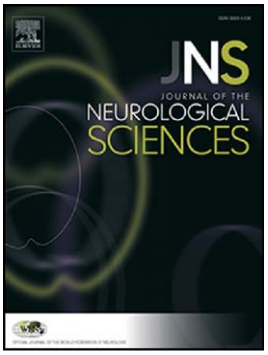
Scales for hyperkinetic disorders: A systematic review

Pietracupa, Sara; Bruno, Elisa; Cavanna, Andrea Eugenio; Falla, Marika; Zappia, Mario; Colosimo, Carlo

DOI:
10.1016/j.jns.2015.08.1544



Scales for hyperkinetic disorders: a systematic review

Sara Pietracupa, Elisa Bruno, Andrea E. Cavanna, Marika Falla, Mario Zappia, Carlo Colosimo

PII: S0022-510X(15)02040-7
DOI: doi: 10.1016/j.jns.2015.08.1544
Reference: JNS 14052

To appear in: *Journal of the Neurological Sciences*

Received date: 8 June 2015
Revised date: 21 August 2015
Accepted date: 28 August 2015

Please cite this article as: Sara Pietracupa, Elisa Bruno, Andrea E. Cavanna, Marika Falla, Mario Zappia, Carlo Colosimo, Scales for hyperkinetic disorders: a systematic review, *Journal of the Neurological Sciences* (2015), doi: 10.1016/j.jns.2015.08.1544

Scales for hyperkinetic disorders: a systematic review

Sara Pietracupa (1), Elisa Bruno (2), Andrea E. Cavanna (3,4,5), Marika Falla (1), Mario Zappia (2), Carlo Colosimo (1)

[1] *Department of Neurology and Psychiatry, Sapienza University of Rome, Italy*

2 Dipartimento di Scienze Mediche, Chirurgiche e Tecnologie Avanzate "G.F. Ingrassia", University of Catania, Via Santa Sofia 78, 95123 Catania, Italy.

[3] *Department of Neuropsychiatry, BSMHFT and University of Birmingham, United Kingdom*

[4] *School of Life and Health Sciences, Aston University, Birmingham, United Kingdom*

[5] *Sobell Department of Motor Neuroscience and Movement Disorders, Institute of Neurology, University College London, United Kingdom*

Running Title
Scales for hyperkinetic disorders

Word count: 8089 (text only)
Number of tables: 12

Correspondence to:

Carlo Colosimo, MD

Dipartimento di Neurologia e Psichiatria

Sapienza Università di Roma,

Viale dell'Università 30, 00185 Roma, Italy

Email: carlo.colosimo@uniroma1.it

Fax +390649979526

Abstract

Hyperkinetic movement disorders represent a heterogeneous group of disorders in which involuntary movements are the prevalent clinical symptoms. The five main categories of hyperkinetic disorders are tremor, dystonia, tics, myoclonus and drug-induced dyskinesia. The severity of hyperkinetic disorders is assessed by all clinicians when they examine a patient; quantifying the severity also provides a means of studying the natural history of a given disorder and the possible effect of new therapeutic interventions. This means that good rating instruments are required in both everyday practice and experimental settings. Unfortunately, the clinical evaluation of these disorders is complicated by the inherent nature and variability over time of involuntary movements. A number of scales have been proposed over the years to study the various hyperkinetic disorders. The aim of this review is to systematically identify all the clinical scales that have been proposed and to classify them according to the criteria developed by the Movement Disorder Society (MDS) task force for rating scales in PD. On the basis of this methodology, a scale may be defined as 'Recommended', 'Suggested' or 'Listed' in decreasing order of value.

We found that, although numerous scales aimed at assessing hyperkinetic disorders have been published, their variability in terms of clinimetric properties, availability and effort required to administer them is high. In this evaluation, we identified scales defined as 'Recommended' for the assessment of all forms of hyperkinetic disorders. The situation highlighted by our analysis varies considerably, with several 'Recommended' scales being available for some conditions such as tics or dystonia, but only one being available for myoclonus. This gap needs to be filled by the scientific community through both the development of new clinical tools and the refinement of existing ones.

Introduction

Hyperkinetic movement disorders represent a heterogeneous group of disorders in which unwanted (involuntary) movements are the prevalent clinical symptoms. These disorders are usually linked to basal ganglia dysfunction (Abdo et al, 2010). The five main categories of dyskinesia are tremor, dystonia, tics, myoclonus and drug-induced dyskinesia.

The severity of hyperkinetic disorders is assessed by all clinicians when they examine a patient. Quantifying the severity also provides a means of studying the natural history of a given disorder and the possible effect of new therapeutic interventions. In recent decades, a number of mechanical and electronic devices, including accelerometers, have been developed jointly by physicians and engineers to measure involuntary movements; more recently, computerized devices have also been designed (Mansur et al, 2007). The advantages of mechanical and electronic measurements are objectivity and consistency even when they are performed by different clinicians. However, as these measurements do not appear be as sensitive as clinical measurements, hyperkinetic disorders continue to be assessed largely by clinical methods. This means that good rating instruments are required in both everyday practice and experimental settings. Unfortunately, the clinical evaluation of these disorders is complicated by the inherent nature and variability over time of involuntary movements. A number of scales have been proposed over the years to study the various hyperkinetic disorders. The aim of this review is to systematically identify all the clinical scales that have been proposed and to classify them according to the criteria developed by the Movement Disorder Society (MDS) task force for rating scales in Parkinson's Disease (PD) (Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease, 2003). The systematic review of the PD rating scales carried out by this task force was conducted according to an established methodology (Goetz et al, 2008). This process includes scale identification, selection and appraisal strategies, using terminology and definitions developed *ad hoc* (Goetz et al, 2008) . On the basis of this methodology, a scale is defined as 'Recommended' if it has been applied to that specific disease population, if there are data on its use in studies other than those collected by the group that developed the scale, and if it has been studied clinimetrically and found to be valid, reliable and sensitive to change. A scale is defined as 'Suggested' if it has been applied to specific populations, but only one of the other criteria applies. A scale is defined simply as 'Listed' if it satisfies only one of the three criteria used to define 'Recommended' scales. Owing to the relative lack of proven treatments for hyperkinetic disorders, the clinimetric criterion for rating

dyskinesia scales does not categorically require responsiveness to be established. Indeed, if a scale fulfills the requirements of reliability and validity, the criterion is considered to be satisfied, although the absence of responsiveness is noted as a weakness of that scale. This classification has been successfully used to assess the validity of the scales used for both motor and non-motor aspects of PD (Colosimo et al, 2010).

This review will follow the same pattern for each one of the conditions studied. Only published or in press peer-reviewed papers or published abstracts form main neurological meetings were evaluated. The Medline database on PubMed was searched for relevant papers and all the scales used to measure a given disorder were identified (as of Medline last accessed on the 10th of November 2014) using the following query: ''Hyperkinetic disorders'' AND 'Assessment' OR 'Scales' OR 'Questionnaires'. For each scale, a search was conducted for the following terms 'Tremor' AND 'Assessment' OR 'Scales' OR 'Questionnaires', 'Dystonia' AND 'Assessment' OR 'Scales' OR 'Questionnaires', 'Chorea' AND 'Assessment' OR 'Scales' OR 'Questionnaires', 'Tics' AND 'Assessment' OR 'scales' OR 'Questionnaires', 'Myoclonus' AND 'assessment' OR 'scales' OR 'Questionnaires', 'Drug-induced dyskinesias' AND 'Assessment' OR 'Scales' OR 'Questionnaires'. In addition for each scale, a search was conducted for the terms 'Tremor', 'Dystonia', 'Chorea', 'Tics', 'Myoclonus', 'Drug induced dyskinesias' and the name of the scale. All scales have been reported in a specific table, though only those defined as 'Recommended' according to the aforementioned criteria will be appraised and discussed in detail in the following text.

Tremor

Twelve scales have been assessed for tremor evaluation (Fahn et al, 1993; Bain et al, 1993; Louis et al, 1997, Elbe et al, 2012; Jankovich et al, 1996; Bove' et al, 2006; Sweet et al, 1974; Findley et al, 1985; Koller et al, 1989; Jefferson et al, 1979; Baruzzi et al, 1983; Ogawa et al, 1987) (Table 1), but only two of them, the Fahn-Tolosa-Marin Tremor rating scale (FTM-TRS) and the Washington Heights-Inwood Genetic Study of Essential Tremor (WHIGET), reached recommendation status (Table 2).

Fahn-Tolosa-Marin Tremor rating scale

Scale description

The Tremor Rating Scale (TRS) was developed by Fahn, Tolosa and Marin to quantify rest, postural and action/intention tremors (Fahn et al, 1993). This scale also evaluates voice tremor as well as handwriting and other tasks, such as hygiene and dressing, assessing the impact of tremor on patients' daily life. The TRS is divided into three parts (A, B and C), yielding a subtotal score that can be combined in an overall TRS score. Part A (scores 1-10) quantifies tremor at rest, with posture holding, and performing an activity, for nine parts of the body (face, tongue, voice, head, upper and lower limbs and trunk), and orthostatic tremor. Severity of tremor is rated by amplitude on a 5-point scale. Part B (scores 1-15) relates to action tremors of the upper extremities, particularly writing and pouring liquids. Drawing tasks require Archimedes' spirals (two types) and draw a straight line between narrow confine (three times). Writing and drawing tremor are scored on a 5-point scale where pouring water from one cup to another is also quantified. Cup size and the amount of water used in the test are specified. Part C assesses functional disability. Its items evaluated the severity of tremor with speaking, eating (feeding), bringing liquids to the mouth, hygienic care, dressing, working, including domestic tasks, and social activities. These scores, with the exception of speaking, are provided by patients, who are asked to evaluated their ability to carry out these tasks by using provided definitions (score 0 to 4). The maximum possible scores are 88 for part A, 36 for part B and 32 for part C, making the maximum possible total score of 156. Finally, in addition to the quantification of tremor in parts A, B, and C, the scoring form allows assessment of overall severity by both the patient and the examiner (global assessment) on a 5-point scale (0 to 4).

