The Clinician's Tardive Inventory (CTI):

A New Clinical Tool for Documenting and Rating Tardive Dyskinesia

Richard M. Trosch, MD; Cynthia L. Comella, MD; Stanley N. Caroff, MD; William G. Ondo, MD; Alicia C. Shillington, PhD, MPH; Brandon J. LaChappelle, MPH; Robert A. Hauser, MD; Christoph U. Correll, MD; and Joseph H. Friedman, MD

Abstract

Objective: Current clinician-rated tardive dyskinesia (TD) symptom scales have not addressed the expanding clinical signs and functional impact of TD. The study objective was to develop and test the reliability of a new integrated instrument.

Methods: A movement disorder neurologist devised the outline of the rating scale. A Steering Committee (5 neurologists and 2 psychiatrists) provided revisions until consensus was reached. The Clinician's Tardive Inventory (CTI) assesses abnormal movements of the eye/eyelid/face, tongue/mouth, jaw, and limb/trunk; complex movements defined as complicated movements different from simple patterned movements or postures; and vocalizations. The CTI rates frequency of symptoms from 0 to 3 (ranging from absent to constant). Functional impairments, including activities of daily living (ADL), social impairment, symptom distress, and physical harm, are rated 0–3 (ranging from unawareness to severe impact). The CTI underwent interrater and test-retest reliability testing between February and June 2022 based on videos and accompanying vignettes, which were reviewed by 2 movement disorder specialists to determine adequacy. Four clinicians rated each video/vignette. Interrater agreement was analyzed via 2-way random-effects intraclass correlation (ICC), and test-retest agreement was assessed utilizing the Kendall tau-b.

Results: Forty-five video/vignettes were assessed for interrater reliability and 16 for test-retest reliability. The most prevalent movements were those of the tongue and mouth (77.8%) and jaw (55.6%). ICCs for movement frequency for anatomic symptoms were as follows: anatomic symptom summary score 0.92, abnormal eye movement 0.89, abnormal tongue/mouth movement 0.91, abnormal jaw movement 0.89, abnormal limb movement 0.76, complex movement 0.87, and abnormal vocalization 0.77; ICCs for functional impairments were as follows: total impairment score 0.92, physical harm 0.82, social embarrassment 0.88, ADLs 0.83, and symptom bother 0.92; Retests were conducted a mean (SD) of 15 (3) days later with correlation coefficients ranging from 0.66 to 0.87.

Conclusions: The CTI is a new integrated instrument with proven reliability in assessing TD signs and functional impacts. Future validation study is warranted.

J Clin Psychiatry 2024;85(1):23m14886

Author affiliations are listed at the end of this article.

Tardive dyskinesia (TD) is a persistent neurologic adverse effect following exposure to dopamine receptor blocking agents (DRBAs), including antipsychotics and antiemetics.[1,2] The reported incidence of TD varies across studies and is dependent on several risk factors, including study population age, DRBA class, dose and duration of exposure, treatment-emergent parkinsonian symptoms or acute DRBA movements, and anticholinergic medication cotreatment. A prospective study of elderly DRBA-naive patients treated with second-generation DRBAs found a 2-year TD incidence between 7.2% and 11.1%.[3] The estimated prevalence of TD is approximately 15%–40% in the United States in those exposed to DRBAs.[4–7]

TD is manifested by a variety of hyperkinetic involuntary movements, including stereotypies, chorea, dystonia, myoclonus, abnormal respiration, akathisia, or pain. Stereotypy is the most frequent TD phenomena and may also include ritualistic repetitive gestures and movements affecting the limbs, trunk, or face.[8–12] Tardive dystonia is manifest as sustained muscle contractions that cause twitching and repetitive movements and abnormal posturing.[13–15] Cervical dystonia, laryngeal dystonia, blepharospasm, bruxism, and lateral trunk flexion are common tardive dystonia movements. Tardive akathisia is experienced as a distressing inner restlessness accompanied by a variety of movements, including tapping, squirming, rocking from foot to foot, truncal rocking, and vocalizations including grunting

