters of gait (5). Siasios *et al.* found that patients with DCM exhibit reduced cadence and gait speed (4), thus gait analysis may be useful in identifying these impairments precisely. However, these tools are not always clinically available due to the cost of the equipment, training required, and time taken to perform a full gait analysis. Other tools such as the 10 sec ST, the FTT and the triangle step test can be alternatives to the 3D gait analysis. Although requiring less space and time to perform compared to 30MWT, limitation in study methodology and evidence hinder their use in clinical practice. Another quantitative tool found was wearable EMG, which assesses muscle activity during a patient's gait cycle (36). EMG on the tibialis anterior muscle demonstrated the highest test-retest reliability, but due to the small sample size, the quality of the reliability was downgraded. Due to costs, technical factors and complexity of EMG, barriers remain for its clinical adoption.

Strengths & limitations

There were some limitations present in this review. Primarily, the review only included studies that calculated a measurement property result for a gait assessment, impacting the generalisability of the results to these study types. Due to the study design being a systematic review, it would not be feasible to review every trial conducted on DCM to calculate a measurement property. Another limitation was the lack of studies available that had assessed a gait assessment tool for DCM, limiting the ability to make a recommendation on the most rigorous gait assessment tool for DCM.

A strength of this review was that it included the validated COSMIN risk of bias tool, and established criteria for assessing the measurement property result. Both the risk of bias tool and measurement property result criteria have also been used in previous reviews of the measurement properties for gait (16).

Clinical implications and future directions

Only five studies included in this review were found to have 'very good' methodological quality according to the COSMIN tool (23,29-32). We identified that the main reasons for lower methodological quality was a lack of pre-defined hypotheses, small sample sizes, and inappropriate analysis e.g., utilising spearman's correlation instead of the ICC to define reliability or using effect size rather than the area under the curve (AUC) to define responsiveness.

Until an increased number of high-quality studies have been completed, we recommend that clinicians should utilise a combined approach of the mJOA score with another objective test, such as the 30MWT, as they are both cost-effective and easy to administer. Although the measurement properties of the 30MWT were just below an adequate level for construct validity, it had very good evidence for adequate reliability. Furthermore, the 30MWT provides a functional, objective walking score, which is applicable to walking distance and provides a functional result. We additionally recognise the emerging 10 second ST, which demonstrates practicality, especially for patients with more severe DCM. This test should be further researched to determine its benefit in clinical practice. Finally, research on objective assessment tools, such as wearable and body tracking technologies, that are capable of tracking kinematics of gait and posture should be a priority. Gait analytical tools are likely to provide high levels of validity, reliability, and responsiveness in the near future, replacing self-reported and clinician assessed gait tools (42-44).

Conclusions

Twelve assessment tools of gait in DCM were identified.

The mJOA in combination with an objective functional test (i.e., 30MWT) is the recommended tool for clinicians assessing gait in DCM, although this may change with an increase in the number of studies completed.

Acknowledgments

Footnote

Provenance and Peer Review: This article was commissioned by Guest Editors (Ralph J. Mobbs, Pragadesh Natarajan and R. Dineth Fonseka) for the series "Objective Monitoring and Wearable Technologies including Sensor-Based Accelerometers and Mobile Health Applications for the Spine Patient" published in *Journal of Spine Surgery*. The article has undergone external peer review.

Reporting Checklist: The authors have completed the PRISMA reporting checklist. Available at https://jss.amegroups.com/article/view/10.21037/jss-21-109/rc

Peer Review File: Available at https://jss.amegroups.com/article/view/10.21037/jss-21-109/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jss.amegroups.com/article/view/10.21037/jss-21-109/coif). The series "Objective Monitoring and Wearable Technologies including Sensor-Based Accelerometers and Mobile Health Applications for the Spine Patient" was commissioned by the editorial office without any funding or sponsorship. WJC serves as an Assistant Managing Editor of *Journal of Spine Surgery*. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Boogaarts HD, Bartels RH. Prevalence of cervical spondylotic myelopathy. Eur Spine J 2015;24 Suppl 2:139-41.
2. Nouri A, Tetreault L, Singh A, et al. Degenerative Cervical Myelopathy: Epidemiology, Genetics, and Pathogenesis. Spine (Phila Pa 1976) 2015;40:E675-93.
3. Tracy JA, Bartleson JD. Cervical spondylotic myelopathy. Neurologist 2010;16:176-87.
4. Siasios ID, Spanos SL, Kanellopoulos AK, et al. The Role of Gait Analysis in the Evaluation of Patients with Cervical Myelopathy: A Literature Review Study. World Neurosurg 2017;101:275-82.
5. Kalsi-Ryan S, Rienmueller AC, Riehm L, et al. Quantitative Assessment of Gait Characteristics in Degenerative Cervical Myelopathy: A Prospective Clinical Study. J Clin Med 2020;9:752.
6. Drew T, Prentice S, Schepens B. Cortical and brainstem control of locomotion. Prog Brain Res 2004;143:251-61.
7. Hartman CJ, Hoh DJ. Pathobiology of Cervical Radiculopathy and Myelopathy. In: Degenerative Cervical Myelopathy and Radiculopathy. Springer, 2019:53-65.
8. Davies BM, Mowforth OD, Smith EK, et al. Degenerative cervical myelopathy. BMJ 2018;360:k186.
9. 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;26:78-84.
10. Bakhsheshian J, Mehta VA, Liu JC. Current Diagnosis and Management of Cervical Spondylotic Myelopathy. Global Spine J 2017;7:572-86.
11. Ghogawala Z, Benzel EC, Riew KD, et al. Surgery vs Conservative Care for Cervical Spondylotic Myelopathy: Surgery Is Appropriate for Progressive Myelopathy. Neurosurgery 2015;62 Suppl 1:56-61.
12. Singh A, Tetreault L, Casey A, et al. A summary of assessment tools for patients suffering from cervical

spondylotic myelopathy: a systematic review on validity, reliability and responsiveness. Eur Spine J 2015;24 Suppl 2:209-28.

13. Revanappa KK, Rajshekhar V. Comparison of Nurick grading system and modified Japanese Orthopaedic Association scoring system in evaluation of patients with cervical spondylotic myelopathy. Eur Spine J 2011;20:1545-51.

14. Kim CR, Yoo JY, Lee SH, et al. Gait analysis for evaluating the relationship between increased signal intensity on t2-weighted magnetic resonance imaging and gait function in cervical spondylotic myelopathy. Arch Phys Med Rehabil 2010;91:1587-92.

15. Cappozzo A, Della Croce U, Leardini A, et al. Human movement analysis using stereophotogrammetry. Part 1: theoretical background. Gait Posture 2005;21:186-96.

16. Anderson DB, Mathieson S, Eyles J, et al. Measurement properties of walking outcome measures for neurogenic claudication: a systematic review and meta analysis. Spine J 2019;19:1378-96.

17. Mokkink LB, Terwee CB, Patrick DL, et al. COSMIN checklist manual. Amsterdam: University Medical Center, 2012.

18. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34-42.

19. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.

20. Stienen MN, Ho AL, Staartjes VE, et al. Objective measures of functional impairment for degenerative diseases of the lumbar spine: a systematic review of the literature. Spine J 2019;19:1276-93.

21. Terwee CB, Jansma EP, Riphagen II, et al. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res 2009;18:1115-23.

22. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1147-57.

23. Yukawa Y, Kato F, Ito K, et al. "Ten second step test" as a new quantifiable parameter of cervical myelopathy. Spine (Phila Pa 1976) 2009;34:82-6.

24. Owen RJ, Zebala LP, Peters C, et al. PROMIS Physical Function Correlation With NDI and mJOA in the Surgical Cervical Myelopathy Patient Population. Spine (Phila Pa 1976) 2018;43:550-5.

25. 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:e0123022.

26. Whitmore RG, Ghogawala Z, Petrov D, et al. Functional outcome instruments used for cervical spondylotic myelopathy: interscale correlation and prediction of preference-based quality of life. Spine J 2013;13:902-7.

27. Longo UG, Berton A, Denaro L, et al. Development of the Italian version of the modified Japanese orthopaedic association score (mJOA-IT): cross-cultural adaptation, reliability, validity and responsiveness. Eur Spine J 2016;25:2952-7.

28. Singh A, Crockard HA. Quantitative assessment of cervical spondylotic myelopathy by a simple walking test. Lancet 1999;354:370-3.

29. Kopjar B, Tetreault L, Kalsi-Ryan S, et al. Psychometric properties of the modified Japanese Orthopaedic Association scale in patients with cervical spondylotic myelopathy. Spine (Phila Pa 1976) 2015;40:E23-8.

30. Fukui M, Chiba K, Kawakami M, et al. Japanese Orthopaedic Association Cervical Myelopathy Evaluation Questionnaire: part 3. Determination of reliability. J Orthop Sci 2007;12:321-6.

31. Nakashima H, Yukawa Y, Ito K, et al. Validity of the 10-s step test: prospective study comparing it with the 10-s grip and release test and the 30-m walking test. Eur Spine J 2011;20:1318-22.

32. Bohm PE, Fehlings MG, Kopjar B, et al. Psychometric properties of the 30-m walking test in patients with degenerative cervical myelopathy: results from two prospective multicenter cohort studies. Spine J 2017;17:211-7.

33. McDermott A, Bolger C, Keating L, et al. Reliability of three-dimensional gait analysis in cervical spondylotic myelopathy. Gait Posture 2010;32:552-8.

34. Numasawa T, Ono A, Wada K, et al. Simple foot tapping test as a quantitative objective assessment of cervical myelopathy. Spine (Phila Pa 1976) 2012;37:108-13.

35. Yonenobu K, Abumi K, Nagata K, et al. Interobserver and intraobserver reliability of the japanese orthopaedic association scoring system for evaluation of cervical compression myelopathy. Spine (Phila Pa 1976) 2001;26:1890-4; discussion 1895.

36. Malone A, Meldrum D, Gleeson J, et al. Reliability of surface electromyography timing parameters in gait in cervical spondylotic myelopathy. J Electromyogr Kinesiol 2011;21:1004-10.

37. Singh A, Crockard HA. Comparison of seven different scales used to quantify severity of cervical spondylotic myelopathy and post-operative improvement. J Outcome Meas 2001-2002;5:798-818.
38. Enoki H, Tani T, Ishida K. Foot Tapping Test as Part of Routine Neurologic Examination in Degenerative Compression Myelopathies: A Significant Correlation between 10-sec Foot-tapping Speed and 30-m Walking Speed. Spine Surg Relat Res 2019;3:207-13.
39. Maezawa Y, Uchida K, Baba H. Gait analysis of spastic walking in patients with cervical compressive myelopathy. J Orthop Sci 2001;6:378-84.
40. Mihara H, Kondo S, Murata A, et al. A new performance test for cervical myelopathy: the triangle step test. Spine (Phila Pa 1976) 2010;35:32-5.
41. Zheng CF, Liu YC, Hu YC, et al. Correlations of Japanese Orthopaedic Association Scoring Systems with Gait Parameters in Patients with Degenerative Spinal Diseases. Orthop Surg 2016;8:447-53.
42. Ghent F, Mobbs RJ, Mobbs RR, et al. Assessment and Post-Intervention Recovery After Surgery for Lumbar Disk Herniation Based on Objective Gait Metrics from Wearable Devices Using the Gait Posture Index. World Neurosurg 2020;142:e111-6.
43. Mobbs RJ, Mobbs RR, Choy WJ. Proposed objective scoring algorithm for assessment and intervention recovery following surgery for lumbar spinal stenosis based on relevant gait metrics from wearable devices: the Gait Posture index (GPi). J Spine Surg 2019;5:300-9.
44. Chakravorty A, Mobbs RJ, Anderson DB, et al. The role of wearable devices and objective gait analysis for the assessment and monitoring of patients with lumbar spinal stenosis: systematic review. BMC Musculoskelet Disord 2019;20:288.