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.

Key Words: botulinum neurotoxins; blepharospasm; dystonias; torticollis; rating scales
Dystonia is one of the most common movement disorders, with an overall prevalence of 16.43 per 100,000 for primary dystonia. 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. 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, 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. 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. Treatment possibilities have greatly expanded in recent years after discovering the efficacy of botulinum neurotoxins and functional surgery. 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 centered on 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. This terminology was also used in recent reviews on rating scales of other MDS task forces. 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 100, 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.

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 subscales 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 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
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.\textsuperscript{14,15,19} The CDIP-58 is good at detecting the impact of BoNT on all 8 health dimensions in patients with cervical dystonia.\textsuperscript{15}

**Strengths and Weaknesses.** The CDIP-58 is a diseasespecific 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 themselves\textsuperscript{15} 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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{20} The TWSTRS scale also showed strong correlation with Tsui scale.\textsuperscript{59} Responsiveness to change has been demonstrated.\textsuperscript{55,59}

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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{60} The CDQ-24 has also been translated and validated in Serbian.\textsuperscript{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,\textsuperscript{16} and also as a HRQoL measure to assess responsiveness to treatment-induced changes.\textsuperscript{60,62}

Clinimetric Properties. There were no relevant ceiling effects, but a considerable floor effect was observed in the social/family life domain.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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.\textsuperscript{64} The VHI has also been 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.\textsuperscript{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.\textsuperscript{65–75}

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

**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. In a study that included patients with laryngeal dystonia, 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. 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 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.

**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, including childhood-onset dystonia.

**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. The FMDRS showed good internal consistency and good level of interrater reliability for the total scores. For separate items, interrater agreement was fair to good, 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 Unified Dystonia Rating Scale, and the Global Dystonia Rating Scale were highly correlated with each other. Responsiveness has been demonstrated in treatment studies.84,91,110

Strengths and Weaknesses. Limitations in the FMDRS include a weighting factor that halves the contribution of dystonia in eyes, mouth, and neck to the total score. The FMDRS does not assess in detail the individual body areas, such as separate ratings for proximal and distal limbs; moreover, included in the FMDRS 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 (FMDRS). 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 FMDRS) 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. 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 confused 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, 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