Clinical Points

- Documenting accurate severity and functional impact of tardive dyskinesia (TD) is essential. Current instruments have not kept up with the expanding spectrum of TD involuntary movements and lack clear definition of severity and ability to assess movement nuances, fluctuation, and impact. Reliability issues persist, hindering everyday practice.
- The Clinician's Tardive Inventory (CTI) can aid TD assessment and examination documentation. Its quantitative nature permits TD severity assessment for practice and trials after planned validation. The CTI is reliable in symptom assessment, including movements, frequency, and functional impacts. It has the potential for telemedicine use based on video ratings reliability.

and moaning.[16,17] Unlike the more common acute akathisia, tardive akathisia, similar to all other tardive subsyndromes, is persistent and does not quickly resolve with discontinuation of the offending agent.[18] Tardive myoclonus is characterized by unpredictable jerking of the face, neck, and limbs, particularly the upper limbs or trunk.[16,19,20] Case series have suggested that tics and tremors may be TD phenomena[15,16,20]; however, the common occurrence of these phenomena in the general population makes confirmation challenging. Chorea, defined by continuous, jerky or brief, irregular, non-patterned, flowing movements randomly moving from one area to another, can occur in TD; however, its occurrence compared to other TD phenomena is less common. A retrospective review of TD videos from a movement disorder clinic found chorea diagnosed in 3.8%, representing the least common TD movement phenomena observed in this series.[19] TD-afflicted individuals may have any of the aforementioned phenomena alone or in combinations. Importantly, TD is associated with poor quality of life, social withdrawal related to stigma, and increased morbidity and mortality.[21–24]

With the recent availability of US Food and Drug Administration (FDA)–approved medications for TD, accurate TD diagnosis is needed to identify patients who could benefit from treatment.[25,26] Clinical trials of new treatments frequently rely primarily on the Abnormal Involuntary Movement Scale (AIMS). However, this instrument has significant limitations, as it has not kept up with the expanding variety of movements seen with the evolving TD phenomenology. The AIMS was initially developed by the National Institute of Mental Health for use in a research setting,[27] but it has noted difficulties of translation into clinical practice.[26] The AIMS lacks adequate instruction and documents movements only by anatomy (eg, arms, trunk), rated on a numeric scale of a physician's determination of "severity." The measure of "severity" is not defined and could be interpreted by the rater to mean movement amplitude, rapidity, frequency, functional impact, or some combination of these features, leading to inconsistency in ratings. It also lacks descriptors of the 0–4 anchors used in scoring, so incremental differences in this ordinal scale are unknown.[26] The AIMS does not require a description of these movements beyond their location (possibly including non-TD phenomena), does not take into consideration symptom fluctuation between assessments,[28] and does not include observed movement frequency or functional impacts. The lack of documentation of the impact of TD on the patient's life, including social, physical, vocational, and psychological functioning, is a significant limitation.[29]

Regardless, the AIMS has been used as a main outcome representing efficacy in TD clinical trials despite issues of reliability and, until recently, having no established minimally clinically important difference (MCID).[30,31] Recent attempts to establish MCID for the AIMS are limited by being conducted in highly selected clinical trial populations with experienced researcher-raters, and therefore value of this outcome for everyday practice is not clear.

Given the need for an updated and integrated scale to facilitate the recognition, assessment, and documentation of the spectrum, frequency, and impact of TD, the objective of this study was to develop and assess the reliability of a new instrument, the Clinician's Tardive Inventory (CTI), to address the shortcomings of existing instruments.

METHODS

CTI Development

A movement disorder neurologist (R.M.T.) devised the outline of the new CTI scale; input from a Steering Committee (5 neurologists and 2 psychiatrists) provided further refinements. The CTI was developed to facilitate assessment and documentation of 6 anatomically manifest TD signs, including the presence, frequency, and functional impacts of abnormal movements of the eye/eyelid/face, tongue/mouth, jaw, and limb/trunk; complex movements (eg, handwringing, self-caressing); and vocalizations. The anatomic assessment includes descriptors to assist both experienced and less experienced clinicians in identifying and documenting observed movements. Arranged anatomically, we attempted to comprehensively, but not exhaustively, include involuntary movements encompassing the recognized TD phenomenon of stereotypy, dystonia, myoclonus, abnormal respirations, chorea, and akathisia. The Steering Committee convened and reviewed the outline of the new instrument. They provided revisions until consensus was reached. Consensus required unanimous agreement by all members. The CTI scores 6 anatomic symptoms and yields a total functional symptoms score. Symptoms scored include abnormal eye/eyelid/facial movements, abnormal tongue and mouth movements, abnormal jaw movements, complex movements, abnormal limb/trunk

