

Dystonia Rating Scales: Critique and Recommendations

Alberto Albanese, MD,^{1,2*} Francesca Del Sorbo, MD,² Cynthia Comella, MD,³ H. A. Jinnah, PhD,^{4,5,6} Jonathan W. Mink, MD, PhD,⁷ Bart Post, MD, PhD,⁸ Marie Vidailhet, MD,⁹ Jens Volkmann, MD, PhD,¹⁰ Thomas T. Warner, FRCP,¹¹ Albert F.G. Leentjens, MD, PhD,¹² Pablo Martinez-Martin, MD,¹³ Glenn T. Stebbins, PhD,¹⁴ Christopher G. Goetz, MD,¹⁴ and Anette Schrag, PhD¹⁵

¹*Istituto di Neurologia, Università Cattolica del Sacro Cuore, Milano, Italy*

²*Neurologia I, Istituto Neurologico Carlo Besta, Milano, Italy*

³*Department of Neurological Sciences, Rush Medical College, Chicago, Illinois, USA*

⁴*Department of Neurology, Emory University, Atlanta, Georgia, USA*

⁵*Department of Human Genetics, Emory University, Atlanta, Georgia, USA*

⁶*Department of Pediatrics, Emory University, Atlanta, Georgia, USA*

⁷*Department of Neurology, University of Rochester Medical Center, Rochester, New York, USA*

⁸*Department of Neurology, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands*

⁹*Department of Neurology, Pitié-Salpêtrière Hospital, Paris, France*

¹⁰*Department of Neurology, University of Würzburg, Würzburg, Germany*

¹¹*Reta Lila Weston Institute of Neurological Studies, University College London (UCL) Institute of Neurology, London, United Kingdom*

¹²*Department of Psychiatry, Maastricht University Medical Center, Maastricht, the Netherlands*

¹³*Alzheimer Disease Research Unit and Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Alzheimer Center Reina Sofía Foundation, Carlos III Institute of Health, Madrid, Spain*

¹⁴*Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, USA*

¹⁵*Department of Clinical Neurosciences, Royal Free Hospital, University College London, London, UK*

ABSTRACT: Many rating scales have been applied to the evaluation of dystonia, but only few have been assessed for clinimetric properties. The Movement Disorders Society commissioned this task force to critique existing dystonia rating scales and place them in the clinical and clinimetric context. A systematic literature review was conducted to identify rating scales that have either been validated or used in dystonia. Thirty-six potential scales were identified. Eight were excluded because they did not meet review criteria, leaving 28 scales that were critiqued and rated by the task force. Seven scales were found to meet criteria to be “recommended”: the Blepharospasm Disability Index is recommended for rating blepharospasm; the Cervical Dystonia Impact Scale and the Toronto Western Spasmodic Torticollis Rating Scale for rating cervical dystonia; the Craniocervical Dystonia Questionnaire for blepharospasm and cervical dystonia; the Voice Handicap Index (VHI) and the Vocal Performance Questionnaire (VPQ) for

laryngeal dystonia; and the Fahn-Marsden Dystonia Rating Scale for rating generalized dystonia. Two “recommended” scales (VHI and VPQ) are generic scales validated on few patients with laryngeal dystonia, whereas the others are disease-specific scales. Twelve scales met criteria for “suggested” and 7 scales met criteria for “listed.” All the scales are individually reviewed in the online information. The task force recommends 5 specific dystonia scales and suggests to further validate 2 recommended generic voice-disorder scales in dystonia. Existing scales for oromandibular, arm, and task-specific dystonia should be refined and fully assessed. Scales should be developed for body regions for which no scales are available, such as lower limbs and trunk.
© 2013 Movement Disorder Society

Key Words: botulinum neurotoxins; blepharospasm; dystonias; torticollis; rating scales

Additional Supporting Information may be found in the online version of the article.

*Correspondence to: Dr. Alberto Albanese, Neurologia I, IRCCS Istituto Neurologico Carlo Besta, Via G. Celoria, 11, 20133 Milano, Italy; alberto.albanese@unicatt.it

Relevant conflicts of interest/financial disclosures: Full financial disclosures and author roles may be found in the Acknowledgments section online.

Received: 10 May 2013; **Accepted:** 22 May 2013

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/mds.25579

Dystonia is one of the most common movement disorders, with an overall prevalence of 16.43 per 100,000 for primary dystonia.¹ This meta-analysis prevalence figure is likely to be an underestimate, as it is based on studies with recruitment of diagnosed cases only and it is clear that under-ascertainment and underdiagnosis is a significant problem.² The broad spectrum of clinical features that encompass dystonia syndromes ranges from severe generalized childhood dystonia, to adult-onset focal dystonias, to secondary dystonias and dystonias as a feature of complex neurological disorders. Dystonia can be localized to a single body region (focal) or has spread to contiguous (segmental) or to noncontiguous (multifocal) regions. In generalized dystonia the trunk and at least 2 other sites are involved,³ whereas hemidystonia affects the body and limbs on one side.

Dystonia is typically considered a movement disorder characterized by motor manifestations, primarily sustained or intermittent muscle contractions causing abnormal, often repetitive, movements, postures, or both.³ However, growing evidence indicates the importance of a non-motor component to dystonia, including abnormalities in sensory and perceptual functions, as well as neuropsychiatric, cognitive and sleep domains.⁴ Treatment possibilities have greatly expanded in recent years after discovering the efficacy of botulinum neurotoxins and functional surgery.^{5,6} Pretreatment evaluation aims at characterizing the severity and topography of motor symptoms and their impact on activities of daily living and provides a baseline reference for posttreatment evaluations. The quality and accuracy of the pretreatment assessment and the choice of assessment tools are crucial as they will affect all subsequent posttreatment comparisons. Precise tools to rate improvement or deterioration are important to assess the patient's disease state as well as outcome after treatment.

To facilitate research and clinical practices aimed at improving the assessment and treatment of dystonia syndromes, the Movement Disorders Society (MDS) convened a task force to evaluate the dystonia rating instruments that have been used in published studies. This review is part of a process to assess scales currently in use for evaluating clinical aspects of movement disorders.

Materials and Methods

Administrative Organization and Critique Process

The MDS Task Force on Rating Scales for Movement Disorders Steering Committee invited the chairman (AA) to form a task force to critique existing

dystonia rating scales and to place them in a clinical and clinimetric context. This task force consisted of 8 members from Europe and North America with diverse background and expertise, including neurologists, a neuropsychiatrist, and a clinical epidemiologist, who had worked extensively in the area of dystonia. This group followed the same procedure as the task forces that appraised other rating scales in movement disorders. Initial discussions among these task force members regarded the construct to be reviewed, in the case its concept was not universally accepted. Then the task force members selected the scales to be included in the review (see criteria in the next paragraph) and identified unresolved issues and limitations of the critiqued scales. A standardized form was drawn up to allow structured assessment of the scales with regard to their descriptive properties, availability, content, use, acceptability, clinimetric properties, and overall impression in patients with dystonia (see online Supporting Information).

Each scale was reviewed by 1 task force member. The completed reviews were then assessed by 2 members and modified according to their suggestions. In the final appraisal of a scale, the task force used the terminology developed for the Appendix of Ancillary Scales to complement the MDS-sponsored revision of the Unified Parkinson's Disease Rating Scale.⁷ This terminology was also used in recent reviews on rating scales of other MDS task forces.⁸ The final assessment was based on consensus among the task force members and the Steering Committee of the Task Force on Rating Scales for Movement Disorders. The following criteria were specifically distilled from the available evidence. Criterion 1: The scale has been applied to dystonia patients; criterion 2: the scale has been used by other groups outside the original developers; criterion 3: the scale has been clinimetrically studied and found to be valid, reliable, and sensitive to change. The official definitions for task force critiques are: "recommended," if the scale fulfills criteria 1, 2, and 3; "suggested," if the scale fulfills criterion 1, but only 1 of the other criteria applies; or "listed," if the scale fulfills criterion 1, but does not meet either of the other 2 criteria.

Scale Selection Process and Literature Search Strategy

We considered all scales and questionnaires that have either been designed or used to rate dystonia, and in addition scales and questionnaires that, based on literature review and expert evaluation, have potential utility in dystonia based on their content, their widespread use, and clinimetric evidence from studies in patients without dystonia. We did not consider scales and questionnaires specifically designed for secondary dystonias

or requiring measurement devices. We included in the main document scales and questionnaires that have been “recommended” for use in dystonia. All scales are listed in the online Supporting Information.

The Medline database on PubMed was systematically searched for relevant papers published up to June 2012 using the following query: (("Dystonia"[MH]) OR ("Dystonia Musculorum Deformans"[MH]) OR ("Dystonic Disorders"[MH])) AND ("scale"[ALL] OR "measure"[ALL] OR "Questionnaire"[ALL]). For each scale, a search was conducted for the terms ("spasmodic dysphonia"[ALL]) OR ("Dystonic Disorders"[MH]) OR ("Dystonia"[MH]) OR ("Blepharospasm"[MH]) AND the name of the scale. In addition, published articles known to the Task Force members were included in this review. Only published or in press peer-reviewed papers or published abstracts were evaluated.

Results

Thirty-six scales and questionnaires were identified. Based on inclusion criteria, 28 measures were considered (Table 1); these were classified as “specific” if developed specifically to rate dystonia, or as “generic” if applicable across different diseases, including dystonia. Among the generic scales, 1 was developed to measure coping in chronically ill populations and 6 were originally developed to quantify the degree of dysphonia and to objectively determine the efficacy of voice therapy in voice disorders. Eight scales were excluded, because evaluating secondary dystonia, requiring measurement devices, or having no potential use in dystonia. The scales are grouped based on the affected body region they intend to explore.

Blepharospasm Scales

Blepharospasm Disability Index

Description of the Scale. The Blepharospasm Disability Index (BSDI)⁹ was developed to improve the Blepharospasm Disability Scale (see online Supporting Information) with respect to ease of use. It is a disease-specific patient-rated disability scale that measures impairment of specific activities of daily living caused by blepharospasm. It consists of 6 items rating specified activities (vehicle driving, reading, watching television, shopping, walking, and doing everyday activities), scored as a 5-point Likert scale relating to the severity of impairment (0, no impairment; 4, no longer possible due to illness), as well as a “not applicable” option. The range of scores is 0 to 24, with higher scores indicating a greater disability. A BSDI mean item score can also be calculated by dividing the total BSDI score by the number of items answered. It is available only in English, although the scale has been used extensively in Europe and Israel.

Scale Application in Dystonia. The BSDI was designed specifically for patients with blepharospasm.

Use by Multiple Groups Outside the Original Developers. The BSDI has been used in several recent trials with botulinum neurotoxins (BoNT).^{10–12}

Clinimetric Properties. The BSDI showed high internal consistency and the retest reliability of the single items was adequate.⁹ The BSDI total score was found to correlate moderately with the Jankovic Rating Scale score.¹² The results of an observational study showed that BSDI was sensitive to change after BoNT treatment.⁹ Data from 2 large randomized trials designed to evaluate the effects of BoNT type A products for blepharospasm showed that the BSDI was sensitive to change but did not detect differences between 2 BoNT products.^{10,11}

Strengths and Weaknesses. The BSDI has been specifically designed to measure disability in blepharospasm due to dystonic movements that affect vision. The scale focuses on daily activities and is easy to use; the scoring system is also rather simple. The BSDI focuses on disability related to sight and does not specifically measure dystonic motor abnormalities; it should be combined with a more specific scale that rates the movement disorder. Concern has been raised regarding poor sensitivity of the scale to mild disability or small changes.¹³

Cervical Dystonia Scales

Cervical Dystonia Impact Scale

Description of the Scale. The Cervical Dystonia Impact Scale (CDIP-58) is a disease-specific patient-rated questionnaire that measures quality of life in patients with cervical dystonia.¹⁴ It was developed for use in clinical research, audit, and treatment trials. It is composed of 58 five-point items grouped into 8 sub-scales that measure symptoms (head and neck movements, pain and discomfort in neck and shoulders, sleep disturbance as a result of torticollis), activity limitations in upper limb activities and walking, and psychosocial features (annoyance, mood, psychosocial functioning). Eight summary scale scores are generated by summing items and are then transformed to a 0 to 100 score. This scale is available only in English.

Scale Application in Dystonia. The CDIP-58 has been specifically developed for patients with cervical dystonia.

Use by Multiple Groups Outside the Original Developers. The CDIP-58 has been used to measure the health impact of cervical dystonia and the impact of treatment in cervical dystonia.^{15–18}

TABLE 1. Classification of the rating measures applied in studies of dystonia

Scales/questionnaire	Type ^b	Criteria ^a			Qualification
		1	2	3	
Blepharospasm					
Blepharospasm Disability Index (BSDI)	Specific	Yes	Yes	Yes	Recommended
Jankovic rating scale	Specific	Yes	Yes	No	Suggested
Blepharospasm Disability Scale	Specific	Yes	Yes	No	Suggested
Cervical dystonia					
Cervical Dystonia Impact Scale (CDIP-58)	Specific	Yes	Yes	Yes	Recommended
Functional Disability Questionnaire	Specific	Yes	Yes	No	Suggested
Toronto Western Spasmodic Torticollis Rating Scale (TWSTRS)	Specific	Yes	Yes	Yes	Recommended
Tsui scale	Specific	Yes	Yes	No	Suggested
Modified Tsui scale	Specific	Yes	No	No	Listed
Freiberg Questionnaire for Dystonia torticollis version	Specific	Yes	No	No	Listed
Disability questionnaire for patients with cervical dystonia	Specific	Yes	No	No	Listed
Body Concept Scale	Specific	Yes	No	Yes	Suggested
Ways of Coping Checklist	Generic	Yes	No	No	Listed
Blepharospasm/cervical dystonia					
Craniocervical Dystonia Questionnaire (CDQ-24)	Specific	Yes	Yes	Yes	Recommended
Oromandibular dystonia					
Oromandibular dystonia questionnaire	Specific	Yes	No	Yes	Suggested
Laryngeal dystonia					
Unified Spasmodic Dysphonia Rating Scale	Specific	Yes	Yes	No	Suggested
Voice Handicap Index (VHI)	Generic	Yes	Yes	Yes	Recommended
Voice Handicap Index 10	Generic	Yes	Yes	No	Suggested
Pediatric Voice Handicap Index	Generic	Uncertain	Yes	No	Listed
Pediatric Voice-Related Quality of Life	Generic	Uncertain	Yes	No	Listed
Voice-Related Quality of Life	Generic	Yes	Yes	No	Suggested
Vocal Performance Questionnaire (VPQ)	Generic	Yes	Yes	Yes	Recommended
Arm dystonia					
Arm Dystonia Disability Scale	Specific	Yes	Yes	No	Suggested
Task-specific dystonia					
Dystonia Evaluation Scale	Specific	Yes	No	No	Listed
Tubiana-Chamagne Score	Specific	Yes	Yes	No	Suggested
Writer's Cramp Rating Scale	Specific	Yes	Yes	No	Suggested
Generalized dystonia					
Global Dystonia rating Scale	Specific	Yes	Yes	No	Suggested
Fahn-Marsden Dystonia Rating Scale (FMDRS)	Specific	Yes	Yes	Yes	Recommended
Unified Dystonia Rating Scale	Specific	Yes	Yes	No	Suggested

^aCriteria are as follows: 1 used in dystonia patients; 2 used by researchers beyond original developers; 3 successful clinimetric testing.

^b"Specific" indicates a measure developed specifically for dystonia, "generic" indicates a measure applicable across different diseases, including dystonia.

Clinimetric Properties. New psychometric techniques (Rasch analyses) revealed that the CDIP-58 performs well and, in addition, traditional psychometric properties such as reliability (internal consistency, item-total correlation, test-retest) and validity (within-scale analyses and comparisons with external measures) have been supported.^{14,15,19} The CDIP-58 is good at detecting the impact of BoNT on all 8 health dimensions in patients with cervical dystonia.¹⁵

Strengths and Weaknesses. The CDIP-58 is a disease-specific validated questionnaire. It is more sensitive in detecting statistical and clinical changes than comparable subscales, although it has not been widely used as an outcome measure. The authors themselves¹⁵ suggest further studies examining the responsiveness of the CDIP-58 as well as refinement of the walking subscale.

Toronto Western Spasmodic Torticollis Rating Scale

Description of the Scale. The Toronto Western Spasmodic Torticollis Rating Scale (TWSTRS) was developed for assessment of cervical dystonia in clinical trials.²⁰ It is composed of 3 subscales that measure symptom severity, disability, and pain. The severity scale, clinician-rated, is composed of 11 items that assess head movements, duration of symptoms, effects of sensory tricks, shoulder elevation and anterior displacement, range of motion, and time in neutral position; the maximal score is 35. The disability scale, patient-rated, comprises 6 items, including daily activities, work, reading, and driving; the maximal score is 30. The pain scale, patient-rated, comprises 3 items including severity, duration, and disability due to pain; the maximal score is 20. Each subscale is scored

independently and a total TWSTRS score (from 0 to 85) is calculated. A training tape for clinicians is available for the severity scale.²¹ The only available version is in English.

Scale Application in Dystonia. The TWSTRS has been developed specifically for patients with cervical dystonia.

Use by Multiple Groups Outside the Original Developers. The TWSTRS scale is the most widely used rating scale for cervical dystonia. Individual subscales and the total TWSTRS score have been used as outcome measures in many treatment trials, evaluating BoNT therapy, pharmacotherapy, and surgery.^{22–58}

Clinimetric Properties. The TWSTRS has been shown to have internal consistency and acceptable interrater agreement. Evidence for validity is shown by moderate within-scale correlations.²⁰ The TWSTRS scale also showed strong correlation with Tsui scale.⁵⁹ Responsiveness to change has been demonstrated.^{25,59}

Strengths and Weaknesses. The TWSTRS assesses the severity of cervical dystonia and includes disability and pain subscales. The TWSTRS includes a videotape protocol allowing its use to evaluate all patients in a standardized fashion.²¹ Despite its value in clinical trials, the TWSTRS scale might be too complex for routine clinical practice. Weaknesses consist of an unclear definition of midline for assessing range of motion, lack of a separate scoring category assessing dystonic tremor, and the specification of duration for the effect of sensory tricks.²¹

Blepharospasm/Cervical Dystonia Scale Craniocervical Dystonia Questionnaire

Description of the Scale. The Craniocervical Dystonia Questionnaire (CDQ-24) is a patient-rated health-related quality of life (HRQoL) measure for craniocervical dystonia, featuring cervical dystonia and blepharospasm. It was developed for use in clinical research.⁶⁰ CDQ-24 measures the impact of craniocervical dystonia on 5 HRQoL domains. It is composed of 24 items, forming 5 subscales: stigma, emotional well-being, pain, activities of daily living, and social/family life. Items are rated on a 5-point scale. Although only the original German version was validated, an exact translation into English, including back-translation, was performed.⁶⁰ The CDQ-24 has also been translated and validated in Serbian.⁶¹

Scale Application in Dystonia. The CDQ-24 has been specifically developed for patients with craniocervical dystonia, who had both cervical dystonia and blepharospasm.

Use by Multiple Groups Outside the Original Developers. The CDQ-24 has been used by multiple groups to measure the impact on quality of life of focal, segmental and even generalized dystonia,¹⁶ and also as a HRQoL measure to assess responsiveness to treatment-induced changes.^{60,62}

Clinimetric Properties. There were no relevant ceiling effects, but a considerable floor effect was observed in the social/family life domain.⁶⁰ The CDQ-24 also showed good reliability properties, internal consistency, and test-retest reliability. Validity was assessed by checking convergent and discriminant validity as well as the dimensional structure of CDQ-24; sensitivity to change was confirmed after BoNT treatment.⁶⁰

Strengths and Weaknesses. The CDQ-24 is a brief and easy instrument. It can be used to evaluate the impact of the disease on areas not covered by generic measures, such as the Short Form (36) Health Survey (SF-36), which are of considerable concern to patients with craniocervical dystonia. The CDQ-24 also evaluates pain, sleep, and depression due to dystonia.

Laryngeal Dystonia Scales

Voice Handicap Index

Description of the Scale. The Voice Handicap Index (VHI) is a patient-rated scale addressing disability related to verbal communication. It was developed to determine the level of disability experienced by patients with different voice disorders.⁶³ The complete VHI has 30 items organized in 3 domains: a 10-item functional subscale, a 10-item emotional subscale, and a 10-item physical subscale. The rating is on a 5-point scale and the total score ranges from 0 to 120. VHI has been translated and clinimetrically tested in German, Mandarin Chinese, Spanish, Dutch, Arabic, Japanese, Hebrew, and Greek. The validity of the French translation has been confirmed, although the quality of translation needs further improvements.⁶⁴ The VHI has been also translated and adapted to Portuguese and Polish.

Scale Application in Dystonia. The VHI is not specific for dystonia-related voice problems; in the original development and validation study, 26% of the patients had neurogenic voice disorders, including vocal fold paralysis and laryngeal dystonia.⁶³

Use by Multiple Groups Outside the Original Developers. VHI has been used to measure outcomes after interventions for a broad range of laryngeal disorders, including cancer and mass lesions, vocal fold polyps and cysts, and laryngeal dystonia.^{65–75}

Clinimetric Properties. In the development and validation study performed on a heterogeneous set of

disorders,⁶³ the VHI proved to have good internal consistency and good test-retest reliability for subscales and total scores. Construct validity was not fully evaluated. The VHI has been used in several studies to assess efficacy of treatments for laryngeal dystonia. However, considering that the VHI was validated on few patients with laryngeal dystonia compared to the total number of patients assessed, it still needs further validation for dystonia.

Strengths and Weaknesses. The VHI is a simple and efficient scale, but as a disability scale, it has no discriminant value and is not specific for dystonia. Therefore, the scale should be further validated specifically in spasmodic dysphonia. The VHI is similar to the Vocal Performance Questionnaire, and direct comparisons have been made showing similar clinimetric properties.

Vocal Performance Questionnaire

Description of the Scale. This scale was designed for use in an evaluation study of voice therapy in cases of nonorganic dysphonia.⁷⁶ The Vocal Performance Questionnaire (VPQ) is a questionnaire designed to allow patients to consider aspects of their own vocal performance and rate their severity. This 12-item questionnaire is designed using an answer format in which the patient selects the statement that best answers each question. The statements are graded in terms of severity of vocal performance. This scale is available only in English.

Scale Application in Dystonia. The VPQ has been used in dystonia in only 1 study,⁷⁷ which evaluated the reliability and validity of the scale in 181 patients with different voice disorders, including an undetermined number of patients with laryngeal dystonia.

Use by Multiple Groups Outside the Original Developers. The VPQ has been used to measure outcomes in interventions in several trials none of which were performed on patients with dystonia.^{78–81}

Clinimetric Properties. The VPQ was found to have good internal consistency in a study that included a large range of voice pathologies except for spasmodic dysphonia.⁸² In a study that included patients with laryngeal dystonia,⁷⁷ the VPQ had high levels of internal consistency and test-retest reliability. The VHI-10—the short form of VHI—and the VPQ were highly correlated in a study that did not include patients with laryngeal dystonia.⁸² Therefore, the VPQ still needs further validation in patients with dystonia.

Strengths and Weaknesses. The VPQ is a simple and efficient scale, but as a disability scale, it has no

discriminant value and is not specific for dystonia. Therefore, the scale should be further validated specifically in spasmodic dysphonia. The VPQ is similar to the VHI (see *Voice Handicap Index*). The value of having 2 scales for the same purpose is questionable and a sensible recommendation would be to merge them or pick 1 for future use.

Generalized Dystonia Scales

Fahn-Marsden Dystonia Rating Scale

Description of the Scale. The Fahn-Marsden Dystonia Rating Scale (FMDRS) is composed of 2 clinician-rated subscales: a movement subscale, based on patient examination, and a disability subscale, based on the patient's report of disability in activities of daily living. The movement subscale rates dystonia severity and provoking factors in 9 body areas, including eyes, mouth, speech and swallowing, neck, trunk, and both arms and legs. All items have a 5-point score. The provoking factor rates the relation of dystonia to action, from 0 (no dystonia at rest or with action) to 4 (dystonia at rest). The score obtained for eyes, mouth, and neck are each multiplied by 0.5, before being entered into the calculation of the total score, in order to down-weight them. The total movement FMDRS subscore is provided by the sum of the products of the provoking, severity, and weighting factors. The maximal total FMDRS score is 120. The disability subscale is composed of 7 items for activities of daily living, such as speech, writing, feeding, eating, hygiene, dressing, and walking. These are rated on a 5-point score (with the exception of walking, which is rated on a 7-point score), providing a maximum disability subscore of 30. Training for administration is recommended.