Key evaluation issues

The interrater and intrarater reliability were evaluated for part A and B (excluding pouring water) in Essential Tremor (ET) patients (Stacy et al, 2007). This evaluation revealed fair (part A; modified Kappa ranging from 0.53-0.62) to poor (part B; modified Kappa ranging from 0.17-0.41) interrater reliability. Intrarater reliability calculated with Spearman's correlation, indicated a good consistency (overall Spearman's correlation= 0.87), although still not as good for part B. Inter and intrarater reliability were separated assessed for neurologists and non-neurologists, showing a lower consistency for the latters. Interrater reliability was also measured in patients with Multiple Sclerosis tremor (Hooper et al, 1998), showing a reliability coefficients ranging from 0.69 to 0.99, calculated via Kendall's

coefficient of concordance to determine overall agreement. The intrarater reliability coefficients (Spearman's correlation coefficient) ranged from 0.85 to 0.97 for the different categories of head tremor, 0.64 to 0.93 for trunk tremor, 0.92 to 0.99 for the right upper limb tremor, 0.81 to 0.99 for the left upper limb tremor, and 0.87 to 1 for the tremor evident when performing upper limb tasks (spiralography and volumetric test). Levels of reliability were high except when certain categories of tremor in the trunk were assessed (postural tremor r = 0.64 and goal related r = 0.72).

Strengths and Weaknesses

The scale appears very similar to another scale to assess tremor, designed by the Tremor Investigation Group: the UTRA scale (Unified Tremor Rating and Assessment, 6). In fact, the TRS was developed incorporating some items used in the UTRA scale to writing and drawing tasks and to the assessment of functional disability. It showed fair-to-poor concordance between raters, especially for non-neurologist, and good intrarater reliability in repeat assessment, although not as good for part B. The scale provides specific instructions to standardize the evaluations, specific definitions for tremor severity assessment and words anchors for the scores. It is a quite complex instrument requiring some time for its use, and can be repeated over time.

Washington Heights-Inwood Genetic Study of Essential Tremor rating scale

Scale description

The Washington Heights-Inwood Genetic Study of Essential Tremor (WHIGET) scale was developed for a community-study on ET (Louis et al, 1997). It is a detailed examination of upper limbs tremor that incorporate multiple test items, in addition to writing and a clinical rating scale. The WHIGET tremor rating scale consists of a six-test 10-minute tremor examination designed to elicit rest tremor (two positions), postural tremor (one test: sustained arm extension), and kinetic tremor (five tests: pouring water between two cups, drinking water from a cup, using a spoon to drink water, finger-to-nose movements, and drawing Archimedes spirals). Each task is first performed with the dominant arm and then with the nondominant arm. The ratings are as follows: 0= no visible tremor, 1= a low amplitude tremor that is barely perceivable or is intermittent, 2= tremor that fulfils three criteria: moderate amplitude and usually present and clearly oscillatory (clearly varies

between alternate extremes with a definable period), and 3= large amplitude, jerky tremor resulting in spilling, and difficulty hitting a target. For kinetic tremor, a rating of 4 was subsequently added to broaden the applicability of this scale to clinical trials. This new rating signified an extremely large amplitude, jerky kinetic tremor resulting in inability to handle liquids, reluctance to touch finger to nose because of fear of self-injury, or inability to draw a spiral. The WHIGET performance-based test of function is a 15-items, 10-minute test developed for the "Columbia University Assessment of Disability In Essential Tremor" study. Each task is performed with the dominant arm and rated according to a 0-to-4 point rating scale (0= no difficulty; 1= mild difficulty; 2= moderate difficulty; 3= severe difficulty; 4= unable to perform the task). A shorter 6-item 5-minute version was also proposed, containing activities that are frequently performed: place keys in lock, drink from a glass, copy sentences, place bills in wallet, dial number on telephone, place coins in slot.

Key evaluation issues

Inter-rater agreement for this scale was assessed in ratings of postural and kinetic tremor (weighted kappa= 0.62–0.78) (Louis et al, 1998). There was also a high degree of test–retest stability when the two neurologists rated tremor at two separate time intervals ($r = 0.98$, $P < 0.00001$) as well as high validity (Louis et al, 1998). A teaching videotape was then developed, accompanied by written instructions. Inter-rater reliability was assessed in raters, after viewing this teaching videotape, for five of the six test items (drawing a spiral, pouring, sustained arm extension, drinking, finger-to-nose maneuver), agreement among raters was nearly perfect or perfect (weighted kappa= 0.86-1), and for the sixth (using a spoon), it was substantial (weighted kappa= 0.79) (Louis et al, 2001). The validity of the 15-item WHIGET performance-based test was tested among ET cases (Louis et al, 1999). The total score of the scale correlated with the total number of questions answered yes ($r=0.44$;$P=0.001$) of a validated 12-items screening questionnaire for ET (Louis et al, 1998); the total score of a 31-item tremor disability questionnaire to assess the functional impact of tremor (Fried et al, 1996). However, before the registration of this videotape, the WHIGET Tremor Rating Scale was revised (adding a score of 4 for kinetic tremor), and it cannot be assumed that the intra-rater reliability and validity of the revised scale is identical to those of the original version. A correlation with this scale to a quantitative measure (motion sensor) showed a good correlation for postural ($r=0.90$) and kinetic ($r=0.80$) tremors in ET patients.

Strengths and weaknesses

The WHIGET are all easily administered, simple, and user friendly, requiring 10 minutes each (5 minutes the WHIGET performance-based scale containing 6 items). Clinimetric properties were assessed and appeared good, demonstrating a good validity, consistency and reproducibility. However, the performance-test scale assess only dominant-arm functions and is influences by kinetic tremor, ignoring tremor in other body parts (nondominant arm, legs, head, voice) and rest and postural tremor. WHIGET has been used by the same authors and by others groups in a prevalence survey on ET (Sur et al, 2009) and in a pilot study on memantine treatment for ET (Handforth, 2010).

Dystonia

Thirty scales have been assessed for evaluation of different types of dystonia (Jankovic et al, 2009; Cano et al, 2004; Consky et al, 1994; Muller et al, 2004; Jacobson et al, 1997; Carding et al, 1999, Burke et al, 1985; Troung et al, 2013; Lindeboomet al, 1995; Page et al, 2007; Tsui et al, 1985, 1986, 1987; Fernandez et al, 2013; Münchau et al, 2001, Merz et al, 2010; Stewart et al, 1997; Morzaria et al, 2012; Jabush et al, 2004; Priori et al, 2001; Comella et al, 2003) (Table 3). Seven of them, the Blepharospasm Disability Index (BDI), the Cervical Dystonia Impact Scale, the Toronto Western Spasmodic Torticollis Rating Scale, the Craniocervical Dystonia Questionnaire (CDIS), the Voice Handicap Index (VHI), the Vocal Performance Questionnaire (VPQ) and the Fahn-Marsden Dystonia Rating Scale (FMDRS), reached the recommendation status (Albanese et al, 2013) (Table 4).

Blepharospasm Disability Index

Scale description

The Blepharospasm Disability Index (BSDI) (Jankovic et al, 2009) was developed to improve the Blepharospasm Disability Scale 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). The range of scores is 0 to 24, with higher

scores indicating a greater disability.

Key evaluation issues

The BSDI showed high internal consistency and the retest reliability of the single items was adequate (Jankovic et al, 2009).

Strengths and Weaknesses

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. Concern has been raised regarding poor sensitivity of the scale to mild disability for small changes.

Cervical Dystonia Impact Scale

Scale description.

This scale 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

Key issues

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 have been supported (Cano et al, 2006; Zetterberg et al, 2009; Cano et al, 2008).

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.

Toronto Western Spasmodic Torticollis Rating Scale

Scale description

The Toronto Western Spasmodic Torticollis Rating Scale (TWSTRS is composed of 3 subscales that measure symptom severity, disability, and pain. The clinician-rated severity scale 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 as refinement of the walking subscale independently and a total TWSTRS score (from 0 to 85) is calculated.

Key evaluation issues

The TWSTRS has been widely used (Münchau et al, 2001), and has been shown to have internal consistency and acceptable interrater agreement. Evidence for validity is shown by moderate within-scale correlations (Consky and Lang, 1994). Responsiveness to change has been demonstrated (Comella et al, 2011; Lew et al, 2010).

Strengths and Weaknesses

The TWSTRS assesses the severity of cervical dystonia and includes disability interrater agreement. Evidence for validity is shown by moderate within-scale correlations (Consky and Lang, 1994). Responsiveness to change has been demonstrated (Comella et al, 2011; Lew et al, 2010). The TWSTRS scale might be too complex for routine clinical practice. A revised version of the TWSTRS has been recently devised and its clinimetric properties are at the moment under scrutiny (Comella, personal communication).

Craniocervical Dystonia Questionnaire

Scale description

The Craniocervical Dystonia Questionnaire (CDQ-24) 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.

Key evaluation issues

The CDQ-24 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 (Muller et al, 2004).

Strengths and Weaknesses

The CDQ-24 is an easy instrument and also evaluates pain, sleep, and depression due to dystonia.

Voice Handicap Index

Scale description

The Voice Handicap Index (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.

Key evaluation issues

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.

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. The VHI is similar to the Vocal Performance Questionnaire, and direct comparisons have been made showing similar clinimetric properties.

Vocal Performance Questionnaire

Scale description

This scale was designed for use in an evaluation study of voice therapy in cases of nonorganic dysphonia (Carding et al, 1999) The Vocal Performance Questionnaire (VPQ) is a 12-item questionnaire 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.

Key evaluation issues

The VPQ was initially found to have good internal consistency in a study that included a large range of voice pathologies except for spasmodic dysphonia (Deary et al, 2004). In a study that included patients with laryngeal dystonia, (Webb et al, 2007) the VPQ had high levels of internal consistency and test-retest reliability. Although the VPQ still needs further validation in patients with organic disorders, it may be Recommended for use in laryngeal 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 organic dystonia.

Fahn-Marsden Dystonia Rating Scale

Scale description.

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 scores 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.

Key evaluation issues

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 (Burke et al, 1985) The FMDRS showed good internal consistency and good level of interrater reliability for the total scores (Comella et al, 2003) For separate items, interrater agreement was fair to good, being lowest for eyes, jaw, face, and larynx (Comella et al, 2003).

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 assesses 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.

Chorea

There are many causes of chorea, but scales have been mainly developed for the disease that represents the prototype of choreatic disorders, Huntington's disease (HD). The only exception is a scale specifically developed for Sydenham’s chorea (Teixeira et al, 2005), the USCRS, which has not been employed by research groups outside the one that proposed it.

Four scales have been identified that are used to assess HD (Huntington Disease Study Group, 1996; Guy, 1976; Marsden and Schachter, 1981) (Table 5), of which three reached the Recommended status (Table 6). These three scales, the Unified Huntington's Disease Rating Scale (UHDRS), the Abnormal Involuntary Movement Scale (AIMS), and the modified motor score of the UHDRS (mMS), will be discussed here in detail.