movements, and abnormal vocalizations. Symptoms are rated from 0 to 3 with 0 = absent; 1 = infrequent/intermittent or present only with activating maneuvers (< 40% present); 2 = frequent intermittent, brief periods without movements (40%–75% present); and 3 = constant or nearly constant (> 75% present). An anatomic symptom summary score is calculated by summing the frequency scores for individual symptoms and ranges from 0 to 18. Functional impairments scored include activities of daily living (ADL), leisure or occupation, social impairment, symptom distress, and physical harm rated 0–3 with 0 = patient is unaware or unaffected, 1 = symptoms mildly impact patient, 2 = symptoms moderately impact patient, and 3 = symptoms severely impact patient. A holistic total functional impairment score ranges from 0 to 3 with 0 = patient is unaware or unaffected and 3 = symptoms severely impact patient and is calculated as the maximum impact score among social impairment ADL, symptom distress, and physical harm. Within each symptom, specific individual movements (eg, tongue darting, fidgety leg movements) and patient-reported functional impairments (eg, dry mouth, difficulty lying in bed) are documented via check boxes as present/absent to serve as clinical documentation of a patient encounter.

CTI Appraisal

Following institutional review board approval, the CTI underwent interrater and test-retest reliability assessment between February and June 2022. Videos of patient TD neurologic examinations were obtained following subject informed consent for use. Forty of the videos were from existing medical records filmed by practices that typically video patient encounters as a part of routine care, and 5 were recorded for the purposes of this study. All videos were drawn from the clinical practices of 2 movement disorder specialists (R.M.T. and J.H.F.). Vignettes were extracted from existing medical records. Before inclusion in the study, each video was reviewed by 2 movement disorder specialists to confirm the diagnosis of TD by consensus and the videos' technical adequacy to demonstrate a TD-consistent movement. Videos had to demonstrate a clearly identifiable movement disorder consistent with TD as determined by a panel of movement disorder clinicians (R.M.T. and J.H.F.). A consensus of 2 experts was required of each case to certify that the diagnosis of TD was met. Videos were edited to a standard length of no more than 3 minutes. Clinical vignettes were created from the subject's medical records to accompany each video and comprised the subjects' documented symptom descriptions and functional, social, or occupational impairments/limitations. Data for vignettes included patient's age at symptom onset, current or most recent DRBA agent exposure, duration of DRBA exposure, symptom occurrence relative to DRBA exposure, presence of other comorbidities that may account for symptoms (eg, edentulous dyskinesia), subjects' subjective description of symptoms, and subjects' reported functional, social, or occupational impairments/limitations due to TD. For subjective symptom and functional impairments that were not available in the medical record, fictional descriptions were created.

Four clinicians from a pool of 5 (4 neurologists and 1 psychiatrist) who were movement disorder specialists (W.G.O., C.L.C., S.N.C., J.H.F., and R.M.T.) were selected to view each of the 45 videos and vignettes and rate each using the newly developed CTI. Raters were provided with training during a 120-minute teleconference, along with sample guidance and instructions for assessment and CTI completion. Clinicians were instructed to complete the review of each video-vignette pair in one setting to approximate the length of time of a clinic-based tardive examination and indicate the presence or absence of symptoms, their frequency, and functional impairments noted. Three raters were involved in reviewing and providing specific edits to refine the scale in the initial development process. The scale development was completed prior to videos' being viewed. Raters assigned to view each video included a mix of individuals who had and had not provided input into the scale development. All raters performed a retest of 3–5 randomly selected videos. One-third of the videos were included in the retest. Non-rating research staff selected retest videos after review of level of detail available from medical records to populate vignettes. Videos were accessed from a password-protected and encrypted central location, and the CTI was programmed for administration online via Qualtrics.