Scale Application in Dystonia. The FMDRS was originally established for the clinical assessment of primary torsion dystonia in adults.⁸³

Use by Multiple Groups Outside the Original Developers. The FMDRS has been used in numerous studies to determine the treatment effects of deep brain stimulation,^{23,34,46,84–109} including childhood-onset dystonia.^{110–114}

Clinimetric Properties. In the original validation study the reliability, interrater agreement, and concurrent validity of the FMDRS were demonstrated for the total score without reporting the level of agreement for ratings of the different body regions.⁸³ The FMDRS showed good internal consistency and good level of interrater reliability for the total scores.¹¹⁵ For separate items, interrater agreement was fair to good, being lowest for eyes, jaw, face, and larynx.¹¹⁵ The modifying ratings for the FMDRS (Provoking Factor)

showed consistently lower levels of agreement than motor severity ratings.¹¹⁵ The total scores for the FMDS, the Unified Dystonia Rating Scale, and the Global Dystonia Rating Scale were highly correlated with each other.¹¹⁵ Responsiveness has been demonstrated in treatment studies.^{84,91,110}

Strengths and Weaknesses. Limitations in the FMDS include a weighting factor that halves the contribution of dystonia in eyes, mouth, and neck to the total score. The FMDS does not assess in detail the individual body areas, such as separate ratings for proximal and distal limbs; moreover, included in the FMDS there is a subjective patient rating for speech and swallowing.

Discussion

We identified 7 scales that fulfilled the predefined criteria for “recommended” scales. One scale rates blepharospasm (BSDI), 2 rate cervical dystonia (CDIP-58, TWSTRS), 1 rates blepharospasm and cervical dystonia (CDQ-24), 2 rate laryngeal dystonia (VHI, VPQ), and 1 rates generalized dystonia (FMDS). Two of these are generic scales (VHI and VPQ) that require further validation specifically in dystonia, while the remaining are disease-specific scales. The task force recommends the 5 specific dystonia scales and suggests that the 2 recommended generic voice-disorder scales be further validated in dystonia. Scales for oromandibular, arm, and task-specific dystonias require further assessment and there are no rating scales for some body areas, particularly the trunk and lower limbs. Eleven of the recommended scales provide objective evaluations, 15 provide measurement of disability or quality of life, and 2 are psychosocial scales.

Each of these scales has been shown to have specific advantages and limitations in dystonia and all have been shown to have adequate clinimetric properties for the assessment of dystonia. However, these scales are useful mainly in assessing the motor aspects of dystonia, and only 2 of them (TWSTRS and FMDS) assess some of the specific motor phenomena of dystonia, such as action specificity, gestes antagonistes, or temporal patterns. Non-motor symptoms such as sensory, sleep, and neuropsychiatric features related to dystonia are partially rated in the TWSTRS, CDIP-58, CDQ-24, and in some of the suggested or listed scales.

None of the reviewed scales is appropriate or sufficient to diagnose a specific dystonia type (eg, specific types of focal or generalized dystonia, paroxysmal, etc.), but these instruments can rate its severity and make comparison within different patient groups. Since most dystonia scales measure specific body regions, they should be applied to well-selected and

homogeneous patient groups. For example, the CDQ-24, a scale that measures the impact of craniocervical dystonia on quality of life, has been used in patients with segmental and generalized dystonia.¹⁶ However, this scale is specific to craniocervical dystonia; its use in patients with dystonia also involving other body regions may lead to misleading results.

Most of the scales used to rate dystonia were designed for adults and then applied to children. Evaluating children with dystonia is difficult, because a wider spectrum of abnormalities may be commonly associated with dystonia. Unlike adults, children frequently have secondary forms that can be confounded with other motor abnormalities, including weakness, spasticity, impaired selective motor control, bradykinesia, choreoathetosis, ataxia, and sensory impairments. Therefore, rating scales in children are designed to evaluate secondary dystonias, including a broad range of movement disorders different from dystonia (ie, the Barry-Albright Dystonia Scale, the Movement Disorder-Childhood Rating Scale) and are not reviewed here. Application of adult dystonia scales to children is further complicated by the impact of development on expressed motor patterns and skills. Thus, further validation of dystonia scales in children with primary dystonia and in those with secondary dystonias is needed.

Future directions will encompass the refinement of existing rating scales to include various specific motor as well as non-motor features of dystonia,³ and fuller clinimetric assessment for oromandibular, arm, and task-specific dystonias. There is also a need for the development of new tools for the dystonia types where no scales are available, such as lower limb and trunk dystonias. The selection of the most appropriate instrument for each particular dystonia type is advocated and the need for training physicians in recognizing the complex phenomenology of dystonia syndromes. Scales need to be evaluated in different populations such as in children versus adults, and primary versus secondary dystonias, and translations should be available. Finally, there is a need for uniform training by developing manuals and training tools for dystonia scales. ■

References

1. Steeves TD, Day L, Dykeman J, Jette N, Pringsheim T. The prevalence of primary dystonia: a systematic review and meta-analysis. *Mov Disord* 2012;27:1789–1796.
2. Albanese A, Asmus F, Bhatia KP, et al. EFNS guidelines on diagnosis and treatment of primary dystonias. *Eur J Neurol* 2011;18:5–18.
3. Albanese A, Bhatia KP, Bressman SB, et al. Phenomenology and classification of dystonia: a consensus update. *Mov Disord* 2013; 863–873.
4. Stamelou M, Edwards MJ, Hallett M, Bhatia KP. The non-motor syndrome of primary dystonia: clinical and pathophysiological implications. *Brain* 2012;135:1668–1681.

5. Hallett M, Benecke R, Blitzer A, Comella CL. Treatment of focal dystonias with botulinum neurotoxin. *Toxicon* 2009;54:628–633.
6. Cloud LJ, Jinnah HA. Treatment strategies for dystonia. *Expert Opin Pharmacother* 2010;11:5–15.
7. Goetz CG, Tilley BC, Shaftman SR, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord* 2008;23:2129–2170.
8. The Unified Parkinson's Disease Rating Scale (UPDRS): status and recommendations. *Mov Disord* 2003;18:738–750.
9. Jankovic J, Kenney C, Grafe S, Goertelmeyer R, Comes G. Relationship between various clinical outcome assessments in patients with blepharospasm. *Mov Disord* 2009;24:407–413.
10. Roggenkamper P, Jost WH, Bihari K, Comes G, Grafe S. Efficacy and safety of a new botulinum toxin type A free of complexing proteins in the treatment of blepharospasm. *J Neural Transm* 2006;113:303–312.
11. Wabbels B, Reichel G, Fulford-Smith A, Wright N, Roggenkamper P. Double-blind, randomised, parallel group pilot study comparing two botulinum toxin type A products for the treatment of blepharospasm. *J Neural Transm* 2011;118:233–239.
12. Jankovic J, Comella C, Hanschmann A, Grafe S. Efficacy and safety of incobotulinumtoxinA (NT 201, Xeomin) in the treatment of blepharospasm—a randomized trial. *Mov Disord* 2011;26:1521–1528.
13. Wabbels B, Jost WH, Roggenkamper P. Difficulties with differentiating botulinum toxin treatment effects in essential blepharospasm. *J Neural Transm* 2011;118:925–943.
14. Cano SJ, Warner TT, Linacre JM, et al. Capturing the true burden of dystonia on patients: the Cervical Dystonia Impact Profile (CDIP-58). *Neurology* 2004;63:1629–1633.
15. Cano SJ, Hobart JC, Edwards M, et al. CDIP-58 can measure the impact of botulinum toxin treatment in cervical dystonia. *Neurology* 2006;67:2230–2232.
16. Zetterberg L, Aquilonius SM, Lindmark B. Impact of dystonia on quality of life and health in a Swedish population. *Acta Neurol Scand* 2009;119:376–382.
17. Jankovic J, Adler CH, Charles PD, et al. Rationale and design of a prospective study: Cervical Dystonia Patient Registry for Observation of OnabotulinumtoxinA Efficacy (CD PROBE). *BMC Neurol* 2011;11:140.
18. Sethi KD, Rodriguez R, Olayinka B. Satisfaction with botulinum toxin treatment: a cross-sectional survey of patients with cervical dystonia. *J Med Econ* 2012;15:419–423.
19. Cano SJ, Warner TT, Thompson AJ, Bhatia KP, Fitzpatrick R, Hobart JC. The cervical dystonia impact profile (CDIP-58): can a Rasch developed patient reported outcome measure satisfy traditional psychometric criteria? *Health Qual Life Outcomes* 2008;6:58.
20. Consky ES, Lang AE. Clinical assessments of patients with cervical dystonia. In: Jankovic J, Hallett M, editors. *Therapy with botulinum toxin*. New York, NY: Marcel Dekker; 1994:211–237.
21. Comella CL, Stebbins GT, Goetz CG, Chmura TA, Bressman SB, Lang AE. Teaching tape for the motor section of the Toronto Western Spasmodic Torticollis Scale. *Mov Disord* 1997;12:570–575.
22. Lalli S, Piacentini S, Franzini A, et al. Epidural premotor cortical stimulation in primary focal dystonia: clinical and 18F-fluoro deoxyglucose positron emission tomography open study. *Mov Disord* 2012;27:533–538.
23. Capelle HH, Blahak C, Schrader C, et al. Bilateral deep brain stimulation for cervical dystonia in patients with previous peripheral surgery. *Mov Disord* 2012;27:301–304.
24. Skogseid IM, Ramm-Petersen J, Volkmann J, Kerty E, Dietrichs E, Roste GK. Good long-term efficacy of pallidal stimulation in cervical dystonia: a prospective, observer-blinded study. *Eur J Neurol* 2012;19:610–615.
25. Comella CL, Jankovic J, Truong DD, Hanschmann A, Grafe S. Efficacy and safety of incobotulinumtoxinA (NT 201, XEOMIN(R)), botulinum neurotoxin type A, without accessory proteins in patients with cervical dystonia. *J Neurol Sci* 2011;308:103–109.
26. Ostrem JL, Racine CA, Glass GA, et al. Subthalamic nucleus deep brain stimulation in primary cervical dystonia. *Neurology* 2011;76:870–878.
27. Chinnapongse R, Pappert EJ, Evatt M, Freeman A, Birmingham W. An open-label, sequential dose-escalation, safety, and tolerability study of rimabotulinumtoxinb in subjects with cervical dystonia. *Int J Neurosci* 2010;120:703–710.
28. Cacciola F, Farah JO, Eldridge PR, Byrne P, Varma TK. Bilateral deep brain stimulation for cervical dystonia: long-term outcome in a series of 10 patients. *Neurosurgery* 2010;67:957–963.
29. Lew MF, Chinnapongse R, Zhang Y, Corliss M. RimabotulinumtoxinB effects on pain associated with cervical dystonia: results of placebo and comparator-controlled studies. *Int J Neurosci* 2010;120:298–300.
30. Truong D, Brodsky M, Lew M, et al. Long-term efficacy and safety of botulinum toxin type A (Dysport) in cervical dystonia. *Parkinsonism Relat Disord* 2010;16:316–323.
31. Quagliato EM, Carelli EF, Viana MA. A prospective, randomized, double-blind study comparing the efficacy and safety of type A botulinum toxins botox and prosigine in the treatment of cervical dystonia. *Clin Neuropharmacol* 2010;33:22–26.
32. Moro E, Piboolnarak P, Arenovich T, Hung SW, Poon YY, Lozano AM. Pallidal stimulation in cervical dystonia: clinical implications of acute changes in stimulation parameters. *Eur J Neurol* 2009;16:506–512.
33. Berman BD, Starr PA, Marks WJ, Ostrem JL. Induction of bradykinesia with pallidal deep brain stimulation in patients with cranial-cervical dystonia. *Stereotact Funct Neurosurg* 2009;87:37–44.
34. Pretto TE, Dalvi A, Kang UJ, Penn RD. A prospective blinded evaluation of deep brain stimulation for the treatment of secondary dystonia and primary torticollis syndromes. *J Neurosurg* 2008;109:405–409.
35. Pinski MO, Volkmann J, Falk D, et al. Electrode implantation for deep brain stimulation in dystonia: a fast spin-echo inversion-recovery sequence technique for direct stereotactic targeting of the GPI. *Zentralbl Neurochir* 2008;69:71–75.
36. Pappert EJ, Germanson T. Botulinum toxin type B vs. type A in toxin-naïve patients with cervical dystonia: Randomized, double-blind, noninferiority trial. *Mov Disord* 2008;23:510–517.
37. Kiss ZH, Doig-Beyaert K, Eliasziw M, Tsui J, Haffenden A, Suchowersky O. The Canadian multicentre study of deep brain stimulation for cervical dystonia. *Brain* 2007;130:2879–2886.
38. Skogseid IM, Malt UF, Roislien J, Kerty E. Determinants and status of quality of life after long-term botulinum toxin therapy for cervical dystonia. *Eur J Neurol* 2007;14:1129–1137.
39. Comella CL, Jankovic J, Shannon KM, et al. Comparison of botulinum toxin serotypes A and B for the treatment of cervical dystonia. *Neurology* 2005;65:1423–1429.
40. Ondo WG, Gollomp S, Galvez-Jimenez N. A pilot study of botulinum toxin A for headache in cervical dystonia. *Headache* 2005;45:1073–1077.
41. Factor SA, Molho ES, Evans S, Feustel PJ. Efficacy and safety of repeated doses of botulinum toxin type B in type A resistant and responsive cervical dystonia. *Mov Disord* 2005;20:1152–1160.
42. Bihari K. Safety, effectiveness, and duration of effect of BOTOX after switching from Dysport for blepharospasm, cervical dystonia, and hemifacial spasm dystonia, and hemifacial spasm. *Curr Med Res Opin* 2005;21:433–438.
43. Dressler D, Bigalke H. Botulinum toxin type B de novo therapy of cervical dystonia: frequency of antibody induced therapy failure. *J Neurol* 2005;252:904–907.
44. Truong D, Duane DD, Jankovic J, et al. Efficacy and safety of botulinum type A toxin (Dysport) in cervical dystonia: results of the first US randomized, double-blind, placebo-controlled study. *Mov Disord* 2005;20:783–791.
45. Kiss ZH, Doig K, Eliasziw M, Ranawaya R, Suchowersky O. The Canadian multicenter trial of pallidal deep brain stimulation for cervical dystonia: preliminary results in three patients. *Neurosurg Focus* 2004;17:E5.
46. Eltahawy HA, Saint-Cyr J, Giladi N, Lang AE, Lozano AM. Primary dystonia is more responsive than secondary dystonia to pallidal interventions: outcome after pallidotomy or pallidal deep brain stimulation. *Neurosurgery* 2004;54:613–619.
47. Meyer CH. Outcome of selective peripheral denervation for cervical dystonia. *Stereotact Funct Neurosurg* 2001;77:44–47.

48. Naumann M, Yakovlev A, Durif F. A randomized, double-masked, crossover comparison of the efficacy and safety of botulinum toxin type A produced from the original bulk toxin source and current bulk toxin source for the treatment of cervical dystonia. *J Neurol* 2002;249:57–63.
49. Munchau A, Palmer JD, Dressler D, et al. Prospective study of selective peripheral denervation for botulinum-toxin resistant patients with cervical dystonia. *Brain* 2001;124:769–783.
50. Lucetti C, Nuti A, Gambaccini G, et al. Mexiletine in the treatment of torticollis and generalized dystonia. *Clin Neuropharmacol* 2000;23:186–189.
51. Brashear A, Lew MF, Dykstra DD, et al. Safety and efficacy of NeuroBloc (botulinum toxin type B) in type A-responsive cervical dystonia. *Neurology* 1999;53:1439–1446.
52. Brin MF, Lew MF, Adler CH, et al. Safety and efficacy of NeuroBloc (botulinum toxin type B) in type A-resistant cervical dystonia. *Neurology* 1999;53:1431–1438.
53. Krauss JK, Toups EG, Jankovic J, Grossman RG. Symptomatic and functional outcome of surgical treatment of cervical dystonia. *J Neurol Neurosurg Psychiatry* 1997;63:642–648.
54. Truong DD, Cullis PA, O'Brien CF, Koller M, Villegas TP, Wallace JD. BotB (botulinum toxin type B): evaluation of safety and tolerability in botulinum toxin type A-resistant cervical dystonia patients (preliminary study). *Mov Disord* 1997;12:772–775.
55. Lew MF, Adornato BT, Duane DD, et al. Botulinum toxin type B: a double-blind, placebo-controlled, safety and efficacy study in cervical dystonia. *Neurology* 1997;49:701–707.
56. Brans JW, Lindeboom R, Snoek JW, et al. Botulinum toxin versus trihexyphenidyl in cervical dystonia: a prospective, randomized, double-blind controlled trial. *Neurology* 1996;46:1066–1072.
57. Brans JW, Lindeboom R, Aramideh M, Speelman JD. Long-term effect of botulinum toxin on impairment and functional health in cervical dystonia. *Neurology* 1998;50:1461–1463.
58. Ranoux D, Gury C, Fondarai J, Mas JL, Zuber M. Respective potencies of Botox and Dysport: a double blind, randomised, crossover study in cervical dystonia. *J Neurol Neurosurg Psychiatry* 2002;72:459–462.
59. Tarsy D. Comparison of clinical rating scales in treatment of cervical dystonia with botulinum toxin. *Mov Disord* 1997;12:100–102.
60. Muller J, Wissel J, Kemmler G, et al. Craniocervical dystonia questionnaire (CDQ-24): development and validation of a disease-specific quality of life instrument. *J Neurol Neurosurg Psychiatry* 2004;75:749–753.
61. Tepavcevic DK, Svetel M, Pekmezovic T, Petrovic I, Kostic VS. Craniocervical Dystonia Questionnaire (CDQ-24): validation and cross-cultural adaptation in Serbian patients. *Coll Antropol* 2009;33:1185–1189.
62. Zetterberg L, Halvorsen K, Farnstrand C, Aquilonius SM, Lindmark B. Physiotherapy in cervical dystonia: six experimental single-case studies. *Physiother Theory Pract* 2008;24:275–290.
63. Jacobson BH, Johnson A, Grywalski C, et al. The Voice Handicap Index (VHI): development and validation. *Am J Speech Lang Pathol* 1997;6:66–70.
64. Woisard V, Bodin S, Puech M. [The Voice Handicap Index: impact of the translation in French on the validation]. Le "Voice Handicap Index": impact de la traduction française sur la validation. *Rev Laryngol Otol Rhinol (Bord)* 2004;125:307–312. [French].
65. Rosen CA, Murry T, Zinn A, Zullo T, Sonbolian M. Voice handicap index change following treatment of voice disorders. *J Voice* 2000;14:619–623.
66. Courey MS, Garrett CG, Billante CR, et al. Outcomes assessment following treatment of spasmodic dysphonia with botulinum toxin. *Ann Otol Rhinol Laryngol* 2000;109:819–822.
67. Spector BC, Netterville JL, Billante C, Clary J, Reinisch L, Smith TL. Quality-of-life assessment in patients with unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg* 2001;125:176–182.
68. Benninger MS, Gardner G, Grywalski C. Outcomes of botulinum toxin treatment for patients with spasmodic dysphonia. *Arch Otolaryngol Head Neck Surg* 2001;127:1083–1085.
69. Lee L, Daughton S, Scheer S, et al. Use of acupuncture for the treatment of adductor spasmodic dysphonia: a preliminary investigation. *J Voice* 2003;17:411–424.
70. Chhetri DK, Mendelsohn AH, Blumin JH, Berke GS. Long-term follow-up results of selective laryngeal adductor denervation-reinnervation surgery for adductor spasmodic dysphonia. *Laryngoscope* 2006;116:635–642.
71. Tsuji DH, Chrispim FS, Imamura R, Sennes LU, Hachiya A. Impact in vocal quality in partial myectomy and neurectomy endoscopic of thyroarytenoid muscle in patients with adductor spasmodic dysphonia. *Braz J Otorhinolaryngol* 2006;72:261–266.
72. Hussain A, Shakeel M. Selective lateral laser thyroarytenoid myotomy for adductor spasmodic dysphonia. *J Laryngol Otol* 2010;124:886–891.
73. Tsuji DH, Takahashi MT, Imamura R, Hachiya A, Sennes LU. Endoscopic laser thyroarytenoid myoneurectomy in patients with adductor spasmodic dysphonia: a pilot study on long-term outcome on voice quality. *J Voice* 2012;26:666–612.
74. Novakovic D, Waters HH, D'Elia JB, Blitzer A. Botulinum toxin treatment of adductor spasmodic dysphonia: longitudinal functional outcomes. *Laryngoscope* 2011;121:606–612.
75. Morzaria S, Damrose EJ. A comparison of the VHI, VHI-10, and V-RQOL for measuring the effect of botox therapy in adductor spasmodic dysphonia. *J Voice* 2012;26:378–380.
76. Carding PN, Horsley IA, Docherty GJ. A study of the effectiveness of voice therapy in the treatment of 45 patients with nonorganic dysphonia. *J Voice* 1999;13:72–104.
77. Webb AL, Carding PN, Deary IJ, MacKenzie K, Steen IN, Wilson JA. Optimising outcome assessment of voice interventions. I: Reliability and validity of three self-reported scales. *J Laryngol Otol* 2007;121:763–767.
78. Steen IN, MacKenzie K, Carding PN, Webb A, Deary IJ, Wilson JA. Optimising outcome assessment of voice interventions. II: Sensitivity to change of self-reported and observer-rated measures. *J Laryngol Otol* 2008;122:46–51.
79. MacKenzie K, Millar A, Wilson JA, Sellars C, Deary IJ. Is voice therapy an effective treatment for dysphonia? A randomised controlled trial. *BMJ* 2001;323:658–661.
80. Rattenbury HJ, Carding PN, Finn P. Evaluating the effectiveness and efficiency of voice therapy using transnasal flexible laryngoscopy: a randomized controlled trial. *J Voice* 2004;18:522–533.
81. Meek P, Carding PN, Howard DH, Lennard TW. Voice change following thyroid and parathyroid surgery. *J Voice* 2008;22:765–772.
82. Deary IJ, Webb A, MacKenzie K, Wilson JA, Carding PN. Short, self-report voice symptom scales: psychometric characteristics of the voice handicap index-10 and the vocal performance questionnaire. *Otolaryngol Head Neck Surg* 2004;131:232–235.
83. Burke RE, Fahn S, Marsden CD, Bressman SB, Moskowitz C, Friedman J. Validity and reliability of a rating scale for the primary torsion dystonias. *Neurology* 1985;35:73–77.
84. Kim JP, Chang WS, Chang JW. The long-term surgical outcomes of secondary hemidystonia associated with post-traumatic brain injury. *Acta Neurochir (Wien)* 2012;154:823–830.
85. Cif L, Gonzalez-Martinez V, Vasques X, et al. Staged implantation of multiple electrodes in the internal globus pallidus in the treatment of primary generalized dystonia. *J Neurosurg* 2012;116:1144–1152.
86. Panov F, Tagliati M, Ozelius LJ, et al. Pallidal deep brain stimulation for DYT6 dystonia. *J Neurol Neurosurg Psychiatry* 2012;83:182–187.
87. Sobstyl M, Zabek M, Dzierzecki S, Koziara H, Mossakowski Z. Chronic bilateral pallidal stimulation in patients with generalized primary dystonia - multi-contact cathodal stimulation is superior to bipolar stimulation mode. Preliminary results. *Neurol Neurochir Pol* 2011;45:252–259.
88. Schrader C, Capelle HH, Kinfe TM, et al. GPi-DBS may induce a hypokinetic gait disorder with freezing of gait in patients with dystonia. *Neurology* 2011;77:483–488.
89. Limotai N, Go C, Oyama G, et al. Mixed results for GPi-DBS in the treatment of cranio-facial and cranio-cervical dystonia symptoms. *J Neurol* 2011;258:2069–2074.
90. Valalik I, Jobbagy A, Bognar L, Csokay A. Effectiveness of unilateral pallidotomy for Meige syndrome confirmed by motion analysis. *Stereotact Funct Neurosurg* 2011;89:157–161.

91. Isaias IU, Volkmann J, Kupsch A, et al. Factors predicting protracted improvement after pallidal DBS for primary dystonia: the role of age and disease duration. *J Neurol* 2011;258:1469–1476.
92. Reese R, Gruber D, Schoenecker T, et al. Long-term clinical outcome in Meige syndrome treated with internal pallidum deep brain stimulation. *Mov Disord* 2011;26:691–698.
93. Azoulay-Zyss J, Roze E, Welter ML, et al. Bilateral deep brain stimulation of the pallidum for myoclonus-dystonia due to epsilon-sarcoglycan mutations: a pilot study. *Arch Neurol* 2011;68:94–98.
94. Sako W, Morigaki R, Mizobuchi Y, et al. Bilateral pallidal deep brain stimulation in primary Meige syndrome. *Parkinsonism Relat Disord* 2011;17:12–125.
95. Gruber D, Kuhn AA, Schoenecker T, et al. Pallidal and thalamic deep brain stimulation in myoclonus-dystonia. *Mov Disord* 2010;25:1733–1743.
96. Vidailhet M, Yelnik J, Lagrange C, et al. Bilateral pallidal deep brain stimulation for the treatment of patients with dystonia-choreoathetosis cerebral palsy: a prospective pilot study. *Lancet Neurol* 2009;8:709–717.
97. Gruber D, Trottenberg T, Kivi A, et al. Long-term effects of pallidal deep brain stimulation in tardive dystonia. *Neurology* 2009;73:53–58.
98. Sensi M, Cavallo MA, Quatrone R, et al. Pallidal stimulation for segmental dystonia: long term follow up of 11 consecutive patients. *Mov Disord* 2009;24:1829–1835.
99. Vasques X, Cif L, Gonzalez V, Nicholson C, Coubes P. Factors predicting improvement in primary generalized dystonia treated by pallidal deep brain stimulation. *Mov Disord* 2009;24:846–853.
100. Sako W, Goto S, Shimazu H, et al. Bilateral deep brain stimulation of the globus pallidus internus in tardive dystonia. *Mov Disord* 2008;23:1929–1931.
101. Hamasaki T, Yamada K, Kuratsu J. Hemidystonia secondary to thalamic hemorrhage treated with GPi stimulation. *Mov Disord* 2008;23:1762–1766.
102. Sobstyl M, Zabek M, Koziara H, Dzierzecki S. Chronic bilateral pallidal stimulation in a patient with DYT-1 positive primary generalized dystonia. A long-term follow-up study. *Neurol Neurochir Pol* 2008;42:50–54.
103. Houeto JL, Yelnik J, Bardinet E, et al. Acute deep-brain stimulation of the internal and external globus pallidus in primary dystonia: functional mapping of the pallidum. *Arch Neurol* 2007;64:1281–1286.
104. Alterman RL, Miravite J, Weisz D, Shils JL, Bressman SB, Tagliati M. Sixty Hertz pallidal deep brain stimulation for primary torsion dystonia. *Neurology* 2007;69:681–688.
105. Kleiner-Fisman G, Liang GS, Moberg PJ, et al. Subthalamic nucleus deep brain stimulation for severe idiopathic dystonia: impact on severity, neuropsychological status, and quality of life. *J Neurosurg* 2007;107:29–36.
106. Ostrem JL, Marks WJ Jr, Volz MM, Heath SL, Starr PA. Pallidal deep brain stimulation in patients with cranial-cervical dystonia (Meige syndrome). *Mov Disord* 2007;22:1885–1891.
107. Kupsch A, Benecke R, Muller J, et al. Pallidal deep-brain stimulation in primary generalized or segmental dystonia. *N Engl J Med* 2006;355:1978–1990.
108. Coubes P, Cif L, El Fertit H, et al. Electrical stimulation of the globus pallidus internus in patients with primary generalized dystonia: long-term results. *J Neurosurg* 2004;101:189–194.
109. Katayama Y, Fukaya C, Kobayashi K, Oshima H, Yamamoto T. Chronic stimulation of the globus pallidus internus for control of primary generalized dystonia. *Acta Neurochir Suppl* 2003;87:125–128.
110. Air EL, Ostrem JL, Sanger TD, Starr PA. Deep brain stimulation in children: experience and technical pearls. *J Neurosurg Pediatr* 2011;8:566–574.
111. Haridas A, Tagliati M, Osborn I, et al. Pallidal deep brain stimulation for primary dystonia in children. *Neurosurgery* 2011;68:738–743.
112. Timmermann L, Pauls KA, Wieland K, et al. Dystonia in neurodegeneration with brain iron accumulation: outcome of bilateral pallidal stimulation. *Brain* 2010;133:701–712.
113. Borggraefe I, Mehrkens JH, Telegravciska M, Berweck S, Botzel K, Heinen F. Bilateral pallidal stimulation in children and adolescents with primary generalized dystonia—report of six patients and literature-based analysis of predictive outcomes variables. *Brain Dev* 2010;32:223–228.
114. Borggraefe I, Boetzel K, Boehmer J, et al. Return to participation - significant improvement after bilateral pallidal stimulation in rapidly progressive DYT-1 dystonia. *Neuropediatrics* 2008;39:239–242.
115. Comella CL, Leurgans S, Wu J, Stebbins GT, Chmura T. Rating scales for dystonia: a multicenter assessment. *Mov Disord* 2003;18:303–312.

Dystonia rating scales: critique and recommendations

Supporting Information

Authors: Alberto Albanese, Francesca Del Sorbo, Cynthia Comella, H.A. Jinnah, Jonathan W. Mink, Bart Post, Marie Vidailhet, Jens Volkmann, Tom Warner, Albert F.G. Leentjens, Pablo Martinez-Martin, Glenn T. Stebbins, Christopher G. Goetz, Anette Schrag

Blepharospasm scales

Blepharospasm Disability Index (BSDI)

1.	Has the scale been used in dystonia?	Yes
2.	Have clinimetric properties of this scale been evaluated in dystonia?	Yes
Are there several versions of the scale? If so, which one has been assessed?		The BSDI is a derivative of an earlier scale called the Blepharospasm Disability Scale ¹ . The earlier scale was abandoned because questions confused patients.
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)		The BSDI is a self-rating scale (0–4 points) for assessment of impairment of specific activities of daily living caused by blepharospasm, ranging from “no impairment”, “slight=moderate=severe impairment” to “no longer possible due to my illness”; it also includes a “not applicable” option. The BSDI rates 6 domains that affect quality of life, such as “Driving a vehicle”, “Reading”, “Watching TV”, “Shopping”, “Walking” and “Doing everyday activities”. The scoring system is: mean item scores on the BSDI are calculated by dividing the sum score by the number of applicable items. The BSDI can be addressed in less than 2 minutes.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?		Patient-rated.

Copyright or public domain?	Copyrighted by Merz Pharma.
How can the scale be obtained (address or website)?	Published in ² .
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	Not clear. Available article online seem to indicate English-language only. However, it is claimed the scale has been used throughout Europe and Israel, so there may be other languages.

Scale properties

<i>Content</i>	
Face validity?	Good
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	It focuses on dystonic movements that affect vision, mostly blepharospasm.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Blepharospasm and apraxia of eyelid opening.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No
Is disability measured?	Yes
Is health-related quality of life measured?	No
Is tremor evaluated?	Tremor is not specifically evaluated.
Does it exclude an important aspect of dystonia? If so, which?	It focuses on disability related to sight. It is not a scale that can specifically measure dystonic movement abnormalities.
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No
Which areas are not covered?	Any areas that don't affect vision.

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Severity of disability
Is there a cut-off score for diagnosis of dystonia?	No
Is this appropriate in dystonia? If not, why?	Yes
Has it been <i>used</i> to measure severity or to screen for dystonia?	Severity of disability.

Acceptability

Length	Very good. It can be addressed in less than 2 minutes (personal judgment).
Clear instructions to raters?	Yes. In the prospective observational pilot study that used the newly developed BSDI, patients uniformly understood this scale well as no missing data or multiple answers occurred with the BSDI ² .
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No. The evaluation of the patients' responses to the BSDI showed that patients in a double-blind, randomized, active-controlled trial used the BSDI properly ² .
Appropriateness of questions/items for dystonia?	Good (personal judgment).
Applicability across dystonia stages: mild – moderate – severe?	Good (personal judgment).

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	None apparent (personal judgment).
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	The internal consistency of the BSDI was high (Cronbach's Alpha = 0.88); the retest reliability of the BSDI single items was adequate (Spearman's rank coefficient = $0.453 < r < 0.595$) ² .
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Good.

Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	No gold standard.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	In a development observational pilot study of the BSDI, the correlation coefficients of the weighted mean scores of the BSDI to the JRS sum scores were $r = 0.734$ (day 0) and 0.792 (day 21). At day 21, the correlation coefficients between the PEGR score and the BSDI weighted mean score ($r = -0.76$) and between the PEGR score and the JRS scores ($r = -0.87$) showed also statistically significant and moderately high interscale correlations ². In a double-blind, randomized, active-controlled study (300 patients with blepharospasm included), the correlation between JRS sum score and BSDI weighted mean score was $r = 0.487$ (baseline) and $r = 0.737$ (control visit), respectively ².
Valid in ethnically and culturally different populations?	Yes (personal judgment).
Valid in both genders and at all ages?	Applicable to both genders, but not to children because it contains items that involve writing, reading, shopping. Never tested in children.
Valid in patients with dementia or significant cognitive impairment?	No

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No
---	----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	The results of an observational study showed that the BSDI was sensitive to change after BoNT treatment (Jankovic et al, 2009). Data from the two randomized trials (Roggenkämper et al, 2006; Wabbels et al, 2010) designed to evaluate the effects of

	BoNTs for blepharospasm using BSDI scale showed that BSDI was sensitive to change but that BSDI may not be sensitive enough to detect differences between two BoNT products. In the randomized, non-inferiority trial (Roggenkämper et al, 2006) mean item scores on the BSDI were calculated at 21 and 109–112 days post-treatment. Both BoNTA products significantly reduced mean item scores on the BSDI versus baseline at the two follow-up time points. Mean item scores on the BSDI at baseline were compared in the two groups (1.67 BOTOX®, 1.60 Xeomin®), and mean improvements at 21 days (-0.83 BOTOX®, -0.82 Xeomin®) and 109–112 days (-0.22 BOTOX®, -0.36 Xeomin®) were not significantly different at the $P < 0.05$ level. In a randomized, controlled trial comparing Xeomin® to BOTOX® (Wabbels et al, 2010) both BoNTA products significantly reduced total BSDI scores versus baseline at 4 and 8 weeks. Mean total baseline scores in the two groups were compared (7.9 BOTOX®, 8.3 Xeomin®), and mean improvements in total BSDI scores at 4 weeks (-2.8 BOTOX®, -1.3 Xeomin®) and 8 weeks (-1.3 BOTOX®, -0.8 Xeomin®) were not significantly different at the $P < 0.05$ level. Mean item scores on the BSDI were also compared at baseline (1.39 BOTOX®, 1.44 Xeomin®), and improvements in mean item scores on the BSDI at 4 weeks (-0.42 BOTOX®, -0.21 Xeomin®) were not significantly different at the $P < 0.05$ level. In a recent randomized, placebo-controlled trial for treatment of blepharospasm, a highly significant treatment difference in favour of incobotulinumtoxinA was observed as measured by the BSDI, indicating meaningful improvement in functional impairment (BSDI improved by 0.5 points (95% CI [0.2–0.7]; $P = 0.002$) compared with placebo) (Jankovic J et al,
--	---

	2011).
Has the minimal clinically important change been assessed in patients with dystonia?	Yes

Overall impression

Advantages and disadvantages	Advantages of the BSDI: it is focused on daily activities and it is ease of use. The scoring system is also rather simple. A useful criticism of the scale raising concerns regarding poor sensitivity of the scale to mild disability or small changes is found in: Wabels et al, J Neural Transm, 2011. This criticism is relevant because the inability of the scale to detect differences between various BTX preparations leads to misleading conclusions that these different preparations have equivalent efficacy.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	This simple disability scale can be used for numerous purposes, including interventions. It focuses on disability related to sight. It is not a scale that can specifically measure dystonic movement abnormalities. It should be combined with a more specific scale that rates movement itself.

Final appraisal of the scale

Tick appropriate	Ranking
Yes	Recommended Requirements are fulfilled if: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Jankovic Rating Scale (JRS)

1.	Has the scale been used in dystonia?	X Yes ☐ No
2.	Have clinimetric properties of this scale been evaluated in dystonia?	X Yes ☐ No
	Are there several versions of the scale?	No

If so, which one has been assessed?	
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	Scale consists of 2 subscales: 1. Severity rated 0-4 2. Frequency rated 0-4 Maximal score: 8 Does not address time frame.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	The scale is combination of patient report and clinician rated. For some items in frequency, must ask patient whether occurring more than 50% of waking time.
Copyright or public domain?	Copyright held by Dr. Jankovic
How can the scale be obtained (address or website)?	Joseph Jankovic, Department of Neurology, Baylor college of Medicine, 6550 Fannin, Suite 1801, Houston, TX 77030. Email: josephj@bcm.edu
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No Not assessed for clinimetric properties in other languages, although may have been translated by Merz for their international studies on incobotulinumtoxinA.

Scale properties

Content

Face validity?	Face validity demonstrated
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Focused on eyes
Is the scale heavily weighted towards one type of dystonia? If so, which?	Blepharospasm
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	Severity, duration

Is disability measured?	Disability is included as part of items 2,3,4 of severity section.
Is health-related quality of life measured?	No
Is tremor evaluated?	NA
Does it exclude an important aspect of dystonia? If so, which?	Not clear if 10 point scale is sufficient (combination of the severity and frequency). Does not evaluate aspects such as increased blinking without external stimuli; Occasional spasms; the combination of increased blinking and spasms). Likewise, frequency does not seem to have equal intervals between scores; eg. goes from eyelid fluttering lasting less than 1 second to eye spasms lasting more than 1 second but eyes open more than 50% of waking time. In patients, this represents a large gap in severity.
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No
Which areas are not covered?	See above.

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Designed to measure severity of motor movements, not a diagnostic tool
Is there a cut-off score for diagnosis of dystonia?	NA
Is this appropriate in dystonia? If not, why?	NA
Has it been <i>used</i> to measure severity or to screen for dystonia?	Has been used to measure severity

Acceptability

Length	Very brief scale, too short
--------	-----------------------------

Clear instructions to raters?	Not clear how much patient input for each item
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	Rating anchors not equally spaced for Likert scale
Appropriateness of questions/items for dystonia?	Appropriate questions
Applicability across dystonia stages: mild – moderate – severe?	Not assessed. Give the large intervals between different measures, moderate patients may not be adequately rated.

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	No floor, ceiling effects although not systematically evaluated.
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Reliability: Kendall Tau only 0.37 at baseline. Following botulinum toxin, Kendall's tau = 0.863. Done by treating clinical rater blinded to drug v placebo. Internal consistency not given ² .
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Face validity demonstrated. Validity assessed using content validity assessing correlations with Global rating.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Criterion validity not established lacking gold standard.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Correlation with Global rating assessed. In the double-blind, randomized, active-controlled study (300 patients with blepharospasm included), the correlation between JRS sum score and BSDI weighted mean score was $r = 0.487$ (baseline) and $r = 0.737$ (control visit), respectively ² .
Valid in ethnically and culturally different populations?	Not tested
Valid in both genders and at all ages?	Not tested
Valid in patients with dementia or significant cognitive impairment?	Not tested

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	Not tested
---	------------

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Responsiveness analyzed and acceptable. Responsiveness assessed using least square means of JRS and Global scale.
Has the minimal clinically important change been assessed in patients with dystonia?	Mean change of -1.75 interpreted as clinically meaningful change in JRS.

<i>Overall impression</i>	The JRS combines historic, patient interview questions as well as motor scoring. This makes the inter-rater reliability falsely high because the patient response does not change.
Advantages and disadvantages	The scale meets some of the minimal analysis for reliability and validity. The lack of items and the concern of the measures for each item not addressed. It has been used as a standard assessment in pivotal clinical trials. There are reservations as this scale combines historic, patient interview questions as well as motor scoring. This makes the inter-rater reliability falsely high because the patient response does not change.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	

Final appraisal of the scale

Tick appropriate	Ranking
Yes: -there are data on its use in studies beyond the group that developed the scale	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Blepharospasm Disability Scale (BDS)

1. Has the scale been used in dystonia?	☒ Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	☒ Yes ☐ No
Are there several versions of the scale? If so, which one has been assessed?	No
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	It is an index of eight items section of the Blepharospasm Rating Scale ¹ . Depends on self-report and gives a single summary score, ranging from 0-26 points. This score is calculated into a % (0-100).
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient report.
Copyright or public domain?	Public.
How can the scale be obtained (address or website)?	Published in ³ .
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No

Scale properties

Content

Face validity?	Good
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Appropriate for blepharospasm
Is the scale heavily weighted towards one type of dystonia? If so, which?	Blepharospasm
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No
Is disability measured?	Sunglasses (0-2), Driving (0-6), Reading (0-3), television (0-3), movies (0-3), shopping (0-3), walking about (0-4), housework (0-3). In brackets the score range for each item.
Is health-related quality of life measured?	No
Is tremor evaluated?	No
Does it exclude an important aspect of dystonia? If so, which?	No
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No
Which areas are not covered?	

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Measures disability in dystonia
Is there a cut-off score for diagnosis of dystonia?	No
Is this appropriate in dystonia? If not, why?	Yes
Has it been <i>used</i> to measure severity or to screen for	Measures disability in dystonia

dystonia?	
-----------	--

Acceptability

Length	8 items. The BDS was easy to assess and could be completed in 4 min with minimal instructions ³ .
Clear instructions to raters?	Yes (personal judgment)
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No (personal judgment)
Appropriateness of questions/items for dystonia?	Good (personal judgment)
Applicability across dystonia stages: mild – moderate – severe?	Not known

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	No floor and ceiling effects mentioned in the publication.
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Developed by Fahn, no clinimetric methods are used ¹ . Validated in only one study by Lindeboom ³ clinical cohort and describing Cronbach's alpha (reliability), correlation coefficients (discriminant validity, convergent validity) and testing change over time (responsiveness). Cronbach's alpha of 0.69 for the total score (this moderate for a scale of 8 items); table 3 of Lindeboom ³ shows Cronbach's alpha and item-total correlation for each individual item. No test-retest and inter-rater reliability

Validity: assessed – not assessed, good – not good (references)

Face and/or content validity? Which areas are not covered?	Good
Criterion validity	No

(compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	a. Convergent validity: Significant correlations were found between BDS and the Blepharospasm Movement Scale (range, Spearman correlation coefficients, 0.58-0.76) and Severity Rating Scale scores (range, Spearman correlation coefficients, 0.65-0.79) ³. b. Evidence of discriminant validity was provided by the significant difference in median score on the BDS between newly admitted patients and patients already under treatment ($p < 0.001$).
Valid in ethnically and culturally different populations?	Not assessed.
Valid in both genders and at all ages?	Not assessed.
Valid in patients with dementia or significant cognitive impairment?	Not assessed.

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No
---	----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Responsiveness to health changes was demonstrated by a significant difference between median BDS scores before treatment and 2 weeks after treatment with botulinum toxin ($p < 0.01$) ³. See discriminant validity and figure 1 in article of Lindeboom ³.
Has the minimal clinically important change been assessed in patients with dystonia?	No MID/MCID.

Overall impression

Advantages and disadvantages	Rather crude measure for disability in blepharospasm, only one article ³ on the clinimetric properties of the scale. Lack of a
------------------------------	---

	“non-applicable” option for any of the individual items of the BDS. It is used in several studies on blepharospasm; see references.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Trials.

Final appraisal of the scale

Tick appropriate	Ranking
Yes: -there are data on its use in studies beyond the group that developed the scale	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Cervical dystonia scales

Cervical Dystonia Impact Scale (CDIP-58)

1. Has the scale been used in dystonia?	☒ Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	☒ Yes ☐ No
Are there several versions of the scale? If so, which one has been assessed?	No.
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	Cervical Dystonia Impact Profile (CDIP-58) is a disease-specific questionnaire that measures quality of life in patients with cervical dystonia. Fifty-eight items made up of eight 5-point scales, ranging from 4 to 10 items in length: three symptom scales (head and neck, pain and discomfort, sleep); two activity scales (upper limb activities, walking); and three psychosocial scales (annoyance, mood, psychosocial functioning). Eight summary scale scores are generated by summing items and then transformed to a 0-100 scale. High scores

	indicate worse health.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient-rated.
Copyright or public domain?	Public
How can the scale be obtained (address or website)?	The questionnaire is published in the original paper ⁴ .
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No

Scale properties

Content

Face validity?	Good ⁴ .
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Neck.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Cervical dystonia.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	There are just two activity scales (upper limb activities, walking).
Is disability measured?	Yes.
Is health-related quality of life measured?	Yes.
Is tremor evaluated?	No.
Does it exclude an important aspect of dystonia? If so, which?	No.
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	Yes, pain and discomfort.
Includes items of sleep disturbance as a result of	Yes.

dystonia?	
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	Yes.
Which areas are not covered?	No.

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Designed to measure the health impact of cervical dystonia.
Is there a cut-off score for diagnosis of dystonia?	No.
Is this appropriate in dystonia? If not, why?	NA.
Has it been <i>used</i> to measure severity or to screen for dystonia?	Used to measure the health impact of cervical dystonia and the impact of treatment in cervical dystonia ⁵⁻⁸ .

Acceptability

Length	Not reported.
Clear instructions to raters?	Yes (personal judgment).
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	Yes (personal judgment).
Appropriateness of questions/items for dystonia?	Yes (personal judgment).
Applicability across dystonia stages: mild – moderate – severe?	Authors that developed the CDIP-58 included in the study people with a wide range of disease severity ⁹ .

<i>Clinimetric properties</i>	New clinimetric techniques (Rasch analyses) revealed that the CDIP-58 performed well (e.g. good mean square infit (MNSQ) <1.5, even spread of item calibrations and hierarchical structure) ⁴ .
-------------------------------	--

Metric flaws (floor and ceiling effects, score distributions)?	Subscale scores spanned the entire scale range. However, the Walking scale fell just outside of the criterion (ceiling effect = 17%) and the Sleep subscale was found to have a more significant ceiling effect (27%). Despite this, responses were not notably skewed (-0.23 to +0.82). These findings indicate good scale-to-sample targeting, thus supporting total and subscale scores as appropriate for all patients representing the full spectrum of cervical dystonia impact ⁹ .
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Cronbach's alpha = 0.92–0.97; Item-total correlation = 0.64–0.91; Test-retest = 0.83–0.96 ⁹ .
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Good.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Not assessed.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Intercorrelations between CDIP-58 subscales ranged from 0.44 – 0.84, suggesting the subscales measured related but different constructs. A few of the correlations fell outside of the predicted range of correlations, and were highly correlated. However, the correlations were not unreasonable given the proximity of the constructs in each of the subscales and scale reliabilities were larger than inter-scale correlations supporting that CDIP-58 subscales measure distinct constructs ⁹ . Correlations between CDIP-58 subscales and hypothesised related scales of the Short Form-36 (SF-36), General Health Questionnaire (GHD) and Hospital Anxiety and Depression Scale (HAD) were consistent with predictions: Medical Outcomes Study SF-36 (r= 0.48–0.66); GHD

	($r = 0.30-0.66$); HAD ($r = 0.31-0.80$) ⁹ .
Valid in ethnically and culturally different populations?	Correlations between CDIP-58 subscales and socio-demographic variables (age, sex, and level of education attained) were in general low (-0.17 to $+0.06$). This finding suggests that responses to the CDIP-58 were not biased by socio-demographic factors ⁹ .
Valid in both genders and at all ages?	As above.
Valid in patients with dementia or significant cognitive impairment?	Not assessed or reported.

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No.
---	-----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	The CDIP-58 is more sensitive in detecting statistical and clinical change than comparable subscales of the SF-36, Functional Disability Questionnaire, and Pain and Activities of Daily Living subscales of the TWSTRS ⁵. Moderate to high effect sizes reveal the CDIP-58 is good at detecting the impact of botulinum toxin on all eight health dimensions ⁵.
Has the minimal clinically important change been assessed in patients with dystonia?	Not assessed or reported.

Overall impression

Advantages and disadvantages	<i>Advantages</i> The CDIP-58 is a disease-specific and validated questionnaire. It is more sensitive in detecting statistical and clinical change than comparable subscales. <i>Disadvantages</i>
------------------------------	---

	This scale has not been widely used as an outcome measure. Authors themselves ⁵ suggest further studies examining the responsiveness of the CDIP-58 as well as refinement of the walking subscale. Only recently works are starting to use CDIP-58 ⁶⁻⁸ .
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Use in clinical research, audit, and treatment trials.

Final appraisal of the scale

Tick appropriate	Ranking
Yes	Recommended Requirements are fulfilled if: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Toronto Western Spasmodic torticollis Rating Scale (TWSTRS)

1. Has the scale been used in dystonia?	X Yes ☐ No Developed by Consky and Lang ¹⁰ , published in a textbook ¹¹ . Frequent use to measure cervical dystonia, with most studies including pivotal clinical trials using as primary outcome either the Severity subscale or the total TWSTRS score.
2. Have clinimetric properties of this scale been evaluated in dystonia?	X Yes ☐ No The clinimetric properties that have been assessed include reliability and convergent validity. For the total score, reliability has also been assessed separately for the motor severity score. The subscales for pain and disability have not been separately assessed.
Are there several versions of the scale? If so, which one has been assessed?	Only one version.

Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	TWSTRS has 3 subscales: Maximal score on total TWSTRS is 85 a. Severity subscale to measure head movement, shoulder elevation and anterior displacement, geste, time in midline and duration. There are 11 items with maximal score =35 b. Disability subscale: 6 items; maximal score 30 c. Pain due to dystonia: 3 items including severity, duration and pain contribution to disability. Maximal score 20
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	The severity is clinician rated; disability and pain are patient rated. For severity, there is a training teaching tape for clinicians
Copyright or public domain?	No copyright on scale. MDS has copyright on training tape
How can the scale be obtained (address or website)?	Scale is available on WEMOVE.org. The training tape at MDS website at movementdisorders.org
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☐ No No official translations

Scale properties

Content

Face validity?	Yes
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Neck
Is the scale heavily weighted towards one type of dystonia? If so, which?	Cervical dystonia
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency,	Motor severity is measured, including severity, duration, frequency and amplitude

amplitude and position of movement, type of movement)?	and position of movement
Is disability measured?	Disability subscale of the TWSTRS measures 6 items.
Is health-related quality of life measured?	No
Is tremor evaluated?	No
Does it exclude an important aspect of dystonia? If so, which?	Includes movement severity, pain and disability
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	Pain scale is included
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No
Which areas are not covered?	See above; also does not measure tremor

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	It is used for a severity measure. Not a diagnostic tool
Is there a cut-off score for diagnosis of dystonia?	No
Is this appropriate in dystonia? If not, why?	NA
Has it been <i>used</i> to measure severity or to screen for dystonia?	Used for severity of dystonia.

Acceptability

Length	Appropriate
Clear instructions to raters?	For the most part, clear instructions. Duration factor is confusing because contains two elements, with scores of 2,3 having two possible definitions.
Ambiguities in instructions to patient/rater?	Some ambiguities in instructions. What defines midline; How long a period should the geste be effective; Should a geste be

Ambiguities in rating anchors?	suggested to the patient; over what time period is the patient to describe disability. The pain scale also includes a disability question that could overlap with the disability scale.
Appropriateness of questions/items for dystonia?	Appropriate
Applicability across dystonia stages: mild – moderate – severe?	Has been used for mild to severe.

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	No apparent floor and ceiling effects
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Has been assessed for internal consistency. Interrater agreement done and acceptable for motor score. A teaching tape of the TWSTRS scoring for the motor symptoms of cervical dystonia was developed ¹². The tape provides investigators with visual representations of each component of the motor section of the TWSTRS as agreed upon by three independently scoring raters. The rate of agreement for the non-dichotomous components was always significant, with a Kendall's coefficient of concordance W ranging between 0.98 and 0.76 ($p < 0.01$ for all measures). For the two dichotomous components, a weighted kappa coefficient was also significant at 0.86 for lateral shift and 0.89 for sagittal shift ($p < 0.01$ for both measures) ¹². TWSTRS scales have been shown to have acceptable inter-rater reliability (inter-class correlation coefficients, ICC, range 0.71-0.85) ¹¹. In detail: Severity scale (3 video raters-26 of 40 patients): ICC = 0.79; Severity scale change in score post Botox (3 video raters-26 of 40 patients): ICC = 0.85; Disability/Pain scales-pre Botox (4 rates-11 patients): ICC = 0.74; Severity scale (3 video raters vs. direct observer scores-20 patients): ICC = 0.71.
Validity: assessed – not assessed, good – not good (references)	

Face and/or content validity? Which areas are not covered?	Face and content validity evaluated for motor severity not pain or disability. However, one study found that the TWSTRS pain and disability measures were associated with global disease burden, more so than the motor severity measured by the Tsui scale ¹³.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Criterion validity not assessed (no real gold standard available)
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	In the original validation study ¹¹, the validity of TWSTRS was assessed by calculating of Pearson correlation coefficients comparing changes in severity scores from videotape raters and the direct observers with the patients' self-reported overall "percent improvement" and changes in disability and pain scores following treatment. The results of validity assessments have been as follow: - Change in total severity score of 3 videotape raters and patients perception overall improvement: $r_s = 0.68$; - Change in total severity score of direct observers and patients perception overall improvement: $r_s = 0.61$; - Change in total severity score of 3 videotape raters and change in disability and pain scores: $r_s = 0.65$. Tsui and TWSTRS score reduction rates correlated highly (Pearson correlation coefficients) with each other (0.57; $p < 0.0001$), and both Tsui (0.59; $p < 0.0001$) and TWSTRS (0.71; $p < 0.0001$) score reduction rates correlated with the global improvement scale. Tsui and TWSTRS severity score reduction rates, which are both objective severity scales, also correlated highly (0.69; $p < 0.0001$). TWSTRS severity showed a smaller mean percentage improvement than Tsui (22 vs. 41%) likely because several individual items in the TWSTRS severity subscale (effect of sensory tricks, shoulder

	displacement, and range of motion) remained unchanged in a large majority of patients despite significant TWSTRS severity improvement. TWSTRS pain showed weak or no correlation with either objective severity scale, probably because in 10 of 42 patients in whom TWSTRS pain was significantly reduced, there was no accompanying improvement in the Tsui scale or TWSTRS severity ¹⁴. Lindeboom and coll. ¹³ found only small correlations between the change in impairment (Tsui) scores and change in pain (TWSTRS-Pain), disability (TWSTRS-Disability), handicap, overall QoL (MOS-20), and global disease burden scores.
Valid in ethnically and culturally different populations?	Not tested adequately in different populations.
Valid in both genders and at all ages?	No DIF analysis for gender or age differences
Valid in patients with dementia or significant cognitive impairment?	Not tested in cognitively impaired.

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No correlations with biological markers
---	---

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Yes Mean Tsui, TWSTRS, and TWSTRS subscale scores all showed significant improvement ($p < 0.0001$) following BTX A as measured by mean score reduction rates ¹⁴. The TWSTRS subscales for motor severity, disability and pain were improved from baseline in both groups treated with different doses of incobotulinumtoxinA vs. placebo at Weeks 4, 8 and final visits. The TWSTRS-Total score was improved from baseline in the active groups vs. placebo at

	Week 8 and at the final visit ¹⁵ .
Has the minimal clinically important change been assessed in patients with dystonia?	MCIC not established

Overall impression

Advantages and disadvantages	The TWSTRS scale is the most utilized rating scale for cervical dystonia. It evaluates severity, disability and pain. The clinimetric properties of severity have been described only for the motor severity.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Not suitable for diagnosis, screening, prevalence aetiological. Is appropriate for treatment trials.

Final appraisal of the scale

Tick appropriate	Ranking
Yes	Recommended Requirements are fulfilled if: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Functional Disability Questionnaire (FDQ)

1. Has the scale been used in dystonia?	☒ Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	☒ Yes ☐ No
Are there several versions of the scale? If so, which one has been assessed?	No.
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	The FDQ measures disability in daily functioning due to dystonia. There are 27 items, forming four factors: social disability, physical activity, self-care, and leisure activities. Each item is rated on a 4-point scale, ranging from 1 ("not affected") to 4 ("severely affected"); a "not applicable" response category (scored 0) is also included. Sub-scores for each factor are

	derived by summation of raw scores. High scores indicate worse health.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient-rated.
Copyright or public domain?	Public domain in the original paper.
How can the scale be obtained (address or website)?	Original paper: Jahanshahi and Marsden 1990 ¹⁶ .
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No

Scale properties

Content

Face validity?	Not assessed.
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	The FDQ was originally designed for patients with cervical problems (cervical dystonia and cervical spondylosis) ¹⁶ and then applied to patients with focal, segmental, and generalized dystonia ^{17,18} .
Is the scale heavily weighted towards one type of dystonia? If so, which?	No.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No.
Is disability measured?	Yes.
Is health-related quality of life measured?	No.
Is tremor evaluated?	No.
Does it exclude an important aspect of dystonia? If so, which?	No.

Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No.
Includes items of sleep disturbance as a result of dystonia?	No.
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No.
Which areas are not covered?	The FDQ comprises only a disability scale.

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	The FDQ measures severity of disability ¹⁶ .
Is there a cut-off score for diagnosis of dystonia?	No.
Is this appropriate in dystonia? If not, why?	NA.
Has it been <i>used</i> to measure severity or to screen for dystonia?	The FDQ has been used in studies to assess disability in daily functioning due to dystonia ¹⁶⁻¹⁸ . It hasn't been used to screen for dystonia.

Acceptability

Length	Acceptable, it takes less than 5 minutes to complete (personal judgment).
Clear instructions to raters?	Yes (personal judgment).
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No (personal judgment).
Appropriateness of questions/items for dystonia?	Yes (personal judgment).
Applicability across dystonia stages: mild – moderate – severe?	Yes (personal judgment).

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Not assessed.
--	---------------

Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	The scale has been shown to have good internal consistency and test-retest reliability ¹⁶: • Cronbach's alpha = 0.92 (27 items); • Test-retest (over a 1 month interval) = 0.93, $p < 0.001$.
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Not assessed.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	No.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Construct validity was examined using principal components analysis (PCA). PCA produced four-factor structure. The four factors accounted for 63% of the variance, and were respectively labelled “social disability”, “physical activity”, “self-care”, and “leisure activities” ¹⁶ .
Valid in ethnically and culturally different populations?	Not assessed.
Valid in both genders and at all ages?	Not assessed.
Valid in patients with dementia or significant cognitive impairment?	Not assessed.

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No.
---	-----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Responsiveness has not been well established, although data are available on responsiveness to BoNT ¹⁹ and change over time ¹⁶ . In a recent study, subscales of the FDQ were found less sensitive than CDIP-58 subscales in detecting changes after

	BoNT/A treatment ⁵ .
Has the minimal clinically important change been assessed in patients with dystonia?	Not assessed.

Overall impression

Advantages and disadvantages	<i>Advantages:</i> - Simple and brief questionnaire - Useful in community-based studies ^{17,18}. <i>Disadvantages:</i> - Not applicable in early ages because it contains items that involve activities such as going to parties, driving etc. - Assessment of the FDQ is not specific for one type of dystonia.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Suitable for use in clinical research, audit, and treatment trials.

Final appraisal of the scale

Tick appropriate	Ranking
Yes - there are data on its use in studies beyond the group that developed the scale	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Tsui scale

1. Has the scale been used in dystonia?	X Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	X Yes ☐ No
Are there several versions of the scale?	Modified Tsui Scale ²⁰ .

If so, which one has been assessed?	
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	The Tsui Scale was the first rating scale specifically designed for cervical dystonia ²¹. Patients were video-taped and a blind assessment was made by a physician unaware of the injection. It measures movement and tremor in cervical dystonia. It contains six items and is designed for video assessment. This scale evaluates the amplitude and duration of neck involuntary movements in the neck, elevation of shoulder, and head tremor.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Clinician-rated
Copyright or public domain?	Public
How can the scale be obtained (address or website)?	In the original paper ²¹ .
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No

Scale properties

Content

Face validity?	Not assessed.
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Cervical.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Cervical dystonia.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	Yes.

Is disability measured?	No.
Is health-related quality of life measured?	No.
Is tremor evaluated?	Yes.
Does it exclude an important aspect of dystonia? If so, which?	Pain.
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No.
Includes items of sleep disturbance as a result of dystonia?	No.
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No.
Which areas are not covered?	Pain.

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Designed to measure severity.
Is there a cut-off score for diagnosis of dystonia?	No.
Is this appropriate in dystonia? If not, why?	NA
Has it been <i>used</i> to measure severity or to screen for dystonia?	Used to measure severity.

Acceptability

Length	Acceptable.
Clear instructions to raters?	Yes (personal judgment).
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No (personal judgment).
Appropriateness of questions/items for dystonia?	Yes (personal judgment).
Applicability across dystonia stages: mild – moderate – severe?	Yes (personal judgment).

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Not assessed.
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	The scale has been shown to have acceptable inter-rater reliability ($r > 0.85$) ²¹⁻²³ .
Validity: assessed – not assessed, good – not good (references):	
Face and/or content validity? Which areas are not covered?	Not assessed.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Not assessed.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	It has been shown that there is a good correlation between the scores obtained with the TWSTRS and the Tsui scale. Pearson's coef: 0.57 ¹⁴ .
Valid in ethnically and culturally different populations?	Not assessed.
Valid in both genders and at all ages?	Not assessed.
Valid in patients with dementia or significant cognitive impairment?	Not assessed.

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No.
---	-----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Not assessed.
Has the minimal clinically important change been assessed in patients with dystonia?	Not assessed.

Overall impression

Advantages and disadvantages	
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Most used scale for BTX Injections and DBS: Toronto Western Spasmodic Torticollis Rating Scale (TWSTRS). The Tsui Scale has been used by researchers beyond original developers mainly for BTX injections, but not recently and it appears that the TWSTRS takes over.

Final appraisal of the scale

Tick appropriate	Ranking
Yes: -there are data on its use in studies beyond the group that developed the scale	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Modified Tsui scale

1. Has the scale been used in dystonia?	☒ Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	☐ Yes ☒ No
Are there several versions of the scale? If so, which one has been assessed?	Yes, there is the original Tsui Scale ²¹ .
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	The Modified Tsui scale is an adaptation of the Tsui scale ²¹ . To increase the sensitivity of the scale the ratings for postural head deviation were performed for four different modes of postural activation (sitting, lying, standing, walking) at each visit.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Clinician-rated
Copyright or public domain?	Public
How can the scale be obtained (address or website)?	In the original paper ²⁰ .

Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No
---	---

Scale properties

Content

Face validity?	Not assessed.
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Cervical.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Cervical dystonia.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	Yes.
Is disability measured?	No
Is health-related quality of life measured?	No
Is tremor evaluated?	Yes
Does it exclude an important aspect of dystonia? If so, which?	Pain
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No
Which areas are not covered?	Pain

Use

Was it <i>designed</i> to measure severity or to screen for	Designed to measure severity.
---	-------------------------------

dystonia?	
Is there a cut-off score for diagnosis of dystonia?	No.
Is this appropriate in dystonia? If not, why?	NA
Has it been <i>used</i> to measure severity or to screen for dystonia?	Used to measure severity.

Acceptability

Length	Acceptable.
Clear instructions to raters?	More precise but more complex compared to the Tsui Scale infect the ratings for postural head deviation are performed for four different modes of postural activation (sitting, lying, standing, walking) (personal judgment).
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No (personal judgment).
Appropriateness of questions/items for dystonia?	Yes (personal judgment).
Applicability across dystonia stages: mild – moderate – severe?	Yes (personal judgment).

Clinimetric properties

Not assessed

Metric flaws (floor and ceiling effects, score distributions)?	
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	
Validity: assessed – not assessed, good – not good (references):	
Face and/or content validity? Which areas are not covered?	
Criterion validity	

(compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	
Valid in ethnically and culturally different populations?	
Valid in both genders and at all ages?	
Valid in patients with dementia or significant cognitive impairment?	

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	
---	--

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	
Has the minimal clinically important change been assessed in patients with dystonia?	

Overall impression

Advantages and disadvantages	Positive points: More precise Takes into account the patient's and the doctor's impressions Negative points: Four different modes of postural activation (sitting, lying, standing, walking): more precise but more complex. Not clinimetrically evaluated. Not widely used.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Treatment trials.

Final appraisal of the scale

Tick appropriate	Ranking
Yes: -it has been applied to dystonia populations	Listed Fulfilled if one of the following applies: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Freiberg Questionnaire for Dystonia (FQD)

1. Has the scale been used in dystonia?	☒ Yes ☐ No The FQD has been developed by Scheidt and colleagues ²⁴ .
2. Have clinimetric properties of this scale been evaluated in dystonia?	☒ Yes ☐ No
Are there several versions of the scale? If so, which one has been assessed?	To our knowledge, there is a torticollis version ²⁴ and a blepharospasm version ²⁵ .
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	The FQD is a multi-item questionnaire consists of open-ended, dichotomous, and 5-point response options. Items refer to demographic variables, medical history, previous treatments, neurologic symptoms, and psychosocial changes caused by the illness. The psychosocial changes are evaluated with 23 items on a five-point scale ranging from 0 = "none" to 4 = "very severe; the psychosocial changes section covers five areas (profession, everyday life, social life, family, and psychological well-being). In the FQD torticollis version patients assess their neurologic symptoms (e.g., deviation of the head, muscle tension and head tremor) on a five-point scale ranging from 0 = "not at all" to 4 = "severely." In the FQD blepharospasm version the neurologic symptoms (e.g., the involuntary closure of the eyes) are rated on a five-point scale ranging from 0 = "not at all" to 4 = "severely." The visual complaints were rated on a five-point scale ranging from 0 = "no complaints" to 4 = "blind."

Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient-rated.
Copyright or public domain?	Not available.
How can the scale be obtained (address or website)?	The questionnaire is not published, even in the original papers ²⁴ .
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No

Scale properties

Content

Face validity?	Not assessed.
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Neck in the torticollis version and eyes in the blepharospasm version.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Cervical dystonia in the torticollis version and blepharospasm in the blepharospasm version.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	Yes. In the FQD torticollis version the patients assess their neurologic symptoms such as deviation of the head, muscle tension and head tremor. In the FQD blepharospasm version the patients assess their neurologic symptoms such as the involuntary closure of the eyes.
Is disability measured?	Yes.
Is health-related quality of life measured?	Yes.
Is tremor evaluated?	Yes, head tremor ²⁴ .
Does it exclude an important aspect of dystonia? If so, which?	Not known.
Includes items of sensory abnormalities (e.g., pain) as a	Yes, pain ²⁴ .

result of dystonia?	
Includes items of sleep disturbance as a result of dystonia?	Not known.
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	Not known.
Which areas are not covered?	A personal judgment is not possible as the scale is not available.

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Questionnaire designed to measure behavioural aspects in subjects with dystonia ²⁴ .
Is there a cut-off score for diagnosis of dystonia?	No.
Is this appropriate in dystonia? If not, why?	NA.
Has it been <i>used</i> to measure severity or to screen for dystonia?	Used to measure psychosocial changes in subjects with dystonia ²⁴ .

Acceptability

Length	Not reported. A personal judgment is not possible as the scale is not available.
Clear instructions to raters?	Not reported or assessed. A personal judgment is not possible as the scale is not available.
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	Not reported. A personal judgment is not possible as the scale is not available.
Appropriateness of questions/items for dystonia?	Not reported. A personal judgment is not possible as the scale is not available.
Applicability across dystonia stages: mild – moderate – severe?	Not reported. A personal judgment is not possible as the scale is not available.

Clinimetric properties

Reliability and validity have been presented on the psychosocial changes

subscale.

Metric flaws (floor and ceiling effects, score distributions)?	Data not found.
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	This scale is shown to have acceptable internal consistency, with Cronbach's alpha > 0.87 (two factors).
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Data not found.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Data not found.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Construct validity supported by factor (principal components) analysis. Two-factor model was adopted, explaining 59.4% of the variance ²⁴ : 1. On the first factor, labelled “illness-related psychological distress”, all items referring to psychological distress due to the illness loaded high (range 0.56-0.83); 2. On the second factor, labelled “functional disability”, items referring to functional disability loaded high (range 0.55-0.81).
Valid in ethnically and culturally different populations?	Not assessed.
Valid in both genders and at all ages?	Not assessed.
Valid in patients with dementia or significant cognitive impairment?	Not assessed.

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No.
---	-----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Not assessed.
Has the minimal clinically important change been assessed in patients with dystonia?	Not assessed.

Overall impression

Advantages and disadvantages	FQD torticollis version includes evaluation of dystonic movements, tremor and pain. Not easy to make other comments as the scale is not available.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Suitable: Treatment trial and clinical assessment. Unsuitable: Screening or diagnosis.

Final appraisal of the scale

Tick appropriate	Ranking
Yes: -it has been applied to dystonia populations	Listed Fulfilled if one of the following applies: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Disability questionnaire for patients with cervical dystonia

1. Has the scale been used in dystonia?	☒ Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	☐ Yes ☒ No
Are there several versions of the scale? If so, which one has been assessed?	No.
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	It is a 16-item questionnaire investigating impairment of activities of daily living due to cervical dystonia during the previous two-week period. Each item is scored on a five-point scale (ranging from 4 = normal performance to 0 = maximum disability). The maximum possible score of the

	disability questionnaire is 64 points.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient-rated.
Copyright or public domain?	Public in the original paper.
How can the scale be obtained (address or website)?	Available in the original publication ²⁶ .
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☒ Yes ☐ No The original questionnaire is in Italian language and has been translated in English. No data are available regarding the translation methodology. Nor the original nor the translated versions have been clinimetrically tested.

Scale properties

Content

Face validity?	Not assessed.
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	The scale is focused on one body area: cervical. One item is indefinite as it investigates difficulty in eating.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Yes. Questionnaire designed for cervical dystonia.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No.
Is disability measured?	Yes.
Is health-related quality of life measured?	No, except one item.
Is tremor evaluated?	Yes.
Does it exclude an important aspect of dystonia? If so, which?	Pain due to dystonia.
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No.

Includes items of sleep disturbance as a result of dystonia?	No.
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No.
Which areas are not covered?	This disability questionnaire do not addresses disability due to pain.

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	No, it is a disability questionnaire.
Is there a cut-off score for diagnosis of dystonia?	No.
Is this appropriate in dystonia? If not, why?	NA.
Has it been <i>used</i> to measure severity or to screen for dystonia?	No. Used as disability questionnaire in one study ²⁶ .

Acceptability

Length	Its length is acceptable: it takes about 5 minutes to be completed (personal judgment).
Clear instructions to raters?	Yes (personal judgment). No evidence found.
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No (personal judgment). No evidence found.
Appropriateness of questions/items for dystonia?	Yes (personal judgment). No evidence found.
Applicability across dystonia stages: mild – moderate – severe?	Yes (personal judgment). No evidence found.

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Not assessed.
--	---------------

Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Not assessed.
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Not assessed.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Not assessed.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Not assessed.
Valid in ethnically and culturally different populations?	Not assessed.
Valid in both genders and at all ages?	Not assessed.
Valid in patients with dementia or significant cognitive impairment?	Not assessed.

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No.
---	-----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Not assessed.
Has the minimal clinically important change been assessed in patients with dystonia?	Not assessed.

Overall impression

Advantages and disadvantages	<i>Advantages:</i> Specifically designed for cervical dystonia; very simple and brief disability questionnaire; can be appropriate for postal surveys as it doesn't need specific
------------------------------	---

	training. <i>Disadvantages:</i> This is not a clinimetrically tested scale, has not been validated up to now. Somatic symptoms of dystonia (pain) are not covered.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	This questionnaire designed to measure disability caused by cervical dystonia may be suitable in treatment trial of dystonia medication (personal judgment).

Final appraisal of the scale

Tick appropriate	Ranking
Yes: -it has been applied to dystonia populations	Listed Fulfilled if one of the following applies: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Body Concept Scale (BCS)

1. Has the scale been used in dystonia?	X Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	X Yes ☐ No
Are there several versions of the scale? If so, which one has been assessed?	No
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	22 semantic differential scales forming 4 factors: speed/strength, postural/movement-related, evaluative/esthetic, tension. Each item is rated on a 7-interval semantic scale. Subscores are derived for each factor by summation of raw scores. High scores indicate more negative body concept.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient
Copyright or public domain?	Public

How can the scale be obtained (address or website)?	From original publication ¹⁶
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes x No

Scale properties

Content

Face validity?	Yes – construct validity supported by moderate correlations of the scale and disfigurement ratings and Beck Depression Inventory (BDI).
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Neck
Is the scale heavily weighted towards one type of dystonia? If so, which?	Neck
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	no
Is disability measured?	no
Is health-related quality of life measured?	no
Is tremor evaluated?	no
Does it exclude an important aspect of dystonia? If so, which?	
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	no
Includes items of sleep disturbance as a result of dystonia?	no
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	Not specifically although scores of BCS may reflect these
Which areas are not covered?	

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	no
Is there a cut-off score for diagnosis of dystonia?	no
Is this appropriate in dystonia? If not, why?	yes
Has it been <i>used</i> to measure severity or to screen for dystonia?	no

Acceptability

Length	yes
Clear instructions to raters?	yes
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	no
Appropriateness of questions/items for dystonia?	yes
Applicability across dystonia stages: mild – moderate – severe?	yes

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Internal consistency (Cronbach's alpha = 0.95), test retest (ICC=0.87) ¹⁶ .
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	No
Criterion validity (compared to gold-standard (Is there a gold	No

standard?) or other criteria for diagnosis)	
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Construct validity is supported by moderate correlations of the scale and disfigurement ratings and the Beck Depression Inventory (0.44 and 0.41, respectively) ¹⁶ .
Valid in ethnically and culturally different populations?	yes
Valid in both genders and at all ages?	yes
Valid in patients with dementia or significant cognitive impairment?	no

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	Yes – BDI and disfigurement
---	-----------------------------

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Tested before and after BTA injection – non-significant improvement found. In patient with cervical dystonia tested before and after surgery (selective peripheral denervation) there was a significant improvement of body concept: compared with before surgery, the patients had a significant less negative body concept ($p= 0.042$) ²⁷ .
Has the minimal clinically important change been assessed in patients with dystonia?	no

Overall impression

Advantages and disadvantages	
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Assessing psychosocial features of torticollis

Final appraisal of the scale

Tick appropriate	Ranking
Yes: - it has been studied clinimetrically and found valid, reliable and sensitive to change	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Ways of coping checklist

1. Has the scale been used in dystonia?	☒ Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	☒ Yes ☐ No
Are there several versions of the scale? If so, which one has been assessed?	No
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	The Ways of Coping Checklist measures coping in cervical dystonia (CD) and was adapted from an original scale devised by Folkman and Lazarus ²⁸ and revised by Felton et al. ²⁹ for use with chronically ill populations. Composed of 64 items, forming 6 factors: wish-fulfilling fantasy, threat minimising, negative reappraisal/coping, cognitive restructuring, religious faith, instrumental coping. Subscores for each factor are derived from summation of raw scores. High scores indicate poorer coping.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient
Copyright or public domain?	Available from researcher
How can the scale be obtained (address or website)?	

Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes x No
---	---

Scale properties

Content

Face validity?	
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Torticollis
Is the scale heavily weighted towards one type of dystonia? If so, which?	Torticollis
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	no
Is disability measured?	no
Is health-related quality of life measured?	no
Is tremor evaluated?	no
Does it exclude an important aspect of dystonia? If so, which?	
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	no
Includes items of sleep disturbance as a result of dystonia?	no
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	no
Which areas are not covered?	

Use

Was it <i>designed</i> to measure severity or to screen for	No – adapted from previous scale not
---	--------------------------------------

dystonia?	designed for dystonia
Is there a cut-off score for diagnosis of dystonia?	no
Is this appropriate in dystonia? If not, why?	
Has it been <i>used</i> to measure severity or to screen for dystonia?	no

Acceptability

Length	lengthy
Clear instructions to raters?	yes
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	no
Appropriateness of questions/items for dystonia?	reasonable
Applicability across dystonia stages: mild – moderate – severe?	yes

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Cronbachs alpha=0.7 for internal consistency
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	no
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	no
Construct validity (correlations with other convergent scales and	Supported by principal components analysis

divergent scales (which?); factor analysis; known-groups comparisons).	
Valid in ethnically and culturally different populations?	yes
Valid in both genders and at all ages?	yes
Valid in patients with dementia or significant cognitive impairment?	no

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	no
---	----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	no
Has the minimal clinically important change been assessed in patients with dystonia?	no

Overall impression

Advantages and disadvantages	
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Not great – only reported use once and not freely available

Final appraisal of the scale

Tick appropriate	Ranking
Yes - it has been applied to dystonia populations	Listed Fulfilled if one of the following applies: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Blepharospasm/cervical dystonia scales

Craniocervical Dystonia Questionnaire (CDQ-24)

1.	Has the scale been used in dystonia?	☒ Yes ☐ No
2.	Have clinimetric properties of this scale been evaluated in dystonia?	☒ Yes ☐ No
Are there several versions of the scale? If so, which one has been assessed?		No.
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests))		The CDQ-24 is a disease-specific health-related quality of life (HRQOL) measure for craniocervical dystonia The CDQ-24 measures the impact of craniocervical dystonia on five HRQOL domains. It was developed for use in clinical research. It is composed of 24 items, forming five subscales: stigma, emotional well-being, pain, activities of daily living, and social/family life. Items are rated on a 5-point scale (18 items based on frequency – never to always; 6 items based on severity – not at all to very severely). High scores indicate worse health.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?		Patient-rated.
Copyright or public domain?		Public in the original paper.
How can the scale be obtained (address or website)?		The English version of the CDQ-24 is published together with the original paper ³⁰ .

Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☒ Yes ☐ No - Although only the original German version was validated, an exact translation into English including back-translation was performed to ensure accuracy ³⁰. - Serbian language: Tepavcević DK et al. ³¹ published the validation of the translated and culturally adapted CDQ-24 questionnaire on a group of Serbian patients (100 patients with idiopathic cervical dystonia and blepharospasm). The linguistic validation of CDQ-24 involved 3 steps, according to an internationally accepted methodology.
---	--

Scale properties

Content

Face validity?	Good (personal judgment).
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Face, head, neck.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Cranio-cervical dystonia, including cervical dystonia (CD) and blepharospasm (BSP).
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No.
Is disability measured?	Yes.
Is health-related quality of life measured?	Yes.
Is tremor evaluated?	No.
Does it exclude an important aspect of dystonia? If so, which?	No (personal judgment).
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	Yes (pain/a burning sensation).
Includes items of sleep disturbance as a result of dystonia?	Yes (sleep disturbance as a result of pain or a pulling sensation).
Includes items of neuropsychiatric features (e.g., anxiety,	Yes (anxiety, depression, sadness).

depression) as a result of dystonia?	
Which areas are not covered?	-

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	The CDQ-24 was developed as a disease specific QOL questionnaire that addresses the perceptions and concerns of patients with craniocervical dystonia including CD and BSP.
Is there a cut-off score for diagnosis of dystonia?	No.
Is this appropriate in dystonia? If not, why?	NA.
Has it been <i>used</i> to measure severity or to screen for dystonia?	Used to measure the impact of dystonia on quality of life ⁶ . Used as HRQOL measure to assess responsiveness to treatment induced changes ^{30,32} .

Acceptability

Length	Its length is acceptable: it takes about 5 minutes to be completed (personal judgment).
Clear instructions to raters?	Yes (personal judgment).
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No (personal judgment).
Appropriateness of questions/items for dystonia?	Yes (personal judgment).
Applicability across dystonia stages: mild – moderate – severe?	Yes (personal judgment).

Clinimetric properties

Müller's study ³⁰ shows that the CDQ-24 fulfils the essential criteria for health measurement rating scales defined by the SAC (Scientific Advisory Group of the Medical Outcomes Trust), including reliability, validity, and responsiveness to

therapy.

Metric flaws (floor and ceiling effects, score distributions)?	There were no relevant ceiling effects and a considerable floor effect was only observed in one domain (social/family life) ³⁰ .
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	The CDQ-24 also showed good reliability properties. For both reliability measures used, internal consistency and test-retest reliability, the criteria for good scale performance were satisfied by all subscales of the CDQ-24 (internal consistency, Cronbach's alpha ranging from 0.77 to 0.89; test-retest reliability, ICC > 0.9) ³⁰ .
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Not assessed.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Not assessed.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Validity was assessed by checking convergent and discriminant validity as well as the dimensional structure of the CDQ-24, was shown to be good or satisfactory in all respects³⁰. <i>Convergent validity:</i> was assessed in a large sample of 231 consecutive patients with CD and BSP by means of correlations of subscales of the CDQ-24 with those subscales of the SF-36 that measure related experiences. Analysis for convergent validity showed moderate to high correlations between those CDQ-24 and SF-36 subscales measuring similar aspects (Pearson's correlation $r = 0.50-0.73$; $p < 0.001$, each)³⁰. <i>Discriminant validity:</i> Patients with CD showed low correlations of the CDQ-24 subscores with the Tsui-score, attaining statistical significance for "ADL" ($r=0.29$, $p=0.01$), "social/family life" ($r=0.19$,

	p=0.04), “emotional well-being” (r=0.18, p=0.04). Correlations of CDQ-4 subscores with pain ratings were higher and reached statistical significance for all CDQ-24 subscales (p,0.05, each) and for the CDQ-24 total score (r=0.48, p,0.001) in CD. In patients with BSP, correlations of the CDQ-24 subscores with the clinical global impression score were also low, attaining statistical significance for two of the five CDQ-24 subscales (ADL r=0.31, p=0.003; emotional wellbeing r=0.21, p=0.047) and the CDQ-24 total score (r=0.25, p=0.019). <i>Dimensional structure of the CDQ-24:</i> A factor analysis performed on all patients who completed the CDQ-24 gave rise to a five factor model according to the eigenvalue criterion (cut-off value 1); this model explained 67.9% of the total variance. With only two exceptions (two items of the ADL subscale) the allocation of all 24 items to the factors—as suggested by their loadings—coincided with the previously defined domain structure of the instrument. The single two items with suboptimal assignment to the ADL subscale showed fairly high side loadings (r=0.38 and 0.39) on the factor ADL. Overall, a confirmation of the hypothesised domain structure of the instrument was obtained.
Valid in ethnically and culturally different populations?	Not assessed.
Valid in both genders and at all ages?	Yes for genders, not sure for ages as there are questions strictly related to adult ages (personal judgment).
Valid in patients with dementia or significant cognitive impairment?	No (personal judgment).
Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No.

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Sensitivity to change was analysed in 51 previously untreated (de novo) patients four weeks and one year following the first botulinum toxin treatment. Sensitivity to change was statistically confirmed by highly significant improvements of all CDQ-24 subscores in the de novo patients from baseline to four week follow up. One year follow up data revealed a stable improvement ³⁰ .
Has the minimal clinically important change been assessed in patients with dystonia?	No.

Overall impression

Advantages and disadvantages	The CDQ-24 is a brief and easy instrument. The CDQ-24 can be used to evaluate the impact of the disease on areas not covered by generic measures, such as SF-36, which are of considerable concern to patients with craniocervical dystonia. The CDQ-24 evaluates also pain, sleep and depression due to dystonia.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	It is a disease-specific patient-based assessment. The developers of this scale have presented data which indicate that it may be more sensitive than generic HRQOL measures in detecting change over time during botulinum toxin treatment ³⁰, and it may turn out to be valuable for that purpose. Therefore, its use is suitable in clinical trials as well as in daily clinical practice. Unsuitable for screening/diagnosis.

Final appraisal of the scale

Tick appropriate	Ranking
Yes	Recommended Requirements are fulfilled if: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to

Oromandibular dystonia scales

Oromandibular dystonia questionnaire (OMDQ-25)

1.	Has the scale been used in dystonia?	■ Yes □ No
2.	Have clinimetric properties of this scale been evaluated in dystonia?	■ Yes □ No
Are there several versions of the scale? If so, which one has been assessed?		no
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)		Patient administered questionnaire containing 25 items rating impairment by oromandibular dystonia in daily life. Based on the answers of 25 questions (graded from 0 to 4) five subscales are generated: the general scale (four items), psychosocial functioning (six), cosmesis (six), speech (three) and eating dysfunction (six). Each item is scaled on a five-point. Likert scale with five different grades of impairment: 0 = Never, 1 = Seldom, 2 = Sometimes, 3 = Often, 4 = Always.³³ The total score is linearly transferred to a range from 0–100, with high scores indicating more severe impairment in QoL. Patients are asked to answer the items according to their subjective perception of symptoms during the preceding 2 weeks.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?		Patient
Copyright or public domain?		Public domain
How can the scale be obtained (address or website)?		Published as Appendix of the original article ³³ .

Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No
---	--

Scale properties

Content

Face validity?	
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Oromanidbular region
Is the scale heavily weighted towards one type of dystonia? If so, which?	Oromanidbular
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	no
Is disability measured?	yes
Is health-related quality of life measured?	yes
Is tremor evaluated?	no
Does it exclude an important aspect of dystonia? If so, which?	
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	Yes (pain).
Includes items of sleep disturbance as a result of dystonia?	no
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	no
Which areas are not covered?	

Use

Was it <i>designed</i> to measure severity or to screen for	Severity/impact on hrQoL
---	--------------------------

dystonia?	
Is there a cut-off score for diagnosis of dystonia?	-
Is this appropriate in dystonia? If not, why?	-
Has it been <i>used</i> to measure severity or to screen for dystonia?	-

Acceptability

Length	10 min (personal judgment).
Clear instructions to raters?	Yes (personal judgment).
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No (personal judgment).
Appropriateness of questions/items for dystonia?	Good
Applicability across dystonia stages: mild – moderate – severe?	all

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	No
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Reliability tests showed internal consistency with satisfying to excellent results on all subscales. The internal consistency of the total score was very good with $\alpha = 0.91$. Test-retest reliability showed very good results for the ten patients included, with the total score as well as all subscores exceeding a Pearson's correlation coefficient >0.9 ³³ .
Validity: <u>assessed</u> – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	
Criterion validity (compared to gold-standard (Is there a gold	No gold standard

standard?) or other criteria for diagnosis)	
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Validity was supported with both convergent validity and discriminant validity³³. To assess the convergent validity of the OMDQ-25, its results were compared to the results of the Glasgow Health Scale Inventory (GHSI). The result for convergent validity on the total scores of OMDQ-25 and GHSI showed to be high ($r=0.83$). Analysing the subscales for convergent validity displayed moderate to high correlations for the general, psycho-social and cosmesis subscores when compared to the assumed equivalent of the GHSI ($r=0.69$, $r=0.72$, $r=0.76$, P always <0.001). Low correlations were found on comparing the speech and eating subscores to the physical/functional score of the GHSI ($r=0.11$, $r=0.14$). Although not significant ($P=0.6$, respectively $P=0.5$), this leads to the conclusion that these subscales measure different constructs. To determine the discriminant validity, OMDQ-25 subscales were related to a clinical grading of severity of disease (mild, moderate, severe). Correlating the total score and all subscores of the OMDQ-25 to the score of a basic clinical grading system showed coefficients constantly <0.85. Thus, discriminant validity is likely to exist between the two scales.
Valid in ethnically and culturally different populations?	Likely (personal judgment).
Valid in both genders and at all ages?	Likely (personal judgment).
Valid in patients with dementia or significant cognitive impairment?	No (personal judgment).
Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Proven by longitudinal study data. Sensitivity to change after Botulinum Toxin A treatment was established for all subscales ($P \leq 0.005$), and in comparison with the Glasgow Benefit Inventory (GBI), all subscales showed large effect sizes ³³ .
Has the minimal clinically important change been assessed in patients with dystonia?	No

Overall impression

Advantages and disadvantages	
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Can be used as specialised hrQoL scale in therapeutic trials for OMD

Final appraisal of the scale

Tick appropriate	Ranking
Yes - it has been studied clinimetrically and found valid, reliable and sensitive to change	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Laryngeal dystonia scales

Unified Spasmodic Dysphonia Rating Scale (USDRS)

1. Has the scale been used in dystonia?	X Yes ☐ No The USDRS ³⁴ was developed using a reiterative process, including voice experts, with 22 in the first round; 25 in the second iterative process and 118 in the third. The scale was developed with the input of these judges in the first 2 consensual evaluations. In the third evaluation, content validity and
---	---

	interrater reliability were determined. This scale has not been widely used, and the final scale has not been adequately assessed for reliability and validity.
2. Have clinimetric properties of this scale been evaluated in dystonia?	X Yes ☐ No Yes, initial clinimetric properties
Are there several versions of the scale? If so, which one has been assessed?	No other versions.
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	USDRS consists of 2 parts a. Reading and conversational speech I. 14 items, scored 0-7 b. Voice and speech task I. 6 items, rated normal or not normal
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Clinician rated. Requires training.
Copyright or public domain?	Public domain.
How can the scale be obtained (address or website)?	Published in Journal of Voice.
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes X No

Scale properties

Content

Face validity?	Face and content validity
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Developed for adductor spasmodic dysphonia
Is the scale heavily weighted towards one type of	Developed for adductor spasmodic

dystonia? If so, which?	dysphonia
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	Motor phenomenon measured for severity, amplitude
Is disability measured?	No
Is health-related quality of life measured?	No
Is tremor evaluated?	Vocal tremor is evaluated
Does it exclude an important aspect of dystonia? If so, which?	Includes spontaneous vocalization on standard passages and ratings of particular vocal tasks.
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No.
Includes items of sleep disturbance as a result of dystonia?	No.
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No.
Which areas are not covered?	Only covers motor features of adductor spasmodic dysphonia.

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Designed to measure severity not as a diagnostic tool.
Is there a cut-off score for diagnosis of dystonia?	Not for diagnosis.
Is this appropriate in dystonia? If not, why?	NA
Has it been <i>used</i> to measure severity or to screen for dystonia?	Has only been used in a few studies, no controlled studies ³⁵⁻³⁷ .

Acceptability

Length	The scale is long and complex for a non-Voice specialist.
--------	---

Clear instructions to raters?	The instructions may be clearer to voice specialists.
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	The ratings do not anchor to SD itself, but are 0-7 (no instance to profound). The difference between 6 (severe) and 7 (profound) is not clear.
Appropriateness of questions/items for dystonia?	Items appear to be very appropriate for spasmodic dysphonia.
Applicability across dystonia stages: mild – moderate – severe?	Does distinguish levels of severity.

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Not identified.
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Internal consistency, interrater reliability assessed. In the development study of USDRS ³⁴ , six symptoms exhibited good to excellent reliability: overall severity (R = 0.909), strained strangled voice quality (R = 0.886), voice arrest (R = 0.862), abrupt voice initiation (R = 0.799), aphonia (R = 0.790), and expiratory effort (R = 0.787). Eight continuous symptoms showed fair to good re- liability: speech intelligibility reduced (R = 0.740), speech rate (R = 0.712), related movements and grimaces (R = 0.685), bursts of loudness (R = 0.670), breathy voice quality (R = 0.629), rough voice quality (R = 0.535), voice tremor (R = 0.500), areas of movement (R = 0.411), and voice loudness (R = 0.410). Two symptoms showed poor reliability. Following these results, authors modified the USDRS by including only those items that showed either excellent or fair to good interrater reliability and high content validity.
Validity: assessed – not assessed, good – not good (references): ³⁴	
Face and/or content validity?	Content validity assessed through expert consensus. Face and content validity were

Which areas are not covered?	satisfactory.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	No criterion validity.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	No construct validity.
Valid in ethnically and culturally different populations?	Not evaluated in different populations.
Valid in both genders and at all ages?	No DIF for gender, age
Valid in patients with dementia or significant cognitive impairment?	Not assessed in dementia

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No correlations with biological markers (no biologic markers for SD).
---	---

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	One study evaluated effect of botulinum toxin over 12 years in 900 AD patients by using the USDRS. The adductor patients had an average benefit of 90% of normal function lasting an average of 15.1 weeks. The abductor patients had an average benefit of 66.7% of normal function lasting an average of 10.5 weeks ³⁶ .
Has the minimal clinically important change been assessed in patients with dystonia?	MCIC not established.

Overall impression

Advantages and disadvantages	This is a very specialized scale for adductor spasmodic dysphonia. It was developed through a reiterative process using consensus panels. It requires training to administer, and has terminology that may not be understood by those other than voice specialists. Although the original scale
------------------------------	---

	has good reliability and content validity, additional studies are needed.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	

Final appraisal of the scale

Tick appropriate	Ranking
Yes: -there are data on its use in studies beyond the group that developed the scale	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Voice Handicap Index (VHI)

1. Has the scale been used in dystonia?	Yes
2. Have clinimetric properties of this scale been evaluated in dystonia?	Yes
Are there several versions of the scale? If so, which one has been assessed?	Voice Handicap Index (VHI) and Voice Handicap Index-10 (VHI-10) (shorter form). Both have been evaluated separately
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	The VHI was developed by Jacobson and colleagues ³⁸ to determine the level of handicap and disability experienced by patients with all types of voice disorders. VHI has been used in a broad range of diagnoses (including laryngeal cancer and mass lesions, vocal fold polyps and cysts, and spasmodic dysphonia) ³⁹ . The original VHI has 30 items organized in 3 domains: 10-item functional subscale, 10-item emotional subscale, and 10-item physical subscale. There are 5 option responses: never, almost never, sometimes, almost always, and always. Scores range from 0 to 120, with 120 representing maximum perceived disability. This scale

	can be completed in 5 minutes.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient rated.
Copyright or public domain?	Public.
How can the scale be obtained (address or website)?	Published in the original paper ³⁸ .
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	These versions of the VHI have been translated and clinimetrically tested: German, Mandarin Chinese, Spanish, Dutch, Arabic, Japanese, Hebrew, and Greek. The validity of the French translation of the VHI has been confirmed but the results prompt to improve the quality of the translation. The VHI has been translated and adapted to Portuguese and Polish.

Scale properties

Content

Face validity?	Good. VHI scores illustrated substantial degrees of perceived handicap in patients with adductor spasmodic dysphonia ⁴⁰ .
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	VHI addresses disability related to verbal communication. Although it misleadingly implies specificity for voice problems such as laryngeal dystonia, it is likely to be sensitive to oromandibular dystonias, but it has not been tested in this population.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Yes, spasmodic dysphonia and oromandibular dystonia.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No
Is disability measured?	Yes

Is health-related quality of life measured?	Yes
Is tremor evaluated?	Not specifically, although it has been tested in voice tremor patients ⁴¹ .
Does it exclude an important aspect of dystonia? If so, which?	No
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	Yes, but only relating to voice loudness, speech clarity, etc...
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No
Which areas are not covered?	Not specific for dystonia-related voice or speech problems. The VHI was developed ³⁸ for patients with all types of voice disorders. The diagnoses of patients participating in the development study of VHI were: mass lesions, neurogenic – including spasmodic dysphonia, laryngectomy, musculoskeletal tension, inflammatory and atypical ³⁸ .

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Severity, but not dystonia itself; severity of disability only
Is there a cut-off score for diagnosis of dystonia?	No
Is this appropriate in dystonia? If not, why?	Yes
Has it been <i>used</i> to measure severity or to screen for dystonia?	Severity, but not dystonia itself; severity of disability only

Acceptability

Length	Good, very brief. Administration time: 5 minutes ^{39,42} .
Clear instructions to raters?	Yes (personal judgment). VHI appears also to be easy to score,

	requiring only the summation of all circled responses (for an index score) or responses from all relevant items (for the three scale dimensions).
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No (personal judgment).
Appropriateness of questions/items for dystonia?	Good (personal judgment).
Applicability across dystonia stages: mild – moderate – severe?	Yes

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Not reported.
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	All tested, and acceptable. In the development and validation study by Jacobson ³⁸, the VHI proved to have good internal consistency and good test-retest reliability. The VHI was validated in a group of 65 patients with a broad range of voice disorders [mass lesions (32%), neurogenic (26%) that included spasmodic dysphonia and vocal fold paralysis, laryngectomy (26%), musculoskeletal tension (8%), inflammatory (5%) and atypical (3%)]. Test-retest stability for subscale and total scores was found to be strong for the functional ($r=0.84$), emotional ($r=0.92$), physical ($r=0.86$) subscales, and total score ($r=0.92$). A 95% confidence interval of 18 points was established that gives users of this scale assurance that changes in total scores between administrations are not due to inherent variability in the VHI. Cronbach's alpha coefficient calculated for the item-total correlation was $r=0.95$. The correlation between the subscales was moderate-strong with Person product-moment correlations ranging from $r=0.70$ to $r=0.79$.

	The VHI was acknowledged in 2002 by the Evidence Practice Center Team (EPC) as a reliable and valid diagnostic tool ⁴³. Despite its wide clinical and research application, a recent study argue the reliability of the VHI when more stringent reliability criteria (coefficients > 0.90) in comparison to EPC (coefficients > 0.80) are used ³⁹. Nevertheless, the results of this study do confirm the EPC's recommendation of overall VHI total score (not dimension scores) for individual decision making. VHI met internal consistency and reproducibility criteria, as defined in the study by Franic DM and colleagues (Franic DM et al, 2005), for the overall instrument score ($\alpha = 0.95$ and $r = 0.92$, respectively) (Jacobson BH et al, 1997), but coefficient α was not reported for its three subscales. Overall, the VHI score can be regarded as a reliable measure for individual decision making. The 95% SEM (standard error of measurement) for internal consistency was narrow (± 6.44), indicating that total VHI score would be useful in individual patient decision making. VHI emotional subscale ($r = 0.92$) ³⁸ was the only dimension to meet test-retest criterion of the study by Franic ³⁹.
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	The items of VHI were selected from patients' reports to ensure that the scale had both content and face validity ³⁸.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	There is no gold standard
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Construct validity was not fully evaluated in the study by Jacobson ³⁸, although the relationship between patient self-perceived severity and VHI scores was determined to be moderately strong. Convergent validity for VHI (subscales and

total score) was supported by significant correlation with SF-36 subscales (mental, social, and emotional role functioning) between different vocal pathologies ⁴⁴⁻⁴⁶. Data are lacking regarding discriminant validity for VHI ³⁹.

Wuyts et al. ⁴⁷ found a significantly high correlation ($r = -0.79$) between VHI and the DSI (dysphonia severity index), an objective measure of vocal quality based on a multiparameter approach. The study involved 68 “normal” speakers and 319 dysphonics. Unfortunately, VHI data for the total and subscales scores is not reported.

The correlation between participants’ scores on the Communicative Participation Item Bank (CPIB) – a new self-report instrument designed to measure the extent to which communication disorders interfere with communicative participation – and total VHI scores using Spearman’s correlation was -0.678 , in this study on spasmodic dysphonia patients. The degree of correlation between the VHI and the CPIB found in this study suggests that the two instruments cover related but not entirely redundant content. The VHI does contain some items that measure communicative participation, and this would account for the degree of correlation observed between the two tools. The lack of complete overlap, however, stems from the observation that the CPIB is designed to contain only items addressing communicative participation, whereas the VHI contains items covering a wider array of constructs, including physical symptoms and personal or emotional consequences ⁴⁸.

The VHI, VHI-10, and V-RQOL are highly correlated, and all reflect the morbidity associated with in adductor spasmodic dysphonia ⁴⁹.

Valid in ethnically and culturally different populations?	Yes. VHI has been used in a variety of races (Caucasians, African, Americans, Hispanics, and Asians) ³⁹ .
Valid in both genders and at all ages?	Applicable to both genders, but not to children because it contains items that involve using telephone, relating to occupation, or substantial introspection; never tested in children. VHI has been used in men and women. VHI was designed for use in adults, and has been used from 18 to 86 years of age ³⁹ .
Valid in patients with dementia or significant cognitive impairment?	No, as a self-rated scale it cannot be used in dementia or mentally retarded subjects.

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No
---	----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	The VHI has shown a significant change following treatment for different voice disorders, including spasmodic dysphonia, indicating that the VHI is a useful instrument to monitor the treatment efficacy for voice disorders. Rosen et al ⁴² in a study including patients with unilateral vocal fold paralysis, muscle tension dysphonia, vocal fold polyp, and vocal cyst indicate that VHI is sensitive to speaker's perception of voice change after different types of intervention (surgery, medical, and voice therapy) and is therefore a useful instrument to measure treatment efficacy. Similarly, Spector et al ⁵⁰ also identified that the VHI is an important tool in the evaluation of patients with unilateral vocal cord paralysis before and after thyroplasty. Several studies have demonstrated the efficacy of different treatments for spasmodic dysphonia through the VHI. Treatments include botulinum toxin ^{49,51,52},

surgery ⁵³⁻⁵⁶ and acupuncture ³⁷.

Tsuji DH, et al 2006: Prospective case series study in patients with adductor spasmodic dysphonia. There was a clear difference in VHI before and after surgery, in partial myectomy and neurectomy of thyroarytenoid muscle.

In a prospective cohort study, Novakovic D. et al. ⁵² examined the longitudinal effect of botulinum toxin (Botox) treatment for adductor spasmodic dysphonia upon functional outcomes and quality of life when using a patient-specific dosing regimen. Mean VHI improvement was 9.6%.

In a prospective case series, responsiveness of the VHI, VHI-10, and Voice-Related Quality of Life (V-RQOL) to the treatment effect of Botox in adductor spasmodic dysphonia was assessed ⁴⁹. The average follow-up time after injection was 7.84 weeks. All the three instruments were highly correlated in subscale and total scores. After treatment, all three instruments showed significant improvement. The VHI, VHI-10, and V-RQOL all reflected the morbidity associated with adductor spasmodic dysphonia and were significantly responsive to the effect of Botox therapy.

In a prospective case series ⁵⁵ lateral laser thyroarytenoid myotomy was found to be a curative procedure for adductor spasmodic dysphonia. All patients reported improvement in different outcome measure, including VHI, over 2.5 years' follow up.

In the pilot study conducted by Tsuji DH and colleagues ⁵⁶, the long-term vocal quality results of endoscopic laser thyroarytenoid myoneurectomy were reported. A significant improvement in VHI was observed after surgery, as compared with baseline values ($P=0.001$). The median and interquartile range for preoperative

	VHI was 99 and 13, respectively and 24 and 42, for postoperative VHI. Chhetri DK et al, 2006: Authors report long-term results of selective laryngeal adductor denervation-reinnervation surgery for the treatment of adductor spasmodic dysphonia. Measured outcomes included VHI-10 scores. Mean VHI-10 scores improved from a mean of 35.6 to 12.7. Lee L et al, 2003: prospective case series that evaluated the effect of acupuncture for the treatment of adductor spasmodic dysphonia. Significant differences were found on all three subtests of the VHI, and the average total scores pre- treatment and post-treatment differed by 17 points. Benninger MS et al, 2001: In this study significant improvements in all 3 subscale scores and the total score on the VHI occurred after treatment with botulinum toxin. Courey MS et al, 2000: A prospective analysis of the effects of botulinum toxin in patients with spasmodic dysphonia. The VHI, improvements in the patients' perception of their functional, physical, and emotional voice handicap reached statistical significance ($p < \text{or} = .0005$).
Has the minimal clinically important change been assessed in patients with dystonia?	Not assessed to my knowledge.

Overall impression

Advantages and disadvantages	Very simple and efficient, but as a disability scale, it has no discriminant value and is not specific for dystonic movement. Considering that the VHI was validated on few patients with laryngeal dystonia compared to the total number of patients assessed, it still needs further validation for dystonia. VHI is similar to the Vocal Performance
------------------------------	--

	Scale, and direct comparisons have been made showing similar clinimetric properties. The value of having two scales for the same purpose is questionable and a sensible recommendation would be to merge them or pick one for future use.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	It has been used successfully to measure outcomes in interventions, e.g. botulinum toxin and various surgeries.

Final appraisal of the scale

Tick appropriate	Ranking
Yes	Recommended Requirements are fulfilled if: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Voice Handicap Index 10 (VHI-10)

1.	Has the scale been used in dystonia?	Yes
2.	Have clinimetric properties of this scale been evaluated in dystonia?	The clinimetric properties of the VHI-10 have been evaluated in a broad range of voice disorders, including spasmodic dysphonia.
Are there several versions of the scale? If so, which one has been assessed?		There is the original 30-item VHI version ³⁸ .
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests))		The VHI-10 is a 10 item, shortened version of the original 30-item VHI ³⁸ . The VHI-10 represents a distilling of 3 sub-scales into one shortened scale. For each question the patient is required to rate each answer using a 5-point scale. The total score is used to indicate the severity of the voice disorders for the patient (maximum score = 40, minimum score = 0). Appendix 2 describes the text of each item in the VHI-10. This scale can be completed in less than 2 minutes.

Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient rated.
Copyright or public domain?	Public.
How can the scale be obtained (address or website)?	Published in ⁵⁷.
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	The VHI-10 has been translated and clinimetrically tested in Chinese and Spanish.

Scale properties

Content

Face validity?	Item analysis of the VHI was performed using the VHI responses of 100 patients with voice disorders (including spasmodic dysphonia) and 159 control subjects. The 10 most robust VHI items were selected using the item analysis and clinical consensus results to form the VHI-10 ⁵⁸.
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	VHI-10 addresses disability related to verbal communication. Although it misleadingly implies specificity for voice problems such as laryngeal dystonia, it is likely to be sensitive to oromandibular dystonias, but it has not been tested in this population.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Yes, spasmodic dysphonia and oromandibular dystonia.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No
Is disability measured?	Yes
Is health-related quality of life measured?	Yes
Is tremor evaluated?	No

Does it exclude an important aspect of dystonia? If so, which?	No
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No
Which areas are not covered?	Not specific for dystonia-related voice or speech problems.

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Severity, but not dystonia itself; severity of disability only.
Is there a cut-off score for diagnosis of dystonia?	No
Is this appropriate in dystonia? If not, why?	Yes
Has it been <i>used</i> to measure severity or to screen for dystonia?	Severity, but not dystonia itself; severity of disability only.

Acceptability

Length	Good, very brief. Administration time: less than 2 minutes (personal judgment).
Clear instructions to raters?	Yes (personal judgment).
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No (personal judgment).
Appropriateness of questions/items for dystonia?	Good (personal judgment).
Applicability across dystonia stages: mild – moderate – severe?	Yes

Clinimetric properties

Metric flaws (floor and ceiling effects, score	Not reported.
--	---------------

distributions)?	
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Rosen et al ⁵⁸ performed statistical analysis comparing the validity of the VHI-10 with the VHI in 819 patients with a wide spectrum of voice disorders, including spasmodic dysphonia. No statistically significant differences between the VHI and the VHI-10 were found. Irrespective of the diagnosis, the correlation between the VHI and the VHI-10 was greater than 0.90 ($p < 0.001$). The ratios of the VHI-10 to VHI scores for a variety of voice disorder categories were analyzed and found to be consistently greater than the expected value (33%). This suggests that the VHI-10 may be a more robust instrument than the VHI ⁵⁸.
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	The 10 most robust VHI items were selected using the item analysis and clinical consensus results to form the VHI-10 ⁵⁸.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	There is no gold standard
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	DYT: The VHI, VHI-10, and V-RQOL are highly correlated, and all reflect the morbidity associated with in adductor spasmodic dysphonia ⁴⁹. The pretreatment VHI domain and total scores were 23.49±8.64 (functional), 23.43±9.51 (physical), 22.76±10.51 (emotional), and 69.95±22.23 (total). The pretreatment VHI-10 was 24.54±7.19. The pretreatment V-RQOL domain and total scores were 41.22±24.50 (social-emotional), 43.70±22.10 (physical functioning), and 47.50±19.26 (total). There was a significant correlation between and within all three QOL instruments. After Botox therapy, the VHI scores improved to 16.49±8.63 (functional), 18.54±9.28 (physical), 16.14±9.80 (emotional),

	and 51.11 ± 21.14 (total). The VHI-10 improved to 15.64 ± 6.90 . The V-RQOL domain and total scores improved to 80.27 ± 18.73 (social-emotional), 83.36 ± 15.18 (physical functioning), and 87.64 ± 12.05 (total). The improvement in QOL scores was statistically significant for all three instruments. The post treatment QOL instruments were again highly correlated, and there remained no difference among the VHI or V-RQOL domains.
Valid in ethnically and culturally different populations?	Yes. The original VHI version has been used in a variety of races (Caucasians, African, Americans, Hispanics, and Asians) ³⁹ .
Valid in both genders and at all ages?	Applicable to both genders, but not to children because it contains items that involve using telephone, relating to occupation, or substantial introspection; never tested in children.
Valid in patients with dementia or significant cognitive impairment?	No, as a self-rated scale it cannot be used in dementia or mentally retarded subjects.

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No
---	----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	In a prospective case series, responsiveness of the VHI, VHI-10, and Voice-Related Quality of Life (V-RQOL) to the treatment effect of Botox in adductor spasmodic dysphonia was assessed ⁴⁹ . The average follow-up time after injection was 7.84 weeks. All the three instruments were highly correlated in subscale and total scores. After treatment, all three instruments showed significant improvement. The VHI, VHI-10, and V-RQOL all reflected the morbidity associated with adductor spasmodic dysphonia and were significantly responsive to the effect

	of Botox therapy. Long-term results of selective laryngeal adductor denervation-reinnervation surgery for the treatment of adductor spasmodic dysphonia was reported ⁵³. Measured outcomes included VHI-10 scores. Mean VHI-10 scores improved from a mean of 35.6 to 12.7.
Has the minimal clinically important change been assessed in patients with dystonia?	Not assessed to my knowledge.

Overall impression

Advantages and disadvantages	Very simple and efficient, but as a disability scale, it has no discriminant value and is not specific for dystonic movement.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	It has been used successfully to measure outcomes in interventions, e.g. botulinum toxin and various surgeries.

Final appraisal of the scale

Tick appropriate	Ranking
Yes: - there are data on its use in studies beyond the group that developed the scale	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Pediatric Voice Handicap Index (pVHI)

1. Has the scale been used in dystonia?	X Uncertain ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	X Uncertain ☐ No
Are there several versions of the scale?	No.
If so, which one has been assessed?	

Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	Three part instrument with seven to nine items each in the physical, functional, and emotional domains with each item rated on a 5 point scale. Time frame is current status.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Parent.
Copyright or public domain?	Copyright.
How can the scale be obtained (address or website)?	Copyright by American Speech, Language, and Hearing Association.
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	X Yes ☐ No Italian and Arabic language.

Scale properties

Content

Face validity?	Yes (personal judgement).
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Voice only ⁵⁹.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Dysphonia ⁵⁹.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No.
Is disability measured?	Yes.
Is health-related quality of life measured?	Yes.
Is tremor evaluated?	No.
Does it exclude an important aspect of dystonia? If so, which?	Not specific for dystonia.

Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No.
Includes items of sleep disturbance as a result of dystonia?	No.
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	Yes.
Which areas are not covered?	

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Measure severity.
Is there a cut-off score for diagnosis of dystonia?	No.
Is this appropriate in dystonia? If not, why?	Probably for pediatric laryngeal dystonia, but the number of patients with laryngeal dystonia included was not specified.
Has it been <i>used</i> to measure severity or to screen for dystonia?	No.

Acceptability

Length	Yes.
Clear instructions to raters?	Yes.
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	Anchors are somewhat ambiguous.
Appropriateness of questions/items for dystonia?	Yes.
Applicability across dystonia stages: mild – moderate – severe?	Unknown.

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Distributions not reported.
Reliability (internal consistency, interrater, intrarater, test-	Test-retest reliability very good. Others not tested.

retest) of total or sub-scales and of individual items	
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Assessed in dysphonic children vs. normal controls. Validity very good.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Not assessed.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Overall correlations between clinician-derived CAPE-V assessments (Consensus Auditory-Perceptual Evaluation of Voice) and proxy-reported pVHI assessments in the whole study population were fair (Spearman rank correlation = 0.34-0.44) ⁶⁰ . This is consistent with a previous report of this finding in children after airway reconstruction (Spearman rank correlation = 0.32-0.53) ⁶¹ .
Valid in ethnically and culturally different populations?	Not assessed.
Valid in both genders and at all ages?	Only validated in children.
Valid in patients with dementia or significant cognitive impairment?	Not assessed.

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No.
---	-----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Not assessed.
Has the minimal clinically important change been assessed in patients with dystonia?	No.

Overall impression

Advantages and disadvantages	Advantage is utility in children and assessment of physical, functional, and
------------------------------	--

	emotional domains. Disadvantage is lack of specificity to dystonia.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Suitable as an outcome measure. Not suitable for screening or diagnosis.

Final appraisal of the scale

Tick appropriate	Ranking
Yes - there are data on its use in studies beyond the group that developed the scale	Listed Fulfilled if one of the following applies: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Pediatric Voice-Related Quality of Life (PVRQOL)

1. Has the scale been used in dystonia?	Uncertain
2. Have clinimetric properties of this scale been evaluated in dystonia?	Uncertain
Are there several versions of the scale? If so, which one has been assessed?	No.
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	10-item instrument adapted from the adult Voice-Related QOL instrument. Severity rated on a 7 point scale for each item and the total score is transformed to a scale of 0 to 100. Subdomains of social-emotional and physical-functional effects can be extracted. Time frame is current status.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Parent.
Copyright or public domain?	Public domain
How can the scale be obtained (address or website)?	Published in ⁶² .

Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No
---	---

Scale properties

Content

Face validity?	Yes (personal judgement).
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	One body area – voice.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Developed for all pediatric voice disorders, but dysphonia was included.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No.
Is disability measured?	No.
Is health-related quality of life measured?	Yes.
Is tremor evaluated?	No.
Does it exclude an important aspect of dystonia? If so, which?	It is not specific to dystonia.
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No.
Includes items of sleep disturbance as a result of dystonia?	No.
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	Yes (anxious, frustrated, depressed).
Which areas are not covered?	

Use

Was it <i>designed</i> to measure severity or to screen for	No. Neither.
---	--------------

dystonia?	
Is there a cut-off score for diagnosis of dystonia?	No.
Is this appropriate in dystonia? If not, why?	Perhaps. In the sample used for reliability and validity testing, only 8% had dysphonia.
Has it been <i>used</i> to measure severity or to screen for dystonia?	Measure health-related QOL.

Acceptability

Length	Yes.
Clear instructions to raters?	Yes.
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	Anchor points are somewhat ambiguous.
Appropriateness of questions/items for dystonia?	Appropriate.
Applicability across dystonia stages: mild – moderate – severe?	Unknown.

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Distributions reasonable, but scores were on the mild end in the sample studied ⁶² .
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Good test-retest reliability (0.8)
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Assessed. Good, but not specific for dysphonia.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Not assessed.
Construct validity (correlations with other convergent scales and	Correlated well with the Pediatric Voice Outcomes Survey (PVOS), a previously

divergent scales (which?); factor analysis; known-groups comparisons).	validate scale. There was a high correlation between the PVOS and the PVRQOL (r=0.70). Discriminant validity supported. Validity of the PVQROL was tested by evaluating its ability to show significant change in voice-related quality-of-life after adenoidectomy (discriminant validity) (P<0.001).
Valid in ethnically and culturally different populations?	Not assessed.
Valid in both genders and at all ages?	Not assessed in adults.
Valid in patients with dementia or significant cognitive impairment?	Not assessed.
Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No.

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Not assessed. It has been used in several studies of non-dystonic children to assess the impact of different therapies on voice quality of life ⁶³⁻⁶⁵ .
Has the minimal clinically important change been assessed in patients with dystonia?	No.

Overall impression

Advantages and disadvantages	Advantage – health-related QOL scale for voice-related problems in children. Disadvantage – not specifically tested for laryngeal dystonia / dysphonia, although some dysphonia patients were included in testing.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes,	Suitable: Treatment trial and clinical assessment. Unsuitable: Screening or diagnosis.

clinical practice for diagnosis/screening). List all.	
---	--

Final appraisal of the scale

Tick appropriate	Ranking
Yes	Listed Fulfilled if one of the following applies: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Voice-Related Quality of Life (V-RQOL)

1. Has the scale been used in dystonia?	☒ Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	☒ Yes ☐ No In the study that evaluated the clinimetric properties of the V-RQOL the patients included had different diagnosis, i.e. inflammatory, neurological (not better specified), mass lesions, others ⁶⁶ . Only responsiveness in dystonia patients ⁶⁷ .
Are there several versions of the scale? If so, which one has been assessed?	No
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	Instrument to Measure Voice-Related Quality of Life. This instrument is called with various names in different papers (e.g., Voice-Related Quality of Life (V-RQOL) Instrument, Voice-Related Quality of Life (V-RQOL) Questionnaire). 10 item, 5 point response
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient
Copyright or public domain?	Public domain
How can the scale be obtained (address or website)?	Publication

Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No
---	---

Scale properties

Content

Face validity?	Yes ⁶⁷ .
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Voice
Is the scale heavily weighted towards one type of dystonia? If so, which?	Spasmodic dysphonia
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	no
Is disability measured?	yes
Is health-related quality of life measured?	yes
Is tremor evaluated?	no
Does it exclude an important aspect of dystonia? If so, which?	Only speech
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	no
Includes items of sleep disturbance as a result of dystonia?	no
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	no
Which areas are not covered?	

Use

Was it <i>designed</i> to measure severity or to screen for	Severity of speech impairment
---	-------------------------------

dystonia?	
Is there a cut-off score for diagnosis of dystonia?	no
Is this appropriate in dystonia? If not, why?	
Has it been <i>used</i> to measure severity or to screen for dystonia?	Measure impact of spasmodic dysphonia

Acceptability

Length	Most patients took about 5 minutes to complete the questionnaire, and could do so independently ⁶⁷ .
Clear instructions to raters?	Yes. Most patients took about 5 minutes to complete the questionnaire, and could do so independently ⁶⁷ .
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No (personal judgment).
Appropriateness of questions/items for dystonia?	OK for spasmodic dysphonia
Applicability across dystonia stages: mild – moderate – severe?	All (personal judgment).

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Not assessed.
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Reliability high (the alpha co- efficient of the final 10-item instrument is 0.89) ⁶⁶ . Both measures of reliability, the internal consistency and reproducibili- ty, are strong for the domain or total scores ⁶⁶ .
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	High face validity ⁶⁶ .
Criterion validity	No gold standard, validated against global

(compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	speech impairment assessment
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	In the study by Hogikyan et al ⁶⁶ construct-related validity is clearly demonstrated by the relationship between self-perceived voice quality rating and mean V-RQOL scores and by the large difference in mean scores for new voice and new non-voice patients. A moderate correlation of 0.51 is also present between the SF-36 social functioning domain and the Social-Emotional domain of the VRQOL Measure. A lower correlation between the physical functioning domain of the SF-36 and the V-RQOL Physical Functioning domain suggests that a general health survey is not sensitive enough to detect impairment in a voice-related physical functioning domain. SF-36 is a general health status instrument and is not disease-specific for voice disorders.
Valid in ethnically and culturally different populations?	Not assessed
Valid in both genders and at all ages?	Not assessed
Valid in patients with dementia or significant cognitive impairment?	No (personal judgment).

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No
---	----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Yes (but not for dystonia alone). The data from the study by Hogikyan et al ⁶⁶ demonstrate a clear relationship between degree of voice improvement following treatment and change in V-RQOL score. <u>DYT</u>: study that measure the V-RQOL of patients with adductory spasmodic dysphonia and the treatment effects of

	botulinum toxin injections ⁶⁷ .
Has the minimal clinically important change been assessed in patients with dystonia?	Yes (but not for dystonia alone)

Overall impression

Advantages and disadvantages	Scale which may be added to document impact of dystonia related speech impairment on ADL and hrQoL
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	

Final appraisal of the scale

Tick appropriate	Ranking
Yes	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Vocal Performance Questionnaire (VPQ)

1. Has the scale been used in dystonia?	Yes ⁶⁸ .
2. Have clinimetric properties of this scale been evaluated in dystonia?	Yes
Are there several versions of the scale? If so, which one has been assessed?	No
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests))	This scale was designed for use in an evaluation study of voice therapy in cases of non-organic dysphonia ⁶⁹ . The VPQ is a questionnaire designed to allow the patients to consider aspects of their own vocal performance and rate the severity of that aspect. It is a 12-item questionnaire. It is designed using an answer format in which the patient selects

	the statement that best answers each question. The statements are graded in terms of severity of vocal performance. A numerical score from 1 to 5 is assigned to each answer and these are summed to provide a single total score of severity. The maximum severity score is 60 and minimum score is 12.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient.
Copyright or public domain?	Public.
How can the scale be obtained (address or website)?	Published by Carding et al. ⁶⁹ .
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	Yes.

Scale properties

Content

Face validity?	Good (personal judgment).
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	VPQ addresses disability related to verbal communication. Although it misleadingly implies specificity for voice problems such as laryngeal dystonia, it is likely to be sensitive to oromandibular dystonias, but it has not been tested in this population.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Yes, spasmodic dysphonia and oromandibular dystonia.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No
Is disability measured?	Yes
Is health-related quality of life measured?	Yes

Is tremor evaluated?	Not specifically, although it has been tested in voice tremor patients
Does it exclude an important aspect of dystonia? If so, which?	No
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	Yes, but only relating to voice loudness, speech clarity, etc.
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	Yes
Which areas are not covered?	Not specific for dystonia-related voice or speech problems

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Severity, but not dystonia itself; severity of disability only
Is there a cut-off score for diagnosis of dystonia?	No
Is this appropriate in dystonia? If not, why?	Yes, but only for spasmodic
Has it been <i>used</i> to measure severity or to screen for dystonia?	Severity, but not dystonia itself; severity of disability only

Acceptability

Length	Good, it takes less than 5 minutes to complete (personal judgment).
Clear instructions to raters?	Yes (personal judgment).
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No (personal judgment).
Appropriateness of questions/items for dystonia?	Good (personal judgment).
Applicability across dystonia stages: mild – moderate – severe?	Good (personal judgment).

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Not assessed.
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Good. In the study by Deary et al. ⁵⁷ on a large range of voice pathologies, the VPQ was found to have good internal consistency (Cronbach alpha was 0.91). Also in the study by Webb et al (2007), VPQ had high levels of internal consistency (Cronbach's alpha coefficient for the total score was 0.81) and test-retest reliability (0.75).
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Good (personal judgment).
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	No gold standard.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Good. In the study by Deary et al. ⁵⁷ on a large range of voice pathologies, the clinimetric properties of the VPQ were examined and compared with those of the Voice Handicap Index-10 (VHI-10). Deary et al. found the VHI-10 and VPQ were highly correlated (disattenuated $r=0.91$).
Valid in ethnically and culturally different populations?	Yes (personal judgment).
Valid in both genders and at all ages?	Applicable to both genders, and may be applicable to children (personal judgment). Unlike the very similar VHI scale, this scale does not contain many items that are not appropriate for children. But never tested in children.
Valid in patients with dementia or significant cognitive impairment?	No (personal judgment).

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No.
---	-----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Carding et al. ⁶⁹ (dysphonic patients of no organic cause were included) found a statistically significant correlation (Spearman rho = 0.648, P<0.05) between the VPQ and the mean severity ratings (as judged by trained listeners). This strong correlation enhances VPQ as valid measure of voice change over time. The VPQ showed statistically significant changes (P<0.05) following a specific therapy program (direct with indirect therapy) in comparison to the other treatment groups (no treatment and indirect therapy, respectively). Therefore, VPQ appears a reliable indicator of voice change following therapy. Other studies assessed the responsiveness to change of the VPQ following voice therapy and surgery in non-organic dysphonia ⁶⁹⁻⁷².
Has the minimal clinically important change been assessed in patients with dystonia?	No

<i>Overall impression</i>	Unfortunately, the authors of the original scale did not formally name their scale, so it sometimes is used under different names, such as "Vocal Rating Scale". This scale is similar to the Voice Handicap Index, and direct comparisons have been made, showing similar clinimetric properties. The value of having two scales for the same purpose is questionable and a sensible recommendation would be to merge them or pick one for future use.
Advantages and disadvantages	Very simple and efficient, but as a disability scale, it has no discriminant value and is not

	specific for dystonic movement. Considering that the VPQ was validated on few patients with laryngeal dystonia compared to the total number of patients assessed, it still needs further validation for dystonia.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	It has been used successfully to measure outcomes in interventions, eg botulinum toxin and various surgeries ^{72,73} . The scale meets criteria for “recommended”, but has been used on dystonia patients in only one paper. For this reason it is not included among the recommended scales in the manuscript.

Final appraisal of the scale

Tick appropriate	Ranking
Yes	Recommended Requirements are fulfilled if: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Arm dystonia scales

Arm Dystonia Disability Scale (ADDS)

1. Has the scale been used in dystonia?	☒ Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	☐ Yes ☒ No
Are there several versions of the scale? If so, which one has been assessed?	No.

Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests))	The ADDS rates the disability in arm dystonia. This scale assesses the impairment of seven activities of daily living that are specific for the upper limb: writing, playing a musical instrument, buttoning, handling utensils and feeding, hygiene, grasping objects, and housework or outside job. Each is scored on a scale of 0 to 4 points: 0, normal; 1, mild difficulty; 2, moderate difficulty; 3, marked difficulty. Calculations: total the points scored, divided by maximum possible for each patient, multiply the quotient by 90, and subtract from 90%. An ADDS score of 100% indicates the absence of any motor impairment, whereas a score of 90% denotes the presence of social restrictions but no limitation in activities. Scores below 90% indicate significant impairment.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Based on the patient's subjective impairment in activities of daily living.
Copyright or public domain?	Public in the original paper.
How can the scale be obtained (address or website)?	Original publication ⁷⁴.
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No

Scale properties

Content

Face validity?	Not assessed.
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Upper limb.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Upper-limb dystonia, including arm dystonia that is specific to a task (playing,

	writing).
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No.
Is disability measured?	Yes.
Is health-related quality of life measured?	No.
Is tremor evaluated?	No.
Does it exclude an important aspect of dystonia? If so, which?	Task specific dystonia (writer's cramp and musician dystonia).
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No.
Includes items of sleep disturbance as a result of dystonia?	No.
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No.
Which areas are not covered?	Non-motor feature related to dystonia are not assessed (personal judgment).

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	To measure severity of disability in arm dystonia.
Is there a cut-off score for diagnosis of dystonia?	No.
Is this appropriate in dystonia? If not, why?	NA.
Has it been <i>used</i> to measure severity or to screen for dystonia?	The ADDS has been used in studies to assess motor impairment in writer's cramp and in pianists with dystonia ^{75,76} . It has not been used to screen for dystonia.

Acceptability

Length	Acceptable, it takes less than 5 minutes to complete (personal judgment).
--------	---

Clear instructions to raters?	Yes (personal judgment).
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No (personal judgment).
Appropriateness of questions/items for dystonia?	Yes (personal judgment).
Applicability across dystonia stages: mild – moderate – severe?	Yes (personal judgment).

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Not assessed.
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Not assessed.
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Not assessed.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Not assessed.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Not assessed.
Valid in ethnically and culturally different populations?	Not assessed.
Valid in both genders and at all ages?	Not assessed.
Valid in patients with dementia or significant cognitive impairment?	Not assessed.

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No.
---	-----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	No specific validation study on ADDS sensitivity to change (over time or treatment) has been found. Clinically, sensitivity to change was demonstrated in three small open studies on limb immobilization treatment, extracorporeal shock wave therapy, and re-training, respectively, in patients with upper limb dystonia ⁷⁷⁻⁷⁹.
Has the minimal clinically important change been assessed in patients with dystonia?	No.

Overall impression

Advantages and disadvantages	<i>Advantages:</i> - Assessment of the ADDS is specific for the focal arm dystonia. - It can be also used in upper limb task-specific dystonia (writer's cramp, musician dystonia). <i>Disadvantages:</i> - The ADDS do not provides fine resolution, especially within the range between "normal" and "moderate difficulty in playing" ⁷⁶; - It cannot distinguish the disability of dystonia from that of spasticity, if the latter is also present.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Useful for quantification of disability due to focal dystonia in playing piano/writing and for asses and monitoring treatment effects, even if ADDS has been found to provide low resolution ⁷⁶ .

Final appraisal of the scale

Tick appropriate	Ranking
Yes: - there are data on its use in studies beyond the group that developed the scale	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Task-specific dystonia scales

Dystonia Evaluation Scale (DES)

1. Has the scale been used in dystonia?	X Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	☐ Yes X No
Are there several versions of the scale? If so, which one has been assessed?	No
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	Five point Likert scale: 0=dystonia as bad as its worst 1=slightly improved 2=moderately improved, 3=almost normal 4=normal Scale is not described in detail. Apparently (judging from the figure in primary article) intermediate values are also used.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient
Copyright or public domain?	Public domain ⁸⁰ .
How can the scale be obtained (address or website)?	Unknown

Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No
---	---

Scale properties

Content

Face validity?	Yes (personal judgement)
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	One body area – used for focal hand dystonia in musicians ⁸⁰ .
Is the scale heavily weighted towards one type of dystonia? If so, which?	Yes. Focal hand dystonia in musicians.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No.
Is disability measured?	No
Is health-related quality of life measured?	No
Is tremor evaluated?	No
Does it exclude an important aspect of dystonia? If so, which?	Duration, variability, task-specificity.
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No
Which areas are not covered?	

Use

Was it <i>designed</i> to measure severity or to screen for	Measure severity.
---	-------------------

dystonia?	
Is there a cut-off score for diagnosis of dystonia?	N/A
Is this appropriate in dystonia? If not, why?	
Has it been <i>used</i> to measure severity or to screen for dystonia?	Yes.

Acceptability

Length	Brief ⁸⁰ .
Clear instructions to raters?	No.
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	Description provided with insufficient detail to judge.
Appropriateness of questions/items for dystonia?	Fair.
Applicability across dystonia stages: mild – moderate – severe?	Yes.

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Insufficient data.
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Unknown.
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Fair (personal judgement).
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Compared to clinician judgement and quantification of finger movements, but data on correlation not provided ⁸⁰ .
Construct validity (correlations with other convergent scales and	Not assessed.

divergent scales (which?); factor analysis; known-groups comparisons).	
Valid in ethnically and culturally different populations?	Not assessed.
Valid in both genders and at all ages?	Not assessed.
Valid in patients with dementia or significant cognitive impairment?	Not assessed.

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	Compared to clinician judgement and quantification of finger movements, but data on correlation not provided ⁸⁰ .
---	--

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Yes ⁸⁰ . A one-way analysis of variance followed by post-hoc Scheffé tests revealed that pre-treatment scores were different from post-treatment ($t[4]=5.2, p<0.01$) and 1-month follow-up ($t[4]=7.2, p<0.01$) scores. The results on the DES were confirmed by the clinical evaluation of the treating neurologist and by the findings from the dexterity/displacement device obtained in the laboratory.
Has the minimal clinically important change been assessed in patients with dystonia?	No.

Overall impression

Advantages and disadvantages	Simple, but insufficient data to judge otherwise.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Seems appropriate for outcome measure in musician's focal hand dystonia, but the scale is little more than a patient CGI.

Final appraisal of the scale

Tick appropriate	Ranking
-------------------------	----------------

Yes: - it has been applied to dystonia populations scale	Listed Fulfilled if one of the following applies: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change
---	---

Tubiana and Chamagne Score (TCS)

1. Has the scale been used in dystonia?	☒ Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	☐ Yes ☒ No
Are there several versions of the scale? If so, which one has been assessed?	No
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	Rating scale of patients' technical performance : 0 = unable to play; 1 = plays several notes but stops because of blockage or lack of facility; 2 = plays short sequences without rapidity and with unsteady fingering; 3=5 plays easy pieces but is unable to perform more technically challenging pieces; 4 = plays almost normally but avoids difficult passages for fear of motor problems; 5 = returns to concert performances).
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Patient
Copyright or public domain?	Public ⁸¹ .
How can the scale be obtained (address or website)?	From the original paper
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No Only available in French.

Scale properties

Content

Face validity?	Not assessed
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Hands
Is the scale heavily weighted towards one type of dystonia? If so, which?	Musician
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	No
Is disability measured?	Functional disability while playing
Is health-related quality of life measured?	No
Is tremor evaluated?	No
Does it exclude an important aspect of dystonia? If so, which?	Yes, phenomenology
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No
Which areas are not covered?	Features specific to dystonia

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Measure severity
Is there a cut-off score for diagnosis of dystonia?	No
Is this appropriate in dystonia? If not, why?	No, it
Has it been <i>used</i> to measure severity or to screen for dystonia?	Measure severity

Acceptability

Length	Acceptable
Clear instructions to raters?	Very simple rules
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	Yes
Appropriateness of questions/items for dystonia?	No
Applicability across dystonia stages: mild – moderate – severe?	Yes

Clinimetric properties

Not assessed

Metric flaws (floor and ceiling effects, score distributions)?	
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	
Valid in ethnically and culturally different populations?	
Valid in both genders and at all ages?	
Valid in patients with dementia or significant cognitive impairment?	

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	
---	--

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	
Has the minimal clinically important change been assessed in patients with dystonia?	

Overall impression

Advantages and disadvantages	Adv: Simplicity Disad: lack of specificity
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	None

Final appraisal of the scale

Tick appropriate	Ranking
Yes: - there are data on its use in studies beyond the group that developed the scale	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Writer's Cramp Rating Scale (WCRS)

1. Has the scale been used in dystonia?	Yes
2. Have clinimetric properties of this scale been evaluated in dystonia?	Only partially (inter-rater reliability)
Are there several versions of the scale? If so, which one has been assessed?	No
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if	The WCRS has 9 items, each with 3 response options. It is organized into 2 domains: part A (writing movement score) scores dystonic symptoms during writing and includes

applicable (e.g. for questionnaires, not tests)	subscores for dystonic posturing of the elbow, wrist, and fingers, the latency of dystonic symptoms, and the occurrence of writing tremor; part B (writing speed) scores the speed of handwriting. The higher the total WCRS score the more severe are dystonic symptoms during handwriting.
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Clinician; some additional instructions or training may be required because inter-rater reliability was modest for several items.
Copyright or public domain?	Public
How can the scale be obtained (address or website)?	Published ⁸² .
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	No

Scale properties

Content

Face validity?	Good (personal judgment).
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	It focuses only on the hand in writing tasks.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Yes, writer's cramp only.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	Yes, the scale assesses the severity of abnormal movements of the fingers, hand and arm.
Is disability measured?	Only one item for writing speed addresses disability.
Is health-related quality of life measured?	No
Is tremor evaluated?	There is one item that addresses tremor.

Does it exclude an important aspect of dystonia? If so, which?	It is relevant only for writer's cramp.
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No
Which areas are not covered?	This scale addresses mostly overt movements and is not likely to be sensitive to disability from stiffness or pain associated with writer's cramp

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	It was designed to address mostly movement severity.
Is there a cut-off score for diagnosis of dystonia?	No
Is this appropriate in dystonia? If not, why?	Yes, but it is limited and probably not sufficient on its own. Some additional measures of functional disability are needed.
Has it been <i>used</i> to measure severity or to screen for dystonia?	It was designed to address severity.

Acceptability

Length	Good. The WCRS can be completed in about 10 minutes ⁸² .
Clear instructions to raters?	Not clear; inter-rater reliability was poor for some items in the study by Wissel et al ⁸² .
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	The published scale does not define "Finger I", "Finger II" and "Finger III" (personal judgment).
Appropriateness of questions/items for dystonia?	Good (personal judgment).

Applicability across dystonia stages: mild – moderate – severe?	Good, but with only 3 options, sensitivity to change is likely to be limited (personal judgment).
---	---

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Not reported or assessed.
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Interrater reliability: good to modest for different items. Interrater reliability assessments for the four blinded raters yielded K coefficients of 0.4 - 0.85 documenting moderate (0.4 - 0.6) or substantial (>0.6) interrater agreement, for all but the items for elbow extension (0.25 – 0.39) and finger I and II extension (0.23 – 0.55) ⁸². In the study by Chen et al ⁸³, the interrater reliability for most subcomponents of the WCRS was in the good to moderate range. For the writer's cramp rating scale, the k coefficient for interrater reliability indicates good agreement for the wrist (k = 0.60) and tremor (k = 0.63) scores; moderate agreement for the finger II (k = 0.56) and latency of dystonia (k = 0.46) scores, and fair agreement for the finger I (k = 0.33) and finger III (k = 0.23) scores. Interrater reliability for the elbow score could not be assessed because none of the patients had dystonic posture at the elbow ⁸³.
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Not reported or assessed. Disability is not fully covered (personal judgment).
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	No gold standard
Construct validity	Poor correlations have been reported with

(correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	more generic arm disability scales. Zeuner KE et al ⁷⁵ rated the severity of writer's cramp using the Arm Dystonia Disability Scale (ADDS) and WCRS and correlated the clinical scores with movement kinematics. Kinematic measures and individual ADDS and WCRS scores did not correlate with each other. Authors conclude that this lack of correlation is not surprising as ADDS, WCRS, and kinematic analysis probe different aspects of motor impairment. Infact, the ADDS characterizes how dystonia affects a set of fine manual tasks, whereas the WCRS scores the manifestation of dystonia during handwriting. Therefore, the clinical scores and kinematic analysis of handwriting provide complementary insights into motor impairment. A correlation was found between the computer based analysis of writing speed and the WCRS subscores part A (dystonic posture, latency of dystonia, and writing tremor; $P < 0.05$) and part B (writing speed; $P < 0.01$) ⁸².
Valid in ethnically and culturally different populations?	Yes (personal judgment).
Valid in both genders and at all ages?	Applicable to both genders, but likely to have a lower age limit cut-off; never tested in children.
Valid in patients with dementia or significant cognitive impairment?	No (personal judgment).

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	No
---	----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Yes. The WCRS as assessed by blinded videotape independent raters showed significant

	decreases for subscores part A (dystonic posture, latency of dystonia, and writing tremor; $P < 0.001$) and B (writing speed; $P < 0.05$) after treatment with botulinum toxin compared with baseline, and these changes correlated with objective measurements of writing speed ⁸². The study by Chen et al ⁸³ confirmed that the WCRS can detect changes in patients following treatment with botulinum toxin. Zeuner et al ⁷⁹ examined 21 patients with writer's cramp that were randomly assigned to two types of re-training: a task specific training (writing movements using a pen attached to the bottom of a finger splint) and a non-task specific training (finger movements using therapeutic putty, but without writing exercises). Eight weeks of re-training improved writer's cramp as indexed by the WCRS ($P = 0.005$). Both types of re-training were equally effective. More severely affected patients benefited most. Taira T et al ^{84,85} performed stereotactic nucleus ventrooralis (Vo) thalamotomy for writer's cramp. The score on the WCRS decreased significantly ($p < 0.001$) after Vo thalamotomy.
Has the minimal clinically important change been assessed in patients with dystonia?	No

Overall impression

Advantages and disadvantages	This is a very simple scale for rating the severity of the movement abnormality in writer's cramp, but has only limited relevance for disability.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes,	It has been used successfully to measure outcomes in interventions, e.g. botulinum toxin and various surgeries.

clinical practice for diagnosis/screening). List all.	
---	--

Final appraisal of the scale

Tick appropriate	Ranking
Yes: - there are data on its use in studies beyond the group that developed the scale	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Generalized dystonia scales

Global dystonia rating scale (GDS)

1. Has the scale been used in dystonia?	☒ Yes ☐ No
2. Have psychometric properties of this scale been evaluated in dystonia?	☒ Yes ☐ No
Are there several versions of the scale?	No
If so, which one has been assessed?	
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	The GDS rates dystonia severity in 10 body areas: 1) eyes and upper face, 2) lower face, 3) jaw and tongue, 4) larynx, 5) neck, 6) shoulder and proximal arm, 7) distal arm and hand including elbow, 8) pelvis and upper leg, 9) distal leg and foot, and 10) trunk. Severity is rated in each body area on a scale with ratings from 0 to 10: 0 no dystonia present in that body area, 1 minimal dystonia, 5 moderate dystonia, and 10 most severe dystonia. There are no modifying ratings or weighting factors in the GDS. The total score is the sum of the scores for all the body areas. The maximal total score of the GDS is 140.

Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Designed to be administered by a clinician. No specific training is required.
Copyright or public domain?	Public domain.
How can the scale be obtained (address or website)?	The GDS is published in Comella et al.⁸⁶. http://www.mdvu.org/library/ratingscales/dystonia/gds.pdf
Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions psychometrically tested?	☐ Yes ☒ No

Scale properties

Content

Face validity?	Not found.
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	The GDS measures the severity of dystonia in individual body areas. Ten body areas are rated: 1) eyes and upper face, 2) lower face, 3) jaw and tongue, 4) larynx, 5) neck, 6) shoulder and proximal arm, 7) distal arm and hand including elbow, 8) pelvis and upper leg, 9) distal leg and foot, and 10) trunk.
Is the scale heavily weighted towards one type of dystonia? If so, which?	Generalized dystonia.
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	The GDS measure the severity of the symptoms of dystonia.

Is disability measured?	No.
Is health-related quality of life measured?	No.
Is tremor evaluated?	No.
Does it exclude an important aspect of dystonia? If so, which?	Not particularly. The GDS was designed as a simple direct assessment of overall severity of dystonia in each body area.
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No.
Includes items of sleep disturbance as a result of dystonia?	No.
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No.
Which areas are not covered?	The GDS does not have a duration rating, in contrast to the Unified Dystonia Rating Scale which has a severity and a duration rating.

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	The GDS was designed as a direct assessment of overall severity of dystonia in each body area.
Is there a cut-off score for diagnosis of dystonia?	No, but 0 indicates no dystonia present in that body area, 1 minimal dystonia, 5 moderate dystonia, and 10 most severe dystonia.
Is this appropriate in dystonia? If not, why?	NA.
Has it been <i>used</i> to measure severity or to screen for dystonia?	Severity.

Acceptability

Length	Its length is acceptable: it takes about 5-10 minutes to be completed (personal judgment). 82% of the rating investigators have found the GDS useful for measuring dystonia severity in an office setting ⁸⁶.
Clear instructions to raters?	The GDS is a tool simple and easy to use. In a study by Comella et al. ⁸⁶ 74% percent of the rating investigators found the GDS extremely or very easy to apply. Some difficulties have been found in the assessment of dystonia in some particular anatomical areas (larynx/speech, upper face/eyes).
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No.
Appropriateness of questions/items for dystonia?	Appropriate as designed specifically for dystonia. The items for the GDS scale consistently measure the domain of interest ⁸⁶.
Applicability across dystonia stages: mild – moderate – severe?	Yes. Comella et al. ⁸⁶ have used the scale in patients encompassing the full spectrum of dystonia severity.

Psychometric properties

Metric flaws (floor and ceiling effects, score distributions)?	No data found.
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	The GDS has been shown to have a high level of internal consistency, with Cronbach's alpha coefficient of 0.91. The GDS showed a high level of inter-rater reliability for the total scores, with the intraclass correlation coefficient of 0.72. The inter-rater

	agreement for individual items (body regions) was fair to excellent using both the Kendall's coefficient of concordance and the generalized weighted κ statistic (Kendall's 0.59–0.86; κ 0.58–0.86). The body regions showing the lowest level of inter-rater agreement were larynx and speech (Kendall's = 0.59), and upper face and eyes (κ = 0.58) ⁸⁶ .
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	Not assessed.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Not assessed.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Correlations with the FMDRS and UDRS have been studied The total scores for the three scales are highly correlated with each other. The scales had Pearson's correlations of 0.983 (UDRS and GDS), 0.977 (FMDRS and UDRS), and 0.980 (FMDRS and UDRS). Overall, 74% of raters found the GDS the easiest to apply ⁸⁶ .
Valid in ethnically and culturally different populations?	Not assessed.
Valid in both genders and at all ages?	Not assessed.
Valid in patients with dementia or significant cognitive impairment?	Not assessed.

Are you aware of any correlations of the scale with	No.
---	-----

biological markers in patients with dystonia?	
---	--

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	This scale has been used in studies of DBS for dystonia ⁸⁷ .
Has the minimal clinically important change been assessed in patients with dystonia?	Not assessed.

Overall impression

Advantages and disadvantages	<i>Advantages:</i> Appropriate for dystonia; the body regions are divided into smaller more defined areas; simple and easy to use; the items of the scale consistently measure the domain of interest; it was actually preferred by the raters for use in clinical practice compared to UDRS and FMDS. <i>Disadvantages:</i> Does not have a duration rating; difficulty in the assessment of dystonia in some particular anatomical areas (larynx/speech, upper face/eyes). The development of a teaching tape that demonstrates the ratings for each body area, especially face and eyes, will increase the inter-rater agreement for dystonia in these areas.
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	The GDS has been designed and used as a measure for dystonia severity ⁸⁶. The GDS with its simplicity and ease of application compared to the UDRS and the FMDS scales ⁸⁶, suggests that this method for rating of dystonia severity may be suitable for assessing change of severity during treatment (personal

	judgment). Tests for validity and responsivity to change are necessary to understand the clinical utility of this scale.
--	--

Final appraisal of the scale

Tick appropriate	Ranking
Yes: - there are data on its use in studies beyond the group that developed the scale	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Fahn-Marsden Dystonia Rating Scale (FMDRS)

1. Has the scale been used in dystonia?	☒ Yes ☐ No
2. Have clinimetric properties of this scale been evaluated in dystonia?	☒ Yes ☐ No
Are there several versions of the scale?	No
If so, which one has been assessed?	
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests)	Dystonia movement and disability subscales Movement scale: rating of dystonia severity and provoking factors in 9 body segments 7-item disability scale based on patient interview
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	Movement scale: Clinician rated Disability scale: Clinician rated Training for administration is recommended.
Copyright or public domain?	Public domain
How can the scale be obtained (address or website)?	Published ⁸⁸ . Multiple download sites

Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☐ No No validated translation
---	--

Scale properties

Content

Face validity?	Yes
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	All body segments
Is the scale heavily weighted towards one type of dystonia? If so, which?	Generalized dystonia
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	Severity and provoking factors
Is disability measured?	Yes
Is health-related quality of life measured?	No
Is tremor evaluated?	No
Does it exclude an important aspect of dystonia? If so, which?	No rating of mobile dystonia, no rating of dystonia tremor or myoclonus, insensitive to focal dystonia
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	No
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No
Which areas are not covered?	

Use

Was it <i>designed</i> to measure severity or to screen for dystonia?	Designed to quantify severity of dystonia
Is there a cut-off score for diagnosis of dystonia?	No
Is this appropriate in dystonia? If not, why?	
Has it been <i>used</i> to measure severity or to screen for dystonia?	Used in multiple clinical studies to assess severity

Acceptability

Length	10-20 minutes
Clear instructions to raters?	Yes
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	No
Appropriateness of questions/items for dystonia?	Good
Applicability across dystonia stages: mild – moderate – severe?	All stages, but multiple body parts should be affected

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Not known
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Intrarater 0.98-0.99; interrater: 0.85-0.96 (depending on training) ⁸⁸ . The FMDRS showed good internal consistency (Cronbach's alpha 0.89) and good level of inter-rater reliability for the total scores (intraclass correlation coefficients of 0.78). For separate items, interrater agreement was fair to excellent (Kendall's 0.55–0.87; κ 0.37–0.91) being lowest for eyes, jaw, face, and larynx. The modifying ratings for the FMDRS (Provoking Factor) showed consistently lower levels of agreement than motor severity ratings. The total scores for the FMDRS, the UDRS and the GDS were highly correlated with each other. The scales had

	Pearson's correlations of 0.983 (UDRS and GDS), 0.977 (FMDRS and UDRS), and 0.980 (FMDRS and UDRS) ⁸⁶ .
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity? Which areas are not covered?	not assessed
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	not assessed
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	0.69-0.98 movement scale compared to global dystonia rating ⁸⁸ .
Valid in ethnically and culturally different populations?	not assessed, but physician rating should be valid
Valid in both genders and at all ages?	yes
Valid in patients with dementia or significant cognitive impairment?	Movement scale: yes Disability scale: yes

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	no
---	----

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	Yes in pharmacological and surgical studies The FMDRS has been used in numerous studies to determine the treatment effects of deep brain stimulation (DBS) in adult and in children with primary or secondary dystonias. In the study by Kim et al ⁸⁹, the long-term surgical outcomes of secondary hemidystonia associated with post-traumatic brain injury was assessed. Burke-Fahn-Marsden Dystonia Rating Scale movement scores had improved by 73.2%

	(range, 38.1-94.1) and disability scores had improved by 75% (range, 60-100) at the 2-year follow-up visit. Results of DBS in children with primary and secondary dystonias are given in a recent study by Air EL ⁹⁰. At 1 year after implantation, patients with DYT1 dystonia had a mean improvement in the Burke-Fahn-Marsden Dystonia Rating Scale movement subscore of 75%, while those with secondary dystonia had only small improvements. In the study by Isaias et al ⁹¹, the FMDRS scores at baseline, 1 and 3 years after DBS were used to evaluate clinical outcome in patients with primary dystonia and bilateral pallidal DBS implants. All subjects showed a significant improvement after DBS implants (mean FMDRS improvement of 74.9% at 1 year and 82.6% at 3 years).
Has the minimal clinically important change been assessed in patients with dystonia?	No

Overall impression

Advantages and disadvantages	
Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	Appropriate for longitudinal and interventional studies of dystonia severity

Final appraisal of the scale

Tick appropriate	Ranking
Yes	Recommended Requirements are fulfilled if: - it has been applied to dystonia populations - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

Unified Dystonia Rating Scale (UDRS)

1. Has the scale been used in dystonia?	X Yes ☐ No This scale was developed by the Dystonia Study Group in a consensus conference in 1997, and was designed to include more body regions for rating, allowing more detailed evaluations of dystonia particularly in the limbs. The scale has been used previously in DBS trials in dystonia ⁹²⁻⁹⁵.
2. Have clinimetric properties of this scale been evaluated in dystonia?	☐ Yes ☐ No Clinimetric properties have been partially evaluated in dystonia
Are there several versions of the scale? If so, which one has been assessed?	One version.
Description of scale (content, number of items and subscales, number of response options (severity or frequency or both), addressed time frame (if applicable (e.g. for questionnaires, not tests))	Two subscales: a. Duration of dystonia for each body area I. 14 body areas, rated 0-4 in 0.5 increments b. 14 items for separate body areas. I. Rated 0-4 c. Total score is summation of the two scores. Maximal score is 112
Rater: Patient or caregiver or clinician. If clinician-rated, training for administration required?	All items are clinician rated.
Copyright or public domain?	No copyright
How can the scale be obtained (address or website)?	Available in the original manuscript ⁸⁶ .

Is the scale available in other languages? If so, in which languages? If so, how was the scale translated (forward-backward, one or more persons/teams)? Were translated versions clinimetrically tested?	☐ Yes ☒ No
---	---

Scale properties

Content

Face validity?	Face validity established
Is the scale focused on one body area or it encompasses dystonia affecting many body sites? Which body areas are rated?	Many body sites
Is the scale heavily weighted towards one type of dystonia? If so, which?	Generalized dystonia
Are the motor phenomena measured? If so, which aspects are rated (e.g., severity, duration, frequency, amplitude and position of movement, type of movement)?	Severity, duration, position
Is disability measured?	No
Is health-related quality of life measured?	No
Is tremor evaluated?	No
Does it exclude an important aspect of dystonia? If so, which?	Only assesses motor features. Does not rate any non-motor features.
Includes items of sensory abnormalities (e.g., pain) as a result of dystonia?	It excludes all the non-motor features of dystonia, including pain.
Includes items of sleep disturbance as a result of dystonia?	No
Includes items of neuropsychiatric features (e.g., anxiety, depression) as a result of dystonia?	No
Which areas are not covered?	All of the above.

Use

Was it <i>designed</i> to measure severity or to screen for	Measures severity of dystonia; is not a
---	---

dystonia?	screen for dystonia
Is there a cut-off score for diagnosis of dystonia?	NA
Is this appropriate in dystonia? If not, why?	NA
Has it been <i>used</i> to measure severity or to screen for dystonia?	Measure of severity.

Acceptability

Length	Somewhat long
Clear instructions to raters?	Clear instructions
Ambiguities in instructions to patient/rater? Ambiguities in rating anchors?	Rating levels are clear
Appropriateness of questions/items for dystonia?	Appropriate for dystonia
Applicability across dystonia stages: mild – moderate – severe?	Applicable in mild to severe

Clinimetric properties

Metric flaws (floor and ceiling effects, score distributions)?	Possible floor effect though not specifically addressed. Most scores grouped at lower ratings in clinimetric testing
Reliability (internal consistency, interrater, intrarater, test-retest) of total or sub-scales and of individual items	Good internal consistency Cronbach's alpha 0.93 and interrater reliability for total score. ICC 0.71. For separate items, agreement among raters, agreement for larynx, lower face and jaw and eyes reduced (weighted $\kappa < 0.75$). Duration with low agreement for arm in addition to the above. Not tested for intrarrater reliability, test retest.
Validity: assessed – not assessed, good – not good (references)	
Face and/or content validity?	Face validity and construct validity with other scales established in primary

Which areas are not covered?	dystonia for total score. ^{86,95} . Factor analysis not done.
Criterion validity (compared to gold-standard (Is there a gold standard?) or other criteria for diagnosis)	Not assessed for criterion validity.
Construct validity (correlations with other convergent scales and divergent scales (which?); factor analysis; known-groups comparisons).	Construct validity: Total score highly correlated with BFM and Global dystonia rating scale.
Valid in ethnically and culturally different populations?	Not assessed in ethnically and culturally different populations
Valid in both genders and at all ages?	Not assessed by gender or age (No DIF)
Valid in patients with dementia or significant cognitive impairment?	Not assessed in dementia

Are you aware of any correlations of the scale with biological markers in patients with dystonia?	
---	--

<i>Responsiveness / Sensitivity to change?</i>	
Demonstrated to be sensitive to change (change over time or to treatment) in patients with dystonia?	In DBS study, shown to improve with treatment similar to BFM ⁹⁵ .
Has the minimal clinically important change been assessed in patients with dystonia?	MCIC not determined

Overall impression

Advantages and disadvantages	UDRS allows for more detailed ratings of extremities than BFM. It is purely a clinician rated scale assessing motor features. Although good interrater reliability overall, less for larynx, lower face and eyes. Duration subscale also less reliable for upper limb. Some clinimetric analysis complete, but still requires additional revision, testing. Not clear if this is more informative than BFM.
------------------------------	---

Which type of study is this scale very suitable for and for which one is it unsuitable (screening, prevalence, aetiological (e.g. case-control or genetic), treatment trial of dystonia medication, correlation with biological markers or other scales, e.g. of dystonia syndromes, clinical practice for diagnosis/screening). List all.	
--	--

Final appraisal of the scale

Tick appropriate	Ranking
Yes - there are data on its use in studies beyond the group that developed the scale	Suggested Requirements are fulfilled if: - it has been applied to dystonia populations And one of both the following apply: - there are data on its use in studies beyond the group that developed the scale - it has been studied clinimetrically and found valid, reliable and sensitive to change

References

1. Fahn S. Rating scales for blepharospasm. Adv Ophthalm Plast Reconstr Surg 1985;4:97-101.
2. Jankovic J, Kenney C, Grafe S, Goertelmeyer R, Comes G. Relationship between various clinical outcome assessments in patients with blepharospasm. Mov Disord 2009;24:407-13.
3. Lindeboom R, de HR, Aramideh M, Speelman JD. The blepharospasm disability scale: an instrument for the assessment of functional health in blepharospasm. Mov Disord 1995;10:444-9.
4. Cano SJ, Warner TT, Linacre JM, Bhatia KP, Thompson AJ, Fitzpatrick R, et al. Capturing the true burden of dystonia on patients: the Cervical Dystonia Impact Profile (CDIP-58). Neurology 2004;63:1629-33.
5. Cano SJ, Hobart JC, Edwards M, Fitzpatrick R, Bhatia K, Thompson AJ, et al. CDIP-58 can measure the impact of botulinum toxin treatment in cervical dystonia. Neurology 2006;67:2230-2.
6. Zetterberg L, Aquilonius SM, Lindmark B. Impact of dystonia on quality of life and health in a Swedish population. Acta Neurol Scand 2009;119:376-82.
7. Jankovic J, Adler CH, Charles PD, Comella C, Stacy M, Schwartz M, et al. Rationale and design of a prospective study: Cervical Dystonia Patient Registry for Observation of OnaBotulinumtoxinA Efficacy (CD PROBE). BMC Neurol 2011;11:140.
8. Sethi KD, Rodriguez R, Olayinka B. Satisfaction with botulinum toxin treatment: a cross-sectional survey of patients with cervical dystonia. J Med Econ 2012;15:419-23.
9. Cano SJ, Warner TT, Thompson AJ, Bhatia KP, Fitzpatrick R, Hobart JC. The cervical dystonia impact profile (CDIP-58): can a Rasch developed patient reported outcome measure satisfy traditional psychometric criteria? Health Qual Life Outcomes 2008;6:58.

10. Consky E, Basinski A, Belle L, Ranawaya R, Lang A. The Toronto Western Spasmodic Torticollis Rating Scale (TWSTRS): assessment of validity and inter-rater reliability. *Neurology* 1990;40(Suppl. 1):445.
11. Consky ES, Lang AE. Clinical assessments of patients with cervical dystonia. In: Jankovic J, Hallett M, editors. *Therapy with botulinum toxin*. New York, NY: Marcel Dekker; 1994. p. 211-37.
12. Comella CL, Stebbins GT, Goetz CG, Chmura TA, Bressman SB, Lang AE. Teaching tape for the motor section of the Toronto Western Spasmodic Torticollis Scale. *Mov Disord* 1997;12:570-5.
13. Lindeboom R, Brans JW, Aramideh M, Speelman HD, De Haan RJ. Treatment of cervical dystonia: a comparison of measures for outcome assessment. *Mov Disord* 1998;13:706-12.
14. Tarsy D. Comparison of clinical rating scales in treatment of cervical dystonia with botulinum toxin. *Mov Disord* 1997;12:100-2.
15. Comella CL, Jankovic J, Truong DD, Hanschmann A, Grafe S. Efficacy and safety of incobotulinumtoxinA (NT 201, XEOMIN((R)), botulinum neurotoxin type A, without accessory proteins) in patients with cervical dystonia. *J Neurol Sci* 2011;308:103-9.
16. Jahanshahi M, Marsden CD. Body concept, disability, and depression in patients with spasmodic torticollis. *Behav Neurol* 1990;3:117-31.
17. Lewis L, Butler A, Jahanshahi M. Depression in focal, segmental and generalized dystonia. *J Neurol* 2008;255:1750-5.
18. Page D, Butler A, Jahanshahi M. Quality of life in focal, segmental, and generalized dystonia. *Mov Disord* 2007;22:341-7.
19. Jahanshahi M, Marsden CD. Psychological functioning before and after treatment of torticollis with botulinum toxin. *J Neurol Neurosurg Psychiatry* 1992;55:229-31.
20. Poewe W, Deuschl G, Nebe A, Feifel E, Wissel J, Benecke R, et al. What is the optimal dose of botulinum toxin A in the treatment of cervical dystonia? Results of a double blind, placebo controlled, dose ranging study using Dysport. German Dystonia Study Group. *J Neurol Neurosurg Psychiatry* 1998;64:13-7.
21. Tsui JK, Eisen A, Stoessl AJ, Calne S, Calne DB. Double-blind study of botulinum toxin in spasmodic torticollis. *Lancet* 1986;2:245-7.
22. Gelb DJ, Lowenstein DH, Aminoff MJ. Controlled trial of botulinum toxin injections in the treatment of spasmodic torticollis. *Neurology* 1989;39:80-4.
23. Moore AP, Blumhardt LD. A double blind trial of botulinum toxin "A" in torticollis, with one year follow up. *J Neurol Neurosurg Psychiatry* 1991;54:813-6.
24. Scheidt CE, Rayki O, Nickel T, Heinen F, Wissel J, Poewe W, et al. Spasmodic torticollis. A multicenter study on behavioural aspects I: introduction and methods. *Behav Neurol* 1996;9:25-31.
25. Scheidt CE, Schuller B, Rayki O, Kommerell G, Deuschl G. Relative absence of psychopathology in benign essential blepharospasm and hemifacial spasm. *Neurology* 1996;47:43-5.

26. Smania N, Corato E, Tinazzi M, Montagnana B, Fiaschi A, Aglioti SM. The effect of two different rehabilitation treatments in cervical dystonia: preliminary results in four patients. *Funct Neurol* 2003;18:219-25.
27. Munchau A, Palmer JD, Dressler D, O'Sullivan JD, Tsang KL, Jahanshahi M, et al. Prospective study of selective peripheral denervation for botulinum- toxin resistant patients with cervical dystonia. *Brain* 2001;124:769-83.
28. Folkman S, Lazarus RS. An analysis of coping in a middle-aged community sample. *J Health Soc Behav* 1980;21:219-39.
29. Felton BJ, Revenson TA, Hinrichsen GA. Stress and coping in the explanation of psychological adjustment among chronically ill adults. *Soc Sci Med* 1984;18:889-98.
30. Muller J, Wissel J, Kemmler G, Voller B, Bodner T, Schneider A, et al. Craniocervical dystonia questionnaire (CDQ-24): development and validation of a disease-specific quality of life instrument. *J Neurol Neurosurg Psychiatry* 2004;75:749-53.
31. Tepavcevic DK, Svetel M, Pekmezovic T, Petrovic I, Kostic VS. Craniocervical Dystonia Questionnaire (CDQ-24): validation and cross-cultural adaptation in Serbian patients. *Coll Antropol* 2009;33:1185-9.
32. Zetterberg L, Halvorsen K, Farnstrand C, Aquilonius SM, Lindmark B. Physiotherapy in cervical dystonia: six experimental single-case studies. *Physiother Theory Pract* 2008;24:275-90.
33. Merz RI, Deakin J, Hawthorne MR. Oromandibular dystonia questionnaire (OMDQ-25): a valid and reliable instrument for measuring health-related quality of life. *Clin Otolaryngol* 2010;35:390-6.
34. Stewart CF, Allen EL, Tureen P, Diamond BE, Blitzer A, Brin MF. Adductor spasmodic dysphonia: standard evaluation of symptoms and severity. *J Voice* 1997;11:95-103.
35. Sulica L, Blitzer A, Brin MF, Stewart CF. Botulinum toxin management of adductor spasmodic dysphonia after failed recurrent laryngeal nerve section. *Ann Otol Rhinol Laryngol* 2003;112:499-505.
36. Blitzer A, Brin MF, Stewart CF. Botulinum toxin management of spasmodic dysphonia (laryngeal dystonia): a 12-year experience in more than 900 patients. *Laryngoscope* 1998;108:1435-41.
37. Lee L, Daughton S, Scheer S, Stemple JC, Weinrich B, Miller-Seiler T, et al. Use of acupuncture for the treatment of adductor spasmodic dysphonia: a preliminary investigation. *J Voice* 2003;17:411-24.
38. Jacobson BH, Johnson A, Grywalski C, Silbergleit A, Jacobson G, Benninger MS, et al. The Voice Handicap Index (VHI): development and validation. *Am J Speech Lang Pathol* 1997;6:66-70.
39. Franic DM, Bramlett RE, Bothe AC. Psychometric evaluation of disease specific quality of life instruments in voice disorders. *J Voice* 2005;19:300-15.
40. Kaptein AA, Hughes BM, Scharloo M, Hondebrink N, Langeveld TP. Psychological aspects of adductor spasmodic dysphonia: a prospective population controlled questionnaire study. *Clin Otolaryngol* 2010;35:31-8.
41. Sulica L, Louis ED. Clinical characteristics of essential voice tremor: a study of 34 cases. *Laryngoscope* 2010;120:516-28.

42. Rosen CA, Murry T, Zinn A, Zullo T, Sonbolian M. Voice handicap index change following treatment of voice disorders. *J Voice* 2000;14:619-23.
43. Biddle AK, Watson LR, Hooper CR, Lohr KN, Sutton SF. Criteria for determining disability in speech-language disorders. *Evid Rep Technol Assess (Summ)* 2002;(52):1-4.
44. Alves CB, Loughran S, MacGregor FB, Dey JI, Bowie LJ. Bioplastique medialization therapy improves the quality of life in terminally ill patients with vocal cord palsy. *Clin Otolaryngol Allied Sci* 2002;27:387-91.
45. Stewart MG, Chen AY, Stach CB. Outcomes analysis of voice and quality of life in patients with laryngeal cancer. *Arch Otolaryngol Head Neck Surg* 1998;124:143-8.
46. Benninger MS, Ahuja AS, Gardner G, Grywalski C. Assessing outcomes for dysphonic patients. *J Voice* 1998;12:540-50.
47. Wuyts FL, De Bodt MS, Molenberghs G, Remacle M, Heylen L, Millet B, et al. The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. *J Speech Lang Hear Res* 2000;43:796-809.
48. Baylor CR, Yorkston KM, Eadie TL, Miller RM, Amtmann D. Developing the communicative participation item bank: Rasch analysis results from a spasmodic dysphonia sample. *J Speech Lang Hear Res* 2009;52:1302-20.
49. Morzaria S, Damrose EJ. A comparison of the VHI, VHI-10, and V-RQOL for measuring the effect of botox therapy in adductor spasmodic dysphonia. *J Voice* 2012;26:378-80.
50. Spector BC, Netterville JL, Billante C, Clary J, Reinisch L, Smith TL. Quality-of-life assessment in patients with unilateral vocal cord paralysis. *Otolaryngol Head Neck Surg* 2001;125:176-82.
51. Benninger MS, Gardner G, Grywalski C. Outcomes of botulinum toxin treatment for patients with spasmodic dysphonia. *Arch Otolaryngol Head Neck Surg* 2001;127:1083-5.
52. Novakovic D, Waters HH, D'Elia JB, Blitzer A. Botulinum toxin treatment of adductor spasmodic dysphonia: longitudinal functional outcomes. *Laryngoscope* 2011;121:606-12.
53. Chhetri DK, Mendelsohn AH, Blumin JH, Berke GS. Long-term follow-up results of selective laryngeal adductor denervation-reinnervation surgery for adductor spasmodic dysphonia. *Laryngoscope* 2006;116:635-42.
54. Tsuji DH, Chrispim FS, Imamura R, Sennes LU, Hachiya A. Impact in vocal quality in partial myectomy and neurectomy endoscopic of thyroarytenoid muscle in patients with adductor spasmodic dysphonia. *Braz J Otorhinolaryngol* 2006;72:261-6.
55. Hussain A, Shakeel M. Selective lateral laser thyroarytenoid myotomy for adductor spasmodic dysphonia. *J Laryngol Otol* 2010;124:886-91.
56. Tsuji DH, Takahashi MT, Imamura R, Hachiya A, Sennes LU. Endoscopic laser thyroarytenoid myoneurectomy in patients with adductor spasmodic dysphonia: a pilot study on long-term outcome on voice quality. *J Voice* 2012;26:666-12.
57. Deary IJ, Webb A, MacKenzie K, Wilson JA, Carding PN. Short, self-report voice symptom scales: psychometric characteristics of the voice handicap index-10 and the vocal performance questionnaire. *Otolaryngol Head Neck Surg* 2004;131:232-5.

58. Rosen CA, Lee AS, Osborne J, Zullo T, Murry T. Development and validation of the voice handicap index-10. *Laryngoscope* 2004;114:1549-56.
59. Zur KB, Cotton S, Kelchner L, Baker S, Weinrich B, Lee L. Pediatric Voice Handicap Index (pVHI): a new tool for evaluating pediatric dysphonia. *Int J Pediatr Otorhinolaryngol* 2007;71:77-82.
60. Johnson K, Brehm SB, Weinrich B, Meinzen-Derr J, de Alarcon A. Comparison of the Pediatric Voice Handicap Index with perceptual voice analysis in pediatric patients with vocal fold lesions. *Arch Otolaryngol Head Neck Surg* 2011;137:1258-62.
61. de Alarcon A, Brehm SB, Kelchner LN, Meinzen-Derr J, Middendorf J, Weinrich B. Comparison of pediatric voice handicap index scores with perceptual voice analysis in patients following airway reconstruction. *Ann Otol Rhinol Laryngol* 2009;118:581-6.
62. Boseley ME, Cunningham MJ, Volk MS, Hartnick CJ. Validation of the Pediatric Voice-Related Quality-of-Life survey. *Arch Otolaryngol Head Neck Surg* 2006;132:717-20.
63. Geneid A, Pakkasjarvi N, Aherto A, Roine R, Sintonen H, Lindahl H, et al. Outcomes of early infancy laryngeal reconstruction on health- and voice-related quality of life. *Int J Pediatr Otorhinolaryngol* 2011;75:351-5.
64. Maturo S, Tse SM, Kinane TB, Hartnick CJ. Initial experience using propranolol as an adjunctive treatment in children with aggressive recurrent respiratory papillomatosis. *Ann Otol Rhinol Laryngol* 2011;120:17-20.
65. Tirado Y, Chadha NK, Allegro J, Forte V, Campisi P. Quality of life and voice outcomes after thyroid ala graft laryngotracheal reconstruction in young children. *Otolaryngol Head Neck Surg* 2011;144:770-7.
66. Hogikyan ND, Sethuraman G. Validation of an instrument to measure voice-related quality of life (V-RQOL). *J Voice* 1999;13:557-69.
67. Hogikyan ND, Wodchis WP, Spak C, Kileny PR. Longitudinal effects of botulinum toxin injections on voice-related quality of life (V-RQOL) for patients with adductory spasmodic dysphonia. *J Voice* 2001;15:576-86.
68. Webb AL, Carding PN, Deary IJ, MacKenzie K, Steen IN, Wilson JA. Optimising outcome assessment of voice interventions, I: Reliability and validity of three self-reported scales. *J Laryngol Otol* 2007;121:763-7.
69. Carding PN, Horsley IA, Docherty GJ. A study of the effectiveness of voice therapy in the treatment of 45 patients with nonorganic dysphonia. *J Voice* 1999;13:72-104.
70. MacKenzie K, Millar A, Wilson JA, Sellars C, Deary IJ. Is voice therapy an effective treatment for dysphonia? A randomised controlled trial. *BMJ* 2001;323:658-61.
71. Rattenbury HJ, Carding PN, Finn P. Evaluating the effectiveness and efficiency of voice therapy using transnasal flexible laryngoscopy: a randomized controlled trial. *J Voice* 2004;18:522-33.
72. Steen IN, MacKenzie K, Carding PN, Webb A, Deary IJ, Wilson JA. Optimising outcome assessment of voice interventions, II: Sensitivity to change of self-reported and observer-rated measures. *J Laryngol Otol* 2008;122:46-51.

73. Meek P, Carding PN, Howard DH, Lennard TW. Voice change following thyroid and parathyroid surgery. *J Voice* 2008;22:765-72.
74. Fahn S. Assessment of the primary dystonias. In: Munsat T, editor. *The quantification of neurological deficit*. Boston: Butterworths; 1989. p. 241-70.
75. Zeuner KE, Peller M, Knutzen A, Holler I, Munchau A, Hallett M, et al. How to assess motor impairment in writer's cramp. *Mov Disord* 2007;22:1102-9.
76. Jabusch HC, Vauth H, Altenmuller E. Quantification of focal dystonia in pianists using scale analysis. *Mov Disord* 2004;19:171-80.
77. Priori A, Pesenti A, Cappellari A, Scarlato G, Barbieri S. Limb immobilization for the treatment of focal occupational dystonia. *Neurology* 2001;57:405-9.
78. Trompetto C, Avanzino L, Bove M, Marinelli L, Molfetta L, Trentini R, et al. External shock waves therapy in dystonia: preliminary results. *Eur J Neurol* 2009;16:517-21.
79. Zeuner KE, Peller M, Knutzen A, Hallett M, Deuschl G, Siebner HR. Motor re-training does not need to be task specific to improve writer's cramp. *Mov Disord* 2008;23:2319-27.
80. Candia V, Elbert T, Altenmuller E, Rau H, Schafer T, Taub E. Constraint-induced movement therapy for focal hand dystonia in musicians. *Lancet* 1999;353:42.
81. Tubiana R, Chamagne P. Les affections professionnelles du membre supérieur chez les musiciens. *Bull Acad Natl Med* 1993;177:203-12.
82. Wissel J, Kabus C, Wenzel R, Klepsch S, Schwarz U, Nebe A, et al. Botulinum toxin in writer's cramp: objective response evaluation in 31 patients. *J Neurol Neurosurg Psychiatry* 1996;61:172-5.
83. Chen R, Karp BI, Goldstein SR, Bara-Jimenez W, Yaseen Z, Hallett M. Effect of muscle activity immediately after botulinum injection for writer's cramp. *Mov Disord* 1999;14:307-12.
84. Taira T, Harashima S, Hori T. Neurosurgical treatment for writer's cramp. *Acta Neurochir Suppl* 2003;87:129-31.
85. Taira T, Hori T. Stereotactic ventrooralis thalamotomy for task-specific focal hand dystonia (writer's cramp). *Stereotact Funct Neurosurg* 2003;80:88-91.
86. Comella CL, Leurgans S, Wu J, Stebbins GT, Chmura T. Rating scales for dystonia: a multicenter assessment. *Mov Disord* 2003;18:303-12.
87. Singh A, Kammermeier S, Mehrkens JH, Botzel K. Movement kinematic after deep brain stimulation associated microlesions. *J Neurol Neurosurg Psychiatry* 2012;83:1022-6.
88. Burke RE, Fahn S, Marsden CD, Bressman SB, Moskowitz C, Friedman J. Validity and reliability of a rating scale for the primary torsion dystonias. *Neurology* 1985;35:73-7.
89. Kim JP, Chang WS, Chang JW. The long-term surgical outcomes of secondary hemidystonia associated with post-traumatic brain injury. *Acta Neurochir (Wien)* 2012;154:823-30.
90. Air EL, Ostrem JL, Sanger TD, Starr PA. Deep brain stimulation in children: experience and technical pearls. *J Neurosurg Pediatr* 2011;8:566-74.

91. Isaías IU, Volkmann J, Kupsch A, Burgunder JM, Ostrem JL, Alterman RL, et al. Factors predicting protracted improvement after pallidal DBS for primary dystonia: the role of age and disease duration. *J Neurol* 2011;258:1469-76.
92. Krauss JK, Lohr TJ, Weigel R, Capelle HH, Weber S, Burgunder JM. Chronic stimulation of the globus pallidus internus for treatment of non-dYT1 generalized dystonia and choreoathetosis: 2-year follow up. *J Neurosurg* 2003;98:785-92.
93. Diamond A, Shahed J, Azher S, Dat-Vuong K, Jankovic J. Globus pallidus deep brain stimulation in dystonia. *Mov Disord* 2006;21:692-5.
94. Pretto TE, Dalvi A, Kang UJ, Penn RD. A prospective blinded evaluation of deep brain stimulation for the treatment of secondary dystonia and primary torticollis syndromes. *J Neurosurg* 2008;109:405-9.
95. Susatia F, Malaty IA, Foote KD, Wu SS, Zeilman PR, Mishra M, et al. An evaluation of rating scales utilized for deep brain stimulation for dystonia. *J Neurol* 2010;257:44-58.