Unified Huntington's Disease Rating Scale

Scale description

The UHDRS (Huntington Disease Study Group, 1996) has been the most widely used scale for HD in routine clinical practice and clinical trials. The final version has four components assessing motor performance, cognitive performance, behavioral abnormalities and functional capacity. The motor section of the UHDRS includes 15 items, scored from 0 to 4 and divided into six subdomains: oculomotor function, dysarthria, chorea, dystonia, gait, and postural stability (Huntington Disease Study Group, 1996). A teaching videotape showing how to assess the motor features is also available, helping to standardize the practical application of the scale, and to enhance inter-intra reliability.

Key evaluation issues

Internal consistency, in each of the four component, has shown high degree (Cronbach's alpha values were 0.95 for the total motor score, 0.90 for the cognitive tests, 0.83 for the behavioral scale, and 0.95 for the functional checklist). Correlation analysis has shown that four components of UHDRS were highly intercorrelated, except for the behavioral score; nevertheless higher mood subscale scores correlated with better motor performance, whereas higher psychosis and obsessive subscale scores correlated with lower functional scores. High degree of intrerrater reliability, assessed by intraclass coefficient, has shown (intraclass correlation coefficient was 0.94 for the total motor score, 0.82 for the chorea score and 0.62 for the dystonia score). Despite its psychometric properties including internal consistency, inter-rater reliability and sensitivity to change are satisfactory, its length and load have been criticized (Martinez-Martin al, 2014). It has been shown that UHDRS is useful for tracking clinical changes longitudinally in patients with HD.

Strengths and weaknesses

The UHDRS assesses relevant clinical domains of HD and was designed for repeated administrations during clinical research studies. The UHDRS may be particularly suitable to follow clinical changes in the setting of controlled trials of experimental interventions. The UHDRS yields several scores assessing the primary features of HD (motor, cognitive, behavioral) as well as the overall functional impact of these features.

Abnormal Involuntary Movement Scale

Scale description

The AIMS is a 12-item clinician-rated instrument to assess the severity of abnormal movements. Seven items assess the abnormal movements in seven body areas (Guy, 1976), each area scored from 0 to 4 (none, minimal, mild, moderate, severe). Three items rates the abnormal movements global severity and patient's awareness, and the disability due to abnormal movements. Two more items focus on dental status. The scale includes specific instructions to standardize the evaluation and requires the examiner to observe the patient sitting quietly at rest and again while carrying out selected motor tasks.

The severity rating is ranked one point lower when the abnormal movements occur only upon activation maneuvers (such as opening and closing mouth, finger tapping, standing or walking), then if they occur already (spontaneously) at rest at the same intensity.

Key evaluation issues

AIMS meets the criterion to use in HD and has been used by several authors studying effects of different drugs on dyskinesia associated with HD (Guy, 1976; Vitale et al, 2007; Ondo et al, 2002; Van Vugt et al, 1997; De Tommaso, 2007). Clinimetric data rely mainly on the high inter-rater and test-retest reliability assessed in tardive dyskinesia (Seet et al. 1993; Whall et al, 1983). The clinimetric properties of the scale have not been specifically examined in HD. Moreover, only the original version of the scale has been clinimetrically assessed, whereas none of the modified versions has undergone validation testing. As a final designation, the AIMS meets the criteria of a Recommended scale, but with limitations that includes limited clinimetric data in HD patients, poor documentation of

phenomenological sub-types of dyskinesia, and no information on the impact of dyskinesia on the patient's quality of life.

Strengths and weaknesses

The AIMS is quick to administer and takes only about 10 minutes to complete. It has been developed for tardive dyskinesia, therefore emphasizes face and neck abnormal movements that may not be the (only) primary focus of dyskinesia in HD. It focuses on dyskinesia localization (anatomy) but does not specified whether dyskinesia is dystonic or choreic. Moreover, the disability rating relies only on clinician's judgement and does not account for patient's perception. AIMS has several modified versions and it is not entirely clear whether clinimetric analyses are uniform across all versions.

Modified Motor Score of the Unified Huntington's Disease rating Scale

Scale description

The mMS is a subscale of the UHDRS Motor Score proposed by NeuroSearch (Waters et al. 2010) which excluded eyes movements, chorea and dystonia and focused only on voluntary movements which correlate strongly with the disability related to HD compared to chorea. It has been proposed with the aim to have a quicker method to assess the response to treatment in patients with HD.

Key evaluation issues

The mMS has been applied in two clinical trial to investigate the potential effect of pridopidine in HD, the MermaidHD (Multinational European Study) and HART studies (which involves U.S. and Canada) (De Yebenes et al, 2011; HSG HART Investigators, 2013). This subscale has shown improved internal consistency compared with total motor score (TMS), and interclass correlation similar to that for TMS, indicating good test-retest reliability(Waters et al. 2010).

Strengths and weaknesses

This subscale is brief and easy to perform and allows to monitor functional modifications over time, which are related to the disability due to worsening of the voluntary movements. However, the clinimetric properties have not been formally assessed, and therefore this brief

scale reaches the Recommended status carrying several limitations.

Tics

Fourteen scales have been assessed for tic evaluation (Gadow & Paolicelli, 1986; Walkup et al, 1992; Gaffney et al, 1994; Woods et al, 2005; Shapiro & Shapiro, 1984; Shapiro et al, 1988; Shytle et al, 2003; Leckman et al, 1988; Leckman et al, 1989; Harcherik et al, 1984; Nolan et al, 1994; Storch et al, 2007; Cavanna et al, 2008; Jagger et al, 1982; Kompoliti et al, 1997; Kurlan et al, 1988) (Table 7), and twelve of them, the Global Tics Rating Scale (GTRS), the Hopkins Motor and Vocal Tic Scale (HMVTS), the Motor Tic, Obsessions and Compulsions, Vocal Tic Evaluation Survey(MOVES), the, the Premonitory Urge for Tics Scale (PUTS), the Shapiro Tourette Syndrome Severity Scale (STSSS), the Tourette's Disorder Scale (TODS), the Tourette Syndrome-Clinical Global Impression (TS-CGI), the Tourette Syndrome Global Scale (TSGS), the Unified Tic Rating Scale (UTRS), the Gilles de la Tourette Syndrome Quality of Life (GTS-QoL) scale, the Yale Global Tic Severity Scale (YGTSS) and the Diagnostic Confidence Index (DCI), reached the recommendation status (Cavanna and Piedad, 2013) (Table 8). We describe below only the rating scales used in the evaluation of motor tics, whereas the GTS-QOL (Cavanna et al, 2008) and the DCI (Robertson et al, 1999) focus on health-related quality of life and clinician's confidence in the diagnosis of Tourette syndrome, respectively.

Global Tics Rating Scale

Scale description

The GTRS is a brief clinician-rated measure of tic frequency (Gadow & Paolicelli, 1986). This checklist contains nine items, with the first five referring to the frequency of motor (three items) and phonic tics (two items) according to body region, which are summed to produce motor and phonic tic frequency subscores, respectively. All items are rated on a scale from 0 (never) to 3 (very much).

Key evaluation issues

When compared with other scales (YGTSS and GTRS) showed relatively small correlation coefficients (.01–.42), suggesting poor concurrent validity. Reliability appeared acceptable for motor and severity subscales with correlation coefficients of above .77, but less so with the vocal tic subscale (.58) (Nolan et al, 1994).

Strengths and weaknesses

The GTRS has been used in interventional trials for patients with Tourette syndrome (Pringsheim & Steeves, 2011). It should be noted that, in the study by Nolan et al. (1994) the GTRS was rated by teachers, who may not have the experience of skilled clinicians in adequately recognizing tics.

The Hopkins Motor and Vocal Tic Scale

Scale description

The HMVTS is a parent- and clinician-rated scale that measures motor and phonic tics using a visual analog scale (VAS) (Walkup et al, 1992). Each tic is given a rating from 0 (not present) to 10 (most severe), with four intermediate ratings: "mild," "moderate," "moderately severe," and "severe." The severity of tics is rated across their frequency, intensity, and the level of interference and impairment they cause. An overall score is also assigned to each motor or phonic tic, ranging from 1 (no tic symptoms) to 5 (worst ever).

Key evaluation issues

Initial psychometric testing demonstrated excellent reliability and validity indices when compared to the YGTSS, STSSS, and behavioral measures (Walkup et al, 1992).

Strengths and weaknesses

The HMVTS has been used in intervention studies to assess tic symptoms (Pringsheim & Steeves, 2011; Singer et al, 1995).

Motor Tic, Obsessions and Compulsions, Vocal Tic Evaluation Survey

Scale description

The MOVES is a self-report assessment of motor and phonic tics, associated phenomena, obsessions, and compulsions (Gaffney et al, 1994). These scores can be combined to produce tic or obsessive-compulsive subscores. Individuals are asked to rate the presence of their symptoms from a list of 20 items in the previous 4 weeks, on a 4-point scale from "never" to "always."

Key evaluation issues

There are some issues with regard to its reliability, however, with some subscales showing below standard correlation coefficients (<.70): tic (.54), associated symptoms (.40), and total scores (.69) (Gaffney et al, 1994).

Strengths and weaknesses

The MOVES is a relatively straightforward scale. Its acceptable validity was demonstrated with strong correlations to two clinician-rated tic scales and two OCB measures, with some sensitivity to clinically relevant changes (Gaffney et al, 1994). Nevertheless, the MOVES is a useful adjunct to clinician ratings as a measure of patient perception of tic symptoms, for example during interventional or phenomenological studies (Haddad et al, 2009; Mu□nchau et al, 2002; Orth et al, 2005).

Premonitory Urge for Tics Scale

Scale description

The PUTS is a relatively brief scale designed to examine the phenomenon of premonitory sensations (also called premonitory urges or sensory tics) in tic disorders (Woods et al, 2005). It contains 10 descriptions of somatic sensations derived from phenomenological descriptions from the literature and clinical experience. The severity of urges is rated on a 4-point ordinal scale ranging from 1 (not at all true) to 4 (very much true). Although psychometric properties appear acceptable in older pediatric patients, initial testing revealed inadequate properties for patients younger than 10 years (Woods et al, 2005).

Recommendation status

A direct translation in Hebrew has recently become available in a study that provided independent testing of the PUTS, showing adequate properties in patients older than 10 years (Steinberg et al, 2010). Despite testing in pediatric populations only, adequate psychometric properties for older children may indicate utility in adults also. More recently, a study using the PUTS and a similar scale (USP-SPS, see below) demonstrated concurrent and discriminant validity (Sutherland Owens et al, 2011).

Strengths and weaknesses

Despite testing in pediatric populations only, adequate psychometric properties for older children may indicate utility in adults also.

Shapiro Tourette Syndrome Severity Scale

Scale description

The STSSS was developed to measure changes in tic symptoms during a clinical trial of pimozide (Shapiro & Shapiro, 1984; Shapiro et al, 1988). This clinician-rated scale addresses five factors: whether tics are noticeable, whether they elicit comments or curiosity, whether the patient is considered odd or bizarre, whether tics interfere with functioning, and whether they lead to incapacitation or to the patient being homebound or hospitalized. The item scores can be summed to produce total ratings, which are assigned a global severity rating on a 6-point scale from 0 (no tics) to 6 (very severe, total sum of ratings >8).

Key evaluation issues

When compared with two measures of tic severity, the STSSS showed excellent reliability and validity (Walkup et al, 1992) and has been successfully used in interventional studies for Tourette syndrome (e.g., Mu□ller-Vahl et al, 2002; Pringsheim & Marras, 2009; Pringsheim & Steeves, 2011)

Strengths and weaknesses

The STSSS take around 5 to 10 minutes to complete. However, the focus is clearly on social impairment, which somewhat limits the overall accurateness of this clinical assessment of tics.

Tourette's Disorder Scale

Scale description

The TODS is an objective measure of the severity of tics plus a wide range of

neuropsychiatric symptoms associated with Tourette syndrome over a period of 1 month (Shytle et al, 2003). The 15-items scale contains questions about tics, inattention, hyperactivity, obsessions, compulsions, aggressions, and emotional symptoms.

Key evaluation issues

Clinician and parent-rated versions are available, both of which have been validated by further psychometric testing (Storch et al, 2007; Storch et al, 2004).

Strengths and weaknesses

Clinician and parent-rated versions are available.

Tourette Syndrome-Clinical Global Impression

Scale description

The CGI is a frequently used measure of disease severity. In preparation for a clinical trial in Tourette syndrome, three disease-specific versions of the CGI were developed (Leckman et al, 1988). These assess symptoms of Tourette syndrome, obsessive-compulsive disorder (OCD), and attention-deficit and hyperactivity disorder (ADHD), based on DSM-III diagnostic criteria. The TS-CGI consists of seven items corresponding with descriptions of no identifiable symptoms (normal) to incapacitating tics or a high level of functional impairment associated with behavioral symptoms (extremely severe).

Key evaluation issues

The TS-CGI showed excellent reliability and validity indices when compared to the STSSS and YGTSS (Walkup et al, 1992) and has been used as an adjunctive clinician measure in interventional studies (Kwon et al, 2011; Pringsheim & Marras, 2009; Pringsheim & Steeves, 2011; Sallee et al, 1997).

Strengths and weaknesses

Very useful in the assessment of both motor and non-motor symptoms across tic disorders.

Tourette Syndrome Global Scale

Scale description

The TSGS is a clinician-rated measure of tics and social functioning in Tourette syndrome (Harcherik et al, 1984). The tic subscore measures simple and complex motor and phonic tics based on their frequency and resulting impairment. Frequency scores are rated on a scale of 0 to 5, with higher scores corresponding to higher tic frequencies. Impairment is measured on a scale of 1 to 5 based on how noticeable tics are and the resulting functional impairment. The social functioning subscore measures the level of functional impairment in behavioral conduct, motor restlessness, and school and learning or work and occupation problems (whichever is relevant). These are rated on a scale of 0 to 25 in increments of 5, with 0 indicating lack of issues in respective domains and 25 indicating severe functional impairment. The global measure of severity is calculated using both the tic and social functioning subscores.

Key evaluation issues

However, in a pharmacological trial of ondansetron, the TSGS detected a significant symptomatic change whereas a gold standard for measuring tic symptoms did not find differences compared to placebo ($p = .002$ vs. .15, respectively; Toren et al, 2005), suggesting potential concurrent validity issues.

Strengths and weaknesses

The advantage of the TSGS is its comprehensiveness in measuring different tic characteristics, as well as its multidimensionality by combining assessments of tics and their effects on social functionality. However, the formula for deriving global scores is relatively complicated and potentially produces social functioning subscores that are disproportionately weighted to tic symptoms (Kompoliti & Goetz, 1997). The TSGS has proven to represent an useful instrument in interventional studies (e.g., Mu□ller-Vahl et al, 2002; Pringsheim & Marras et al, 2009; Sallee et al, 1997).

Unified Tic Rating Scale

Scale description

The UTRS contains subscales rated by patients and/or informants and clinicians, which are summed to indicate overall tic severity. A 2-minute tic count is also incorporated to measure motor and vocal tics during conversation with the patient. The subscales contain items on the anatomical distribution of tics, types, frequency, intensity, and level of interference and suppression. Measures of ADHD and OCD symptoms and global functioning are also included.

Key evaluation issues

The 2-minute tic count component was used in a phenomenological study in Tourette syndrome (Nolan et al, 1994), which demonstrated the versatility of the UTRS by focusing on relevant subscales. Revisions and piloting are currently under way, with a view toward improving the reliability, dimensionality, internal consistency, and validity (Kurlan & McDermott, 2005).

Strengths and weaknesses

The multidimensionality of the UTRS and the inclusion of both subjective and objective data are its greatest advantages.

Yale Global Tic Severity Scale

Scale description

The YGTSS is the most widely used measure of tic severity in TS (Porta et al, 2009; Pringsheim et al, 2009; Pringsheim & Marras, 2009; Pringsheim & Steeves, 2011) and other tic disorders (Leckman et al, 1989). The YGTSS is based on a semistructured interview focusing on tic symptoms over the past week, where the clinician is asked to record patients' motor and phonic tics. Subsequently, the tic symptoms are rated separately based on their number, frequency, intensity, complexity, and interference from 0 (no tic symptoms) to 5 (severe). The tic severity subscore consists of the motor and phonic tic severity scores. This

is summed with the impairment subscore, which rates the severity of functional impairment from 0 to 50, to produce the total score (0–100).

Key evaluation issues

The YGTSS allows a multidimensional overview of tic characteristics, as well as the level of functional interference. The separation of motor and phonic tics is particularly useful, also for diagnostic purposes. Independent testing confirmed reliability and validity indices using the STSSS and TS-CGI (Walkup et al, 1992). A childhood and adolescent sample confirmed this scale's excellent psychometric properties, particularly its internal consistency and validity (Storch et al, 2005). A factor analytic study also confirmed and validated the structure that was initially identified (Storch et al, 2007). Cutoff scores for clinically relevant treatment response have been proposed: a reduction by 35% in total YGTSS scores or 6 or 7 points in the tic severity subscale (Storch et al, 2011).

Strengths and weaknesses

The YGTSS is an ideal instrument for busy routine clinical practice, as it takes only 15 minutes to complete.

Myoclonus

Three scales have been assessed for myoclonus (Chadwick et al, 1977; Troung & Fahn, 1988; Fructh et al, 2002; Table 9), but only one, the Unified Myoclonus Rating Scale UMRS) reached the recommendation status (Table 10).

Unified Myoclonus Rating Scale

Scale description

This scale is a further development of the Truong and Fahn myoclonus scale (Truong & Fahn, 1988). The UMRS was modified to address some of the shortfalls from the original scale. The paper describing UMRS is published in a non peer reviewed book. (Frucht et al, 2002). The UMRS is a quantitative 73-item clinical rating instrument developed to evaluate myoclonus and response, of patients with myoclonus, to anti-myoclonic therapy. The scale

contains a patient questionnaire, a handwriting and spiral sheet, rating instructions, a score sheets and a videotape. The scale consists of eight sections: section 1, patient questionnaire (11 items); section 2, myoclonus at rest (frequency and amplitude, 16 items); section 3, stimulus sensitivity of myoclonus (17 items); section 4, severity of myoclonus with action (frequency and amplitude, 20 items); section 5, performance on functional tests (5 items); section 6, physicians rating of patients global disability (1 item); section 7, presence of negative myoclonus (1 item); section 8, severity of negative myoclonus (1 item). Each item is rated on a scale of 0 to 4, with the exception of sections 3 and 7 (rated present or absent) and section 8 rated on a scale of 0-3.

Key evaluation issues

Members of the Myoclonus Study group performed statistical validation of the UMRS. Twenty patients with chronic myoclonus, having different etiologies and severity, were videotaped while UMRS was performed. Eighteen neurologists, experts in movement disorders, rated two tapes (ten patients per neurologist). A Cronbach's alpha (measure of reliability) was calculated for each section of the UMRS and confirmed that the scale possesses excellent interrater reliability.

Strengths and weaknesses

The scale is easy to use and can be completed in less than 15 minutes. Intrarater reliability has not been tested formally (Frucht et al, 2002).

Drug-induced dyskinesia

Seven scales have been assessed for tardive dyskinesia (DID) evaluation (Simpson et al, 1979; Sprague et al, 1984; Sprague et al, 1991; Smith et al, 1983; Lund et al, 1991; Guy, 1976; Chouinard and Margolese, 2005; Lindström et al, 2001) (Table 11). Only three of them received the recommendation status, the Simpson Tardive Dyskinesia Rating Scale (TDRS), the Abbreviated Dyskinesia Scale (ADS), the Dyskinesia Identification System – Coldwater (DIS-Co)

Dyskinesia Identification System Condensed User Scale (DISCUS) and the Extrapyramidal Symptoms Rating Scale (ESRS) Extrapyramidal Symptoms Rating Scale abbreviated

version (ESRS-A) (table 12). The review of the scales used to measure L-Dopa induced dyskinesia in PD has been the focus of a specific article as part of the global effort to review all scales used in PD (Colosimo et al, 2010), and will not be further discussed here.

TDRS: Simpson Tardive Dyskinesia Rating Scale
ADS: Abbreviated Dyskinesia Scale

Scale description

The TDRS is a clinician-rated instrument to assess the severity of abnormal movements in four anatomical regions, each area scored from 0 to 6 (absent, possibly present, mild, moderate, moderately severe, very severe). Fourteen items rate the severity of abnormal movements in the facial region (including tongue and lips), six rate the neck/trunk region, six rate the upper limbs, six rate the lower limbs; two additional items rate the severity of abnormal movements involving the entire body. The scale includes specific instructions to standardize the evaluation. The highest severity of the abnormal movements is rated. The scale does not provide word-anchors to explain the designations of absent, minimal, mild, moderate and severe, so that these final designations may be biased by the rater's experience. The scale was developed exclusively for rating tardive dyskinesia in psychiatric patients. Although it primarily focuses on anatomy, it characterizes movements using descriptive definitions of the various items, often referring also to phenomenological terms as 'choreoathetoid', 'ballistic', 'torticollis', and 'tics', suggesting it aims at encompassing different forms of hyperkinesias. There is no patient input into the ratings, relying on the clinician's judgment rather than patient perceptions. Because it is a scale developed for tardive dyskinesia, it emphasizes face and neck movements that may not be the primary focus of dyskinesia in PD. The same group presented an abbreviated dyskinesia rating scale (ADS) consisting of only 13 items, characterised by a less focal reference to anatomical sites. Both scales offered the opportunity for raters to add additional individualized items (Simpson et al, 1979).

Key evaluation issues

Clinimetric data on TDRS and ADS rely mainly on inter-rater coefficients (0.98 and 0.97, respectively), whereas test-retest reliability has not been evaluated in detail. Convergent validity and responsiveness to change have also been evaluated. Given that this scale has

been applied to TD populations, that data on its use are available beyond the group that developed the scale, and that it showed good inter-rater reliability, convergent validity and responsiveness to change, both TDRS and ADS are considered Recommended for use in TD.

Strengths and weaknesses

The TDRS is quick to administer and takes about 15 minutes to complete. Intra-rater (test-retest reliability) was not evaluated by authors.

DIS-Co: Dyskinesia Identification System – Coldwater
DISCUS: Dyskinesia Identification System Condensed User Scale

Scale description

The DIS-Co is a 34-item scale originally based on data from 519 "institutionalized mentally-retarded residents", 250 of whom had never received antipsychotic medications. The 34 items are grouped in 10 body areas and scored on a five-point scale. Later, the scale was revised by the same authors, based on data collection from subsequent studies. They developed a method for selecting items for the rating scale based on six qualities, including interrater reliability, stability, and relationship with medication (Sprague et al, 1984, Sprague et al, 1991).

Key evaluation issues

The resulting 15-item scale, the DISCUS, has been used to assess tardive dyskinesia. Both have been thoroughly evaluated on psychometric properties: the DIS-Co showed 0.78interrater reliability, 0.77 test-retest reliability (2-week stability), good construct and convergent validity, and good responsiveness to change; the DISCUS showed 0.92 interrater reliability and 0.40 2-week stability, based on 400 individuals with developmental disability (Sprague et al, 1991).

Strengths and weaknesses

This scale has been applied to TD populations, particularly those with developmental

disabilities, and it showed good inter-rater reliability, convergent validity and responsiveness to change. The limitation is the exclusive use in patients with learning disabilities.

ESRS: Extrapyramidal Symptoms Rating Scale
ESRS-A: Extrapyramidal Symptoms Rating Scale abbreviated version

Scale description

The ESRS is a clinician-rated instrument to assess the severity of parkinsonism, dyskinesia, dystonia and akathisia, each manifestation scored from 0 to 5 (absent, minimal, mild, moderate, severe, extreme) with welldefined explanatory word-anchors. Its most commonly used form (ESRS-A) comprises 28 items for the comprehensive evaluation of movement disorders. The scale includes specific instructions to standardize the evaluation. Ratings are made through a combination of clinical interview and motor examination, after which a score can be obtained. The severity of ratings should be related to both frequency and severity/intensity of the phenomenon being evaluated. The highest severity of the abnormal movements is rated. Each rating should represent the most appropriate rating based on overall clinical judgment for the anchor points pertaining to each symptom, with a temporal component considered secondarily (Chouinard and Margolese, 2005).

Key evaluation issues

Clinimetric data on ESRS rely mainly on inter-rater coefficients (0.86-0.91). Convergent validity and responsiveness to change have also been evaluated. During a cross-scale comparison, AIMS and ESRS were found to have a 96% agreement between TD-defined cases by DSM-IV TD criteria. This scale has been applied to TD populations, that data on its use are available beyond the group that developed the scale, and that it showed good inter-rater reliability, convergent validity and responsiveness to change.

Strengths and weaknesses

The ESRS is quick to administer and takes about 15 minutes to complete.

Conclusions

The adoption of valid rating scales for the assessment of patients with hyperkinetic disorders is needed to be able to share information, quantify health status and fully understand the trends and outcomes in clinical trials (Martinez-Martin et al, 2014). Although numerous scales aimed at assessing hyperkinetic disorders have been published, their variability in terms of clinimetric properties, availability and effort required to administer them is high. In this review, we identified scales defined as 'Recommended' for the assessment of all forms of hyperkinetic disorders. The situation highlighted by our systematic review nevertheless varies considerably, with several 'Recommended' scales being available for some conditions such as tics or dystonia, but only one being available for myoclonus. This gap needs to be filled by the scientific community through both the development of new clinical tools and the refinement of existing ones. Further work is also required to adapt and validate 'Recommended' scales to specific patient populations according to different age groups and clinical stratifications.

Acknowledgements

The authors wish to thank Dr Davide Martino (London, UK) for his help in identifying scales used to assess tics and Drug-induced dyskinesia.

References

Abdo WF, van de Warrenburg BP, Burn DJ, Quinn NP, Bloem BR. The clinical approach to movement disorders. Nat Rev Neurol 2010; 6:29-37.

Albanese A, Del Sorbo F, Comella C, Jinnah HA, Mink JW, Post B, Vidailhet M, Volkmann J, Warner TT, Leentjens AF, Martinez-Martin P, Stebbins GT, Goetz CG, Schrag A. Dystonia rating scales: critique and recommendations. Mov Disord 2013; 28(7):874-83.

Bain PG, Findley LG, Atchison P, Behari M, Vidailhet M, Gresty M, Rothwell JC, Thompson PD, Marsden CD. Assessing tremor severity. J Neurol Neurosurg Psychiatry 1993; 56:868–73.

Baruzzi A, Procaccianti G, Martinelli P, Riva R, Denoth F, Montanaro N, Lugaresi E. Phenobarbital and propranolol in essential tremor: a double-blind controlled clinical trial. Neurology 1983; 33:296-300.

Bové M, Daamen N, Rosen C, Wang CC, Sulica L, Gartner-Schmidt J. Development and Validation of the Vocal Tremor Scoring System. The Laryngoscope 2006; 116:1662-7.

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.

Cano SJ, Warner TT, Linacre JM, Bathia KP, Thompson AJ, Fitzpatrick R, Hobart JC. Capturing the true burden of dystonia on patients: the Cervical Dystonia Impact Profile (CDIP-58). Neurology 2004; 63:1629–33. .

Cano SJ, Hobart JC, Edwards M, Fitzpatrick R, Bathia KP, Thompson AJ, Warner TT. CDIP-58 can measure the impact of botulinum toxin treatment in cervical dystonia. Neurology 2006; 67:2230–2. .

Cano SJ, Warner TT, Thompson AJ, Bathia 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 and Quality of Life Outcomes 2008; 6:58.

Carding PN, Wilson JA, MacKenzie K, Deary IJ. Measuring voice outcomes: state of the science review. The Journal of Laryngology & Otology 1999; 123(8):823-9.

Cavanna AE, Schrag A, Morley D, Orth M, Robertson MM, Joyce E, Critchley HD, Selai C. The Gilles de la Tourette Syndrome-Quality of Life scale (GTS-QOL): development and validation. Neurology 2008; 71: 1410-6.

Cavanna AE, Piedad JCP. Clinical Rating Instruments in Tourette Syndrome. Oxford Med 2013.

Chadwcik D, Hallett M, Harris R, Jenner P, Reynolds EH, Marsden CD. Clinical, biochemical, and physiological features distinguishing myoclonus responsive to 5-

hydroxytryptophan, tryptophan with a monoamine oxidase inhibitor, and clonazepam. Brain 1977; 100:455-87.

Chouinard G, Margolese HC. Manual for the Extrapyramidal Symptom Rating Scale (ESRS).Schizophr Res 2005;76:247-65.

Cohen DJ, Leckman JF, Shaywitz BA. The Tourette's syndrome and other tics. In: Shaffer D, Ehrhardt AA & Greenhill L (Eds.). The clinical guide to child psychiatry. New York: Free Press, 1984; pp. 566-73.

Colosimo C, Martínez-Martín P, Fabbrini G, Hauser RA, Merello M, Miyasaki J, Poewe W, Sampaio C, Rascol O, Stebbins GT, Schrag A, Goetz CG. Task force report on scales to assess dyskinesia in Parkinson's disease: critique and recommendations. Mov Disord 2010; 25:1131-42.

Comella CL, Leurgans S, Wuu J, Stebbins GT, Chmura T; Dystonia Study Group. Rating scales for dystonia: a multicenter assessment. Mov Disord 2003; 18:303–12.

Comella CL, Jankovic J, Truong DD, Hanschmann A, Grafe S; U.S. Xeomin Cervical Dystonia Study Group. 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.

Consky ES, Lang AE. Clinical assessments of patients with cervical dystonia. In Jankovic J, Hallett M, eds. Therapy with Botulinum Toxin. New York, NY: Marcel Dekker, 1994; pp. 211–37.

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.

De Yebenes JG, Landewehrmeyer B, Squitieri F, Reilmann R, Rosser A, Barker RA, Saft C, Magnet MK, Sword A, Rembratt A, Tedroff J; MermaiHD study investigators. Pridopidine for the treatment of motor function in patients with Huntington's disease (MermaiHD): a phase 3, randomised, double-blind, placebo-controlled trial. Lancet Neurology 2011; 10:1049-57.

de Tommaso M, Difruscolo O, Sciruicchio V, Specchio N, Livrea P. Two Years' Follow-up of Rivastigmine Treatment in Huntington Disease. Clinical

Neuropharmacology 2007; 30(1):43-6.

Elble R, Comella C, Fahn S, Hallett M, Jankovic J, Juncos JL, Lewitt P, Lyons K, Ondo W, Pahwa R, Sethi K, Stover N, Tarsy D, Testa C, Tintner R, Watts R, Zesiewicz T. Reliability of a new scale for essential tremor. Mov Disord. 2012 Oct;27(12):1567-9. doi: 10.1002/mds.25162

Fahn S, Tolosa E, Marin C, Parkinson's Disease and Movement Disorders, 2nd edition. Clinical rating scale for tremor. In: Jankovic J, Tolosa E, New York, Williams & Wilkins, 1993; pp. 271-80.

Fernandez HH, Pagan F, Danisi F, Greeley D, Jankovic J, Verma A, Sethi K, Peppert EJ onbehalf of the XCiDaBLE Study Group. Prospective Study Evaluating IncobotulinumtoxinA for Cervical Dystonia or Blepharospasm: Interim Results from the First 145 Subjects with Cervical Dystonia. Tremor Other Hyperkinetic Movements (NY) 2013; 17:3.

Findley LJ, Cleeves L. Phenobarbitone in essential tremor. Neurology 1985; 35:1784-7.

Fried LP, Bandeen-Roche K, Williamson JD, Prasada-Rao P, Chee E, Tepper S, Rubin GS. Functional decline in older adults: expanding methods of ascertainment. J Gerontol A Biol Sci Med Sci 1996; 51:M206-14.

Frucht SJ, Leurgans SE, Hallett M, Fahn S. The Unified Myoclonus Rating Scale. Adv Neurol 2002; 89:361-76.

Gadow KD, Paolicelli LM. Global Tic Rating Scale. Stony Brook, NY: State University of New York Department of Psychiatry, 1986. Unpublished results.

Gaffney GR, Sieg K, Hellings J. The MOVES: a self-rating scale for Tourette's syndrome. J Child Adolesc Psychopharmacol 1994; 4:269-80.

Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, Poewe W, Sampaio C, Stern MB, Dodel R, Dubois B, Holloway R, Jankovic J,

Kulisevsky J, Lang AE, Lees A Leurgans S, LeWitt PA, Nyenhuis D, Olanow CW, Rascol O, Schrag A, Teresi JA, van Hilten JJ, LaPelle N; Movement Disorder Society UPDRS Revision Task Force. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord 2008; 23:2129-70.

Guy W. ECDEU assessment manual for psychopharmacology. Washington, DC: US Government Printing Office. Abnormal Involuntary Movement Scale, 1976. pp. 534-7

Haddad AD, Umoh G, Bhatia V, Robertson MM. Adults with Tourette's syndrome with and without attention deficit hyperactivity disorder. Acta Psychiatrica Scand 2009; 120:299-307.

Handforth A, Bordelon Y, Frucht SJ, Quesada A. A Pilot Efficacy and Tolerability Trial of Memantine for Essential Tremor. Clin Neuropharmacol 2010; 33(5):223-6.

Harcherik DF, Leckman JF, Detlor J, Cohen DJ. A new instrument for clinical studies of Tourette's syndrome. J Am Acad Child Psychiatry 1984; 23:153-60.

Hooper J, Taylor R, Pentland B, Whittle IR. Rater Reliability of Fahn's Tremor Rating Scale in Patients With Multiple Sclerosis. Arch Phys Med Rehabil 1998; 79:1076-9.

Unified Huntington's Disease Rating Scale: Reliability and Consistency. Huntington Study Group. Mov Disord 1996; 11:136-42.

Huntington Study Group HART Investigators. A Randomized, Double-Blind, Placebo-Controlled Trial of Pridopidine in Huntington's Disease. Mov Disord 2013; 28:1407-15.

Jabusch HC, Vauth H, Altenmüller E. Quantification of focal dystonia in pianists using scale analysis. Mov Disord 2004; 19:171-80.

Jacobson BH, Johnson A, Grywalski C, Silbergleit A, Jacobson G, Benninger MS,

Newman CW. The Voice Handicap Index (VHI): development and validation. Am J Speech Lang Pathol 1997;6:66–70.

Jagger J, Prusoff BA, Cohen DJ, Kidd KK, Carbonari CM, John K. The epidemiology of Tourette's syndrome: a pilot study. Schizophr Bull 1982; 8: 267-78.

Jankovic J, Schwartz K, Clemence W, Aswad A, Mordaunt J. A randomized, double-blind, placebo-controlled study to evaluate botulinum toxin type A in essential hand tremor. Mov Disord 1996; 11:250-6.

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.

Jefferson D, Jenner P, Marsden CD. Relationship between plasma propranolol concentration and relief of essential tremor. Journal Neurol Neurosurg Psychiatry 1979; 42:831-7.

Koller WC, Vetere-Overfield B. Acute and chronic effects of propranolol and primidone in essential tremor. Neurology 1989; 39:1587-8.

Kompoliti K, Goetz CG. Tourette syndrome. Clinical rating and quantitative assessment of tics. Neurol Clin 1997; 15:239-54.

Kwon HJ, Lim WS, Lim MH, Lee SJ, Hyun JK, Chae JH, Paik KC. 1-Hz low frequency repetitive transcranial magnetic stimulation in children with Tourette's syndrome. Neurosci Lett 2011; 492:1-4.

Kurlan R, McDermott M, Como P. Tourette Syndrome Unified Rating Scale. New York: Tourette Syndrome Association, 1988.

Kurlan R, McDermott MP. Rating tic severity. In: Kurlan R (ed), Handbook of Tourette's syndrome and related tic and behavioral disorders, 2nd Ed. New York: Marcel Dekker, 2005.

Leckman JF, Towbin KE, Ort SI, Cohen DJ. Clinical assessment of tic disorder

severity. In: Cohen DJ, Bruun RD, Leckman JF (eds.), Tourette's syndrome and tic disorders: clinical understanding and treatment. New York: John Wiley & Sons, 1988.

Leckman JF, Riddle MA, Hardin MT, Ort SI, Swartz KL, Stevenson J, Cohen DJ. The Yale Global Tic Severity Scale: initial testing of a clinician-rated scale of tic severity. J Am Acad Child Adolesc Psychiatry 1989; 28:566-73.

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.

Lindeboom R, De Haan R, Aramideh M, Speelman JD. The blepharospasm disability scale: an instrument for the assessment of functional health in blepharospasm. Mov Disord 1995; 10:444-9.

Lindström E, Lewander T, Malm U, Malt UF, Lublin H, Ahlfors UG. Patient-rated versus clinician-rated side effects of drug treatment in schizophrenia. Clinical validation of a self-rating version of the UKU Side Effect Rating Scale. Nord J Psychiatry 2001;55 (S44):5-69

Louis ED, Ottman R, Ford B, Pullman S, Martinez M, Fahn S, Hauser WA. The Washington Heights-Inwood Genetic Study of Essential Tremor: methodological issues in essential-tremor research. Neuroepidemiology 1997; 16:124-33.

Louis ED, Ford B, Lee H, Andrews H. Does a screening questionnaire for essential tremor agree with the physicians' examination? Neurology 1998; 50:1351-7.

Louis ED, Ford B, Bismuth B. Reliability between two observers using a protocol for diagnosing essential tremor. Mov Disord 1998; 13:287–93.

Louis ED, Wendt KJ, Albert SM, Pullman SL, Yu Q, Andrews H. Validity of a Performance-Based Test of Function in Essential Tremor. Arch Neurol 1999; 56:841-6.

Louis ED, Barnes L, Wendt KJ, Ford B, Sangiorgio M, Tabbal S, Lewis L, Kaufmann P, Moskowitz C, Comella CL, Goetz CC, Lang AE. A Teaching Videotape for the Assessment of Essential Tremor. Mov Disord 2001; 16:89–93.

Lund CE, Mortimer AM, Rogers D, McKenna PJ. Motor, volitional and behavioural disorders in schizophrenia. 1: Assessment using the Modified Rogers Scale. Br J Psychiatry. 1991 Mar;158:323-7, 333-6.Mansur PH, Cury LK, Andrade AO, Pereira AA, Miotto GA, Soares AB, Naves EL. A review on techniques for tremor recording and quantification. Crit Rev Biomed Eng 2007; 35:343-62.

Marsden CD & Schachter M. Br. Assessment of extrapyramidal disorders.Br J Clin Pharmac 1981; (11):129-51.

Martínez-Martín P, Rodrigues-Blazquez, Kurtis MM. Scales and tests in movement disorders. In: Wolters, EC; Baumann, CR; Parkinson disease and other movement disorders: motor behavioural disorders and behavioural motor disorders. VU University Press Amsterdam, The Netherlands: Chicago: International Association of Parkinsonism and Related Disorders; Distributed by Independent Publishers Group, 2014.

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.

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.

Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease. The Unified Parkinson’s Disease Rating Scale (UPDRS): status and recommendations. Mov Disord. 2003; 18:738–50.

Müller J, Wissel J, Kemmler G, Voller B, Bodner T, Schneider A, Wenning GK, Poewe W. Craniocervical dystonia questionnaire (CDQ-24): development and

validation of a diseasespecific quality of life instrument. J Neurol Neurosurg Psychiatry 2004;75:749–753.

Müller-Vahl KR, Schneider U, Koblenz A, Jöbges M, Kolbe H, Daldrup T, Emrich HM.Treatment of Tourette's syndrome with Delta 9-tetrahydrocannabinol (THC): a randomized crossover trial. Pharmacopsychiatry 2002; 35:57-61.

Münchau A, Palmer JD, Dressler D, O'Sullivan JD, Tsang KL, Jahanshashi M, Quinn NP, Lees AJ, Bathia KP. Prospective study of selective peripheral denervation for botulinum-toxin resistant patients with cervical dystonia. Brain 2001; 124:769-83.

Münchau A, Bloem BR, Thilo KV, Trimble MR, Rothwell JC, Robertson MM. Repetitive transcranial magnetic stimulation for Tourette syndrome. Neurology 2002; 59:1789-91.

Nolan EE, Gadow KD, Sverd J. Observations and ratings of tics in school settings. J Abnorm Child Psychol 1994; 22: 579-93.

Ogawa N, Takayama H, Yamamoto M. Comparative studies on the effects of beta-adrenergic blockers in essential tremor. J Neurol 1987; 235:31-3.

Ondo WG, Tintner R, Thomas M, Jankovic J. Tetrabenazine treatment for Huntington's disease-associated chorea. Clin Neuropharmacol 2002; 25:300-2.

Orth M, Kirby R, Richardson MP, Snijders AH, Rothwell JC, Trimble MR, Robertson MM, Münchau A. Subthreshold rTMS over pre-motor cortex has no effect on tics in patients with Gilles de la Tourette syndrome. Clin Neurophysiol 2005; 116:764-8.

Page D, Butler A, Jahanshahi M. Quality of life in focal, segmental, and generalized dystonia. Mov Disord 2007; 22:341-7.

Porta M, Brambilla A, Cavanna AE, Servello D, Sassi M, Rickards H, Robertson MM.Thalamic deep brain stimulation for treatment-refractory Tourette syndrome: two-year outcome. Neurology 2009; 73:1375-80.

Pringsheim T, Lang A, Kurlan R, Pearce M, Sandor P. Understanding disability in Tourette syndrome. Dev Med Child Neurol 2009; 51:468-72.

Pringsheim T, Marras C. Pimozide for tics in Tourette's syndrome. Cochrane Database Syst Rev 2009; 2:CD006996.

Pringsheim T, Steeves T. Pharmacological treatment for attention deficit hyperactivity disorder (ADHD) in children with comorbid tic disorders. Cochrane Database Syst Rev 2011; 4:CD007990.

Priori A, Pesenti A, Cappellari A, Scarlato G, Barbieri S. Limb immobilization for the treatment of focal occupational dystonia. Neurology 2001; 57:405-9.

Robertson MM, Banerjee S, Kurlan R, Cohen DJ, Leckman JF, McMahon W, Pauls DL, Sandor P, van de Wetering BJ.The Tourette syndrome diagnostic confidence index: development and clinical association. Neurology 1999; 53:2108-12.

Sallee FR, Nesbitt L, Jackson C, Sine L, Sethuraman G. Relative efficacy of haloperidol and pimozide in children and adolescents with Tourette's disorder. Am J Psychiatry1997; 154:1057-62.

Seet RA, DeSensi EG, Zubenko GS. Reliability and applicability of movement disorders rating scales in the elderly. J Neuropsychiatry Clin Neurosci 1993; 5:56-60.

Shapiro AK, Shapiro E. Controlled study of pimozide vs. placebo in Tourette's syndrome. J Am Acad Child Psychiatry 1984; 23: 161-73.

Shapiro AK, Shapiro ES, Young JG, Feinberg TE.. Measurement in tic disorders. In: Shapiro AK, Shapiro ES, Young JG, Feinberg TE (eds.), Gilles de la Tourette syndrome (2nd ed.). New York, Raven Press, 1988, pp. 451-480.

Shytle RD, Silver AA, Sheehan KH, Wilkinson BJ, Newman M, Sanberg PR, Sheehan D. The Tourette's Disorder Scale (TODS): development, reliability, and validity. Assessment 2003; 10:273-87.

Simpson GM, Lee JH, Zoubok B, Gardos G. A rating scale for tardive dyskinesia. Psychopharmacology 1979; 64:171–9.

Singer HS, Brown J, Quaskey S, Rosenberg LA, Mellits ED, Denckla MB.The treatment of attention-deficit hyperactivity disorder in Tourette's syndrome: a double-blind placebo-controlled study with clonidine and desipramine. Pediatrics1995; 95:74-81.

Smith RC, Allen R, Gordon J, Wolff J. A rating scale for tardive dyskinesia and parkinsonian symptoms. Psychopharmacol Bull 1983; 19:266–76.

Sprague RL, Kalachnik JE, Breuning SE, Davis VJ, Ullman RK, Cullari S, Davidson NA, Ferguson DG, Hoffner BA. The dyskinesia identification system--Coldwater (DIS-Co): a tardive dyskinesia rating scale for the developmentally disabled. Psychopharmacol Bull 1984; 20:328–38.

Sprague RL, Kalachnik JE. Reliability, validity, and a total score cutoff for the dyskinesia identification system: condensed user scale (DISCUS) with mentally ill and mentally retarded populations. Psychopharmacol Bull 1991; 27:51–58

Stacy MA, Elble RJ, Ondo WG, Wu SC, Humihan J; TRS study group. Assessment of Interrater and Intrarater Reliability of the Fahn–Tolosa–Marin Tremor Rating Scale in Essential Tremor. Mov Disord 2007; 22:833-8.

Steinberg T, Shmuel Baruch S, Harush A, Dar R, Woods D, Piacentini J, Apter A.. Tic disorders and the premonitory urge. J Neural Transm 2010; 117:277-84.

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.

Storch EA, Murphy TK, Geffken GR, Soto O, Sajid M, Allen P, Roberti JW, Killiany EM, Goodman WK. Further psychometric properties of the Tourette's Disorder Scale-Parent Rated version (TODS-PR). Child Psychiatry Hum Dev 2004; 35:107-20.

Storch EA, Murphy TK, Geffken GR, Sajid M, Allen P, Roberti JW, Goodman WK. Reliability and validity of the Yale Global Tic Severity Scale. Psychol Assess 2005; 17:486-91.

Storch EA, Murphy TK, Fernandez M, Krishnan M, Geffken GR, Kellegren AR, Goodman WK. Factor-analytic study of the Yale Global Tic Severity Scale. Psychiatry Res 2007; 149:231-7.

Storch EA, Merlo LJ, Lehmkuhl H, Grabill KM, Geffken GR, Goodman WK, Murphy TK. Further psychometric examination of the Tourette's Disorder Scales. Child Psychiatry Hum Dev 2007; 38:89-98.

Storch EA, De Nadai AS, Lewin AB, McGuire JF, Jones AM, Mutch PJ, Shytle RD, Murphy TK. Defining treatment response in pediatric tic disorders: a signal detection analysis of the yale global tic severity scale. J Child Adolesc Psychopharmacol 2011; 21:621-7.

Sur H, Ilhan S, Erdoğan H, Oztürk E, Taşdemir M, Börü UT. Prevalence of essential tremor: A door-to-door survey in Sile, Istanbul, Turkey. Parkinsonism Relat Disord 2009; 15:101-4.

Sutherland Owens AN, Miguel EC, Swerdlow NR. Sensory gating scales and premonitory urges in Tourette syndrome. Scientific World J 2011; 11:736-41.

Sweet RD, Blumberg J, Lee JE, McDowell FH. Propranolol treatment of essential tremor. Neurology 1974; 24:64-7.

Teixeira AL Jr, Maia DP, Cardoso F. UFMG Sydenham's chorea rating scale (USCRS): reliability and consistency. Mov Disord 2005; 20:585-91.

Toren P, Weizman A, Ratner S, Cohen D, Laor N. Ondasentron treatment in Tourette's disorder: a 3-week, randomized, double-blind, placebo-controlled study. J Clin Psychiatry 2005; 66:499-503.

Truong DD, Fahn S. Therapeutic trial with glycine in myoclonus. Mov Disord 1988; 3:222-32.

Tsui JK, Eisen A, Mak E, Carruthers J, Scott A, Calne DB. A pilot study on the use of botulinum toxin in spasmodic torticollis. Can J Neurol Sci 1985; 12:314–6.

Tsui JK, Eisen A, Stoessl AJ, Calne S, Calne DB. Double-blind study of botulinum toxin in spasmodic torticollis. Lancet 1986; 2:245–7.

Tsui JK, Fross RD, Calne S, Calne DB. Local treatment of spasmodic torticollis with

botulinum toxin. Can J Neurol Sci 1987; 14:533–55.

Van Vugt JP, Siesling S, Vergeer M, van der Velde EA, Roos Ra. Clozapine versus placebo in Huntington's disease: a double blind randomised comparative study. J Neurol Neurosurg Psychiatry 1997; 63:35-9.

Vitale C, Marconi S, Di Maio L, De Michele G, Longo K, Bonavita V, Barone P. Short-term continuous infusion of Apomorphine Hydrochloride for treatment of Huntington's chorea: A double blind, randomized cross-over trial. Mov Disord 2007; 22:2359-64.

Walkup JT, Rosenberg LA, Brown J, Singer HS. The validity of instruments measuring tic severity in Tourette's syndrome. J Am Acad Child Adolesc Psychiatry 1992; 31:472-7.

Waters S., Tedroff J. and Kieburtz K. Validation of the Modified Motor Score (mMS): A Subscale of the Unified Huntington's Disease Rating Scale (UHDRS) Motor Score. Neurotherapeutics 2010; 7(1):144.

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.

Whall AL, Engle V, Edwards A, Bobel L, Haberland C. Development of a screening program for tardive dyskinesias: feasibility issue. Nurs Res 1983; 32:151-6.

Woods DW, Piacentini J, Himle MB, Chang S. Premonitory Urge for Tics Scale (PUTS): initial psychometric results and examination of the premonitory urge phenomenon in youths with Tic disorders. J Dev Behav Pediatr 2005; 26:397-403.

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. 6

Zeuner KE, Peller M, Knutzen A, Holler I, Münchau A, Hallett M, Deuschl G, Siebner HR. How to assess motor impairment in writer's cramp. Mov Disord 2007;

22:1102-9

Tables

Table 1. Tremor: main characteristics of the scales assessed.

Scale	*Time to complete*	*Patient historical rating*	*Clinical examination*	*Administration burden*
FTM-TRS	15*	yes	yes	±
WHIGET	20*	yes	yes	+
Bain TRS	10*	no	yes	±
UTRA	10*	yes	yes	+
TETRAS	10*	yes	yes	+
VTSS	1*	no	yes	+
Sweet et al. TRS	5*	no	yes	+
Findley et al. TRS	5*	yes	yes	+
Koller et al. TRS	5*	no	yes	+
Jefferson et al. TRS	5*	yes	yes	±
Baruzzi et al. TRS	2*	no	yes	+
Ogawa et al. TRS	2*	yes	yes	+

Abbreviations: FTM-TRS, Fahn Tolosa Marin Tremor Rating Scale; WHIGET, Washington Heights-Inwood Genetic Study of Essential Tremor ; UTRA, Unified Tremor Rating Assessment; TETRAS, The Essential Tremor Rating Assessment Scale; VTSS, Vocal Tremor Scale Assessment.

**mean estimated time to complete (not specified); administration burden was rated as follows: ''+''(easy, e.g., summing up of the items), ''±'' (moderate, e.g., visual analogue scale (VAS) or simple formula), ''-'' (difficult, e.g., VAS in combination with formula, or complex formula).*

Table 2. Tremor: classification of the scales assessed.

Scale	*Applied in ET*	*Applied beyond authors*	*Successfull clinimetric testing*	*Qualification*
FTM-TRS	yes	yes	yes	Recommended
WHIGET	yes	yes	yes	Recommended
Bain TRS	yes	no	yes	Suggested
UTRA	yes	yes	no	Suggested
TETRAS	no	no	yes	Listed
VTSS	no	no	yes	Listed
Sweet et al. TRS	yes	no	no	Listed
Findley et al. TRS	yes	no	no	Listed
Koller et al. TRS	yes	no	no	Listed
Jefferson et al. TRS	yes	no	no	Listed
Baruzzi et al. TRS	yes	no	no	Listed
Ogawa et al. TRS	yes	no	no	Listed

Table 3. Dystonia: main characteristics of the scales assessed.

Scale	*Time to complete*	*Patient historical rating*	*Clinical examination*	*Administration burden*
BSDI	n.a.	yes	no	+
Jankovic rating scale	n.a.	no	yes	+
Blepharospasm Disability Scale	4*	yes	yes	+
CDIP-58	n.a	yes	no	+
Functional Disability Questionnaire	n.a.	yes	no	+
TWSTRS	n.a	no	yes	+
Tsui scale	n.a.	no	yes	±
Modified Tsui scale		no	yes	+
Freiberg Questionnaire for Dystonia torticollis version	n.a.	yes	no	±
Disability questionnaire for patients with cervical dystonia	n.a.	yes	no	+
Body Concept Scale	n.a.	yes	no	-
Ways of Coping Checklist	n.a.	yes	no	+
CDQ-24	n.a.	yes	no	+
Oromandibular dystonia questionnaire	n.a.	yes	no	+
Unified Spasmodic Dysphonia Rating Scale	n.a.	no	yes	+
VHI	n.a.	yes	yes	+
Voice Handicap Index 10	n.a.	yes	no	+
Voice-Related Quality of Life	n.a.	yes	no	+
VPQ	n.a.	yes	no	+
Arm Dystonia	<1*	yes	no	±

Disability Scale				
Dystonia Evaluation Scale	<1*	yes	no	+
Tubiana-Chamagne Score	<1*	yes	no	+
Writer's Cramp Rating Scale	n.a.	no	yes	±
Global Dystonia rating Scale	n.a.	no	yes	+
FMDRS	n.a.	yes	yes	±
Unified Dystonia Rating Scale	n.a.	yes	yes	+

Table 4. Dystonia: classification of the scales assessed.

Scale	*Applied in Dystonia*	*Applied beyond authors*	*Succesfull clinometric testing*	*Qualification*
BSDI	yes	yes	yes	Recommended
Jankovic rating scale	yes	yes	no	Suggested
Blepharospasm Disability Scale	yes	yes	no	Suggested
CDIP-58	yes	yes	yes	Recommended
Functional Disability Questionnaire	yes	yes	no	Suggested
TWSTRS	yes	yes	yes	Recommended
Tsui scale	yes	yes	no	Suggested
Modified Tsui scale	yes	no	no	Listed
Freiberg Questionnaire for Dystonia torticollis version	yes	no	no	Listed
Disability questionnaire for patients with cervical dystonia	yes	no	no	Listed
Body Concept Scale	yes	no	yes	Suggested

Ways of Coping Checklist	yes	no	no	Listed
CDQ-24	yes	yes	yes	Recommended
Oromandibular dystonia questionnaire	yes	no	yes	Suggested
Unified Spasmodic Dysphonia Rating Scale	yes	yes	no	Suggested
VHI	yes	yes	yes	Recommended
Voice Handicap Index 10	yes	yes	no	Suggested
Voice-Related Quality of Life	yes	yes	no	Suggested
VPQ	yes	yes	yes	Recommended
Arm Dystonia Disability Scale	yes	yes	no	Suggested
Dystonia Evaluation Scale	yes	no	no	Listed
Tubiana-Chamagne Score	yes	yes	no	Suggested
Writer's Cramp Rating Scale	yes	yes	no	Suggested
Global Dystonia rating Scale	yes	yes	no	Suggested
FMDRS	yes	yes	yes	Recommended
UDRS	yes	yes	no	Suggested

Abbreviations: BSDI, Blepharospasm Disability Index; CDIP-58, Cervical Dystonia Impact Profile; TWSTRS, Toronto Western Spasmodic Torcicollis Rating Scale; CDQ-24, Craniocervical Dystonia Questionnaire; VHI, Voice Handicap Index; VPQ, Vocal Performance Questionnaire; FMDRS, Fahn-Marsden Dystonia Rating Scale, UDRS, Unified Dystonia Rating Scale.

Table 5. Chorea: main characteristics of the scales assessed.

Scale	*Time to complete*	*Patient historical rating*	*Clinical examination*	*Administration burden*
AIMS	10’	No	Yes	+
Marsden & Quinn	n.a.	No	Yes	+
UHDRS	≈ 30’	Semi-objective	Yes	+
mMS	brief	No	Yes	+

Table 6. Chorea: classification of the scales assessed.

Scale	*Applied in HD*	*Applied beyond original authors*	*Successful clinimetric testing*	*Qualification*
UHDRS	Yes	Yes	Yes	Recommended
AIMS	Yes	Yes	Yes[a]	Recommended
mMS	Yes	Yes	No	Recommended[b]
Marsden & Quinn	Yes	Yes	No	Suggested

Abbreviations: UHDRS, Unified Huntington's Disease Rating Scale; [a] *AIMS has several modified versions and it is not entirely clear whether clinimetric analyses are uniform across all versions. mMS, modified motor score of the UHDRS*
[b] *With several limitations*

Table 7. Tics: main characteristics of the scale assessed.

Scale	*Respondent*	*Administration time (minutes)*	*Patient historical rating*	*Clinical examination*	*Administration burden*
DCI	Clinician	n.a.	Yes	No	+
GTRS	Informant, clinician	n.a.	No	No	+
GTS-QoL	Patient	10-15	No	No	±
HMVST	Clinician (child/adolescent, adult)	n.a.	No	No	±
MOVES	Child/adolescent, adult	1-5	No	No	+
PUTS	Child (>10 years)/ adolescent, Adults	5	No	No	+
STSSS	Clinician	5-10	No	No	+
TODS	Informant (parent), clinician (child/adolescent)	5	No	No	+
TS-CGI	Clinician	n.a.	Yes	No	+
TSGS	Clinician (child/adolescent, adult)	n.a.	Yes	No	±
TSQ	Child/adolescent, adult	n.a.	Yes	No	+
TSSL	Patient, informant	n.a.	Yes	No	+
UTRS	Clinician	n.a.		No	+
YGTSS	Clinician (child/adolescent, adult)	15-20	Yes	No	+

Abbreviations: DCI, Diagnostic Confidence Index; GTRS, Global Tic Rating Scale; GTS-QoL, Gilles de la Tourette Syndrome Quality of Life scale; HMVST, The Hopkins Motor and Vocal Tic Scale; MOVES, Motor Tic, Obsessions and Compulsions, Vocal Tic Evaluation Survey; PUTS, Premonitory Urge Tics Scale; STSSS, Shapiro Tourette Syndrome Severity Scale; TODS, Tourette's Disorder Scale; TS-CGI, Tourette Syndrome Global Clinical

Impression; TSGS, Tourette Syndrome Global Scale; TSQ, Tourette Syndrome Questionnaire; TSSL, Tourette Syndrome Symptom List; UTRS, Unified Tic Rating Scale; YGTSS, Yale Global Tic Severity Scale.

Table 8. Tics: classification of the scales assessed.

Scale	*Applied in TS*	*Applied beyond original authors*	*Successful clinimetric testing*	*Qualification*[n]
DCI	Yes	Yes	Yes	Recommended
GTRS	Yes	Yes	Yes	Recommended
GTS-QoL	Yes	Yes	Yes	Recommended
HMVST	Yes	Yes	Yes	Recommended
MOVES	Yes	Yes	Yes	Recommended
PUTS	Yes	Yes	Yes	Recommended
STSSS	Yes	Yes	Yes	Recommended
TODS	Yes	Yes	Yes	Recommended
TS-CGI	Yes	Yes	Yes	Recommended
TSGS	Yes	Yes	Yes	Recommended
TSQ	Yes	Yes	No	Suggested
TSSL	Yes	Yes	No	Suggested
UTRS	Yes	Yes	Yes	Recommended
YGTSS	Yes	Yes	Yes	Recommended

Table 9. Myoclonus: main characteristics of the scales assessed.

Scale	*Time to complete (minutes)*	*Patient historical rating*	*Clinical examination*	*Administration burden*
Chadwick & Marsden	?	no	yes	?
Truong & Fahn	?	no	yes	?
UMRS	15 minutes	yes	yes	+

Abbreviations: UMRS, Unified Myoclonus Rating Scale.

Table 10. Myoclonus: classification of the scales assessed.

Scale	*Applied in myoclonus*	*Applied beyond authors*	*Successful clinimetric testing*	*Qualification*
Chadwick & Marsden	x	x	0	suggested
Troung & Fahn	x	x	0	suggested
UMRS	x	x	x	recommended

Table 11. DID: main characteristics of the scales assessed.

Scale	*Time to complete (minutes)*	*Patient historical rating*	*Clinical examination*	*Administration burden#*
AIMS	15	No	Yes	+
TDRS/ADS	15-20/10-15	No	Yes	+
DIS-co/DISCUS	20/15	No	Yes	+
ESRS/ESRS-A	20*/10*	No	Yes	+
Smith Tardive Dyskinesia Scale	15-20*	No	Yes	+
Modified Rogers Scale	10	No	Yes	+
UKU-SERS	10-30*	Yes (patient version also available)	Yes	+

Abbreviations: AIMS, Abnormal Involuntary Movement Scale; TDRS, (Simpson) Tardive Dyskinesia Rating Scale; MDS-UPDRS, revised version of the UPDRS; ADS, Abbreviated Dyskinesia Scale; DIS-Co, Dyskinesia Identification System-Coldwater; DISCUS, Dyskinesia Identification System Condensed User Scale; UKU-SERS, UKU-Side Effect Rating Scale

Table 12. DID: classification of the scales assessed.

Scale	*Applied in DID*	*Applied beyond original authors*	*Successful clinimetric testing*	*Qualification*
AIMS	X	X	X[a]	Recommended
TDRS/ADS	X	X	0[b]	Suggested
DIS-co/DISCUS	X	X	X	Recommended [c]
ESRS/ESRS-A	X	X	X	Recommended
Smith Tardive Dyskinesia Scale	X	0	0 [b]	Suggested
Modified Rogers Scale	X	0	unclear	Listed
UKU-SERS	X	0	X	Suggested

[a] *Has several modified versions and it is not entirely clear whether clinimetric analyses are uniform across all versions.*

[b] *Intra-rater (test-retest) reliability not evaluated by authors; construct validity not assessed*

[c] *Used only in patients with learning disabilities*

Highlights

- Hyperkinetic disorders represent a heterogeneous group of conditions in which involuntary movements are the prevalent clinical symptoms. They include tremor, dystonia, tics, myoclonus and drug-induced dyskinesia

- Good rating instruments for hyperkinetic disorders are required in both everyday practice and experimental settings

- Unfortunately, the evaluation of these disorders is complicated by the inherent nature and variability over time of involuntary movements

- On the basis of the methodology developed by the MDS task force for rating scales, a scale was defined as ‘Recommended’, ‘Suggested’ or ‘Listed’ in decreasing order of value

- We identified scales defined as ‘Recommended’ for the assessment of all hyperkinetic disorders, with several ‘Recommended’ scales being available for conditions such as tics or dystonia, but only one for myoclonus