Statistical Considerations

To determine the N = 45 sample size calculation, we examined 3 assumptions. In each, we used the null for intraclass correlation (ICC) of 0.40 (poor "agreement"). We tested, against the null, 3 alternatives assuming an ICC of 0.6, 0.7, and 0.75 (at least moderate "agreement"). We set significance at .05 and tested assumptions to achieve a minimum 80% power using an ICC of 0.6 as our alternative hypothesis. Assuming an N = 45, with 4 observer ratings of each subject, we would have > 80% power to detect an intraclass correlation of 0.6 (at least modest "agreement") under the alternative hypothesis when the intraclass correlation under the null hypothesis is 0.4 (poor "agreement") using an *F* test with a significance level of 5%. ICCs for movements and agreement between raters were at the main symptom category level (eg, any tongue or mouth movement) as opposed to the individual movement level (eg, intraoral tongue movements or tongue darting).

By evaluating the interrater consistency in using CTI scoring, the reliability of the instrument was assessed. A high interrater consistency for individual items was assumed to indicate that the CTI, based on expert clinician development, has utility in identifying the presence or absence and severity of symptoms by clinicians. ICC was calculated by mean squares (ie, estimates of the population

variances based on the variability among a given set of measures) obtained through analysis of variance. For this specific analysis, assuming our raters are drawn from a larger population of raters, a 2-way random-effects model was used (ICC) to generalize our reliability results to raters who possess characteristics similar to those of the selected raters in the reliability study. The inferred reliability from the ICC values was classified as follows: > 0.90 = excellent agreement between raters, 0.75–0.90 = good agreement, and < 0.75–0.50 = moderate agreement.[32] Individual movements were considered present if a minimum 3 of 4 raters indicated as such and the proportion of patients in the sample displaying movements as agreed upon by at least 3 raters was reported.

Test-retest reliability was conducted to determine consistency between two measurements (tests) using the CTI when the same rater evaluated the same video at two different points in time. Test-retest agreement assessment utilized The Kendall tau-b correlation coefficient (CC) based on a Fisher *r*-to-*z* transformation. The correlation in observations between the two tests was used as an estimate of test-retest reliability and coefficients were computed pairwise between the raters to evaluate the agreement. The Kendall tau-b varies between −1 (100% negative association, or perfect inversion) and +1 (100% positive association, or perfect agreement). A value of zero indicates the absence of association. For testing a hypothesis of positive association, we interpret 1 = perfect reliability, ≥ 0.9 = excellent reliability, ≥ 0.8 < 0.9 = good reliability, ≥ 0.7 < 0.8 = acceptable reliability, ≥ 0.6 < 0.7 = questionable reliability, ≥ 0.5 < 0.6 = poor reliability, < 0.5 = unacceptable reliability, and 0 = no reliability.[33]

RESULTS

Forty-five video and vignette pairs were assessed. The most commonly observed movements were those of the tongue and mouth (77.8%) (including lip pursing/puckering/curling/protrusion, tongue darting [fly catcher's tongue], lip licking, and intraoral tongue movements [bon-bon sign]) and jaw (55.6%) (including opening/closing/displacement to one side, chewing, and grinding). Least commonly observed were vocalizations (13.3%) (irregular or noisy respirations, grunting, spasmodic dysphonia, moaning, or humming) (Table 1). The mean (SD) time to complete the CTI was 7 (6) minutes.

Interrater Reliability

Each video and vignette pair was evaluated by 4 raters. ICC for the anatomic symptom summary score total was 0.922. ICCs for individual anatomic symptoms ranged from a high of 0.907 for abnormal tongue/mouth movement frequency to a low of 0.763 for abnormal limb/trunk movement frequency. The ICC for the total functional impairment score was 0.920. ICCs for individual functional impairments ranged from a high of 0.922

Table 1.
Proportion of Patients Displaying of Movements in 45 Rated Videos

Type of Movement	n (%)
Any Tongue and Mouth Movements	35 (77.8)
Lip (pursing/puckering/curling/protrusion)	24 (53.3)
Tongue darting (fly catcher's tongue)	19 (42.2)
Lip licking	14 (31.1)
Intraoral tongue movements (bon-bon sign)	12 (26.7)
Retracting or elevating corners of mouth (bridling)	11 (24.4)
Lip smacking	9 (20.0)
Enlarged tongue (macroglossia)	5 (11.1)
Puffing out cheeks or blowing	0
Any Jaw Movements	25 (55.6)
Involuntary abnormal jaw movements (opening/closing/displacement to one side)	16 (35.6)
Chewing	5 (11.1)
Grinding	0
Any Complex Movements	17 (37.8)
Leg swinging, bouncing, tapping, repetitive crossing/uncrossing	17 (37.8)
Fidgety hand movements/picking	7 (15.6)
Trunk movement	3 (6.7)
Handwringing or fingers intertwined	1 (2.2)
Self-caressing	0
Inability to stand still	0
Postural adjustment	0
Unable to remain seated (akathisia) (standing/pacing)	0
Any Trunk and Limb Movements	16 (35.6)
Repetitive/patterned foot or hand movements	7 (15.6)
Neck posturing	4 (8.9)
Truncal flexion/extension	2 (4.4)
Twisting limb movements (arm/hand/leg/feet)	1 (2.2)
Pelvic thrusting	1 (2.2)
Jerking limb movements at (rest posture/end intention) (myoclonus)	0
Any Eye/Eyelid/Facial Movements	14 (31.1)
Forceful eyelid closure (blepharospasm)	9 (20.0)
Excessive blinking (blepharoclonus)	8 (17.8)
Repetitive brow elevation/depression	3 (6.7)
Facial grimace (sustained)	2 (4.4)
Eyes driven up (oculogyric deviations)	1 (2.2)
Any Abnormal Vocalizations	6 (13.3)
Irregular or noisy respirations	5 (11.1)
Grunting	0
Spasmodic dysphonia	0
Moaning	0
Humming	0

Table 2.

Interrater Reliability for the Clinician's Tardive Inventory (CTI)

Variable	ICC	95% CI Lower	95% CI Upper	*P*
Total Anatomic Symptom Score	0.922	0.873	0.954	<.001
Abnormal eye movement frequency	0.893	0.831	0.936	<.001
Abnormal tongue/mouth movement frequency	0.907	0.798	0.953	<.001
Abnormal jaw movement frequency	0.894	0.761	0.948	<.001
Complex movement frequency	0.868	0.736	0.931	<.001
Abnormal limb/trunk movement frequency	0.763	0.514	0.879	<.001
Abnormal vocalizations frequency	0.774	0.622	0.870	<.001
Total Functional Impairment Score	0.920	0.809	0.962	<.001
Harm functional impairment score	0.824	0.722	0.895	<.001
Social embarrassment impairment score	0.882	0.765	0.938	<.001
Activities of daily living impairment score	0.826	0.680	0.905	<.001
Symptom bother impairment score	0.922	0.875	0.954	<.001

Abbreviation: ICC = intraclass correlation coefficient.

Table 3.

Intrarater (Test-Retest) Reliability for the Clinician's Tardive Inventory (CTI)

Variable	CC	95% CI Lower	95% CI Upper	P
Anatomic Symptom Score	0.892	0.767	0.953	<.001
Abnormal eye movement frequency	0.872	0.785	0.925	<.001
Abnormal tongue/mouth movement frequency	0.761	0.615	0.857	<.001
Abnormal jaw movement frequency	0.781	0.644	0.869	<.001
Complex movement frequency	0.767	0.624	0.861	<.001
Abnormal limb/trunk movement frequency	0.805	0.681	0.885	<.001
Abnormal vocalizations frequency	0.774	0.633	0.865	<.001
Total Functional Impairment Score	0.707	0.467	0.850	.002
Harm global impression score	0.739	0.519	0.867	<.001
Social embarrassment global impression score	0.854	0.715	0.928	<.001
Activities of daily living global impression score	0.857	0.720	0.930	<.001
Symptom bother global impression score	0.659	0.394	0.823	.004

Abbreviation: CC = correlation coefficient.

for symptom bother to a low of 0.826 for functional impairment of activities of daily living (Table 2).

Intrarater Reliability

All raters performed retests of 3–5 randomly selected videos. Retests were conducted a mean (SD) of 15 (3) days later. Correlation coefficient for the anatomic symptom summary score total was 0.892. Correlation coefficients for anatomic symptoms ranged from a high of 0.872 for abnormal eye frequency, to a low of 0.767 for complex movements. CCs for functional impairments ranged from a high of 0.857 for impaired activities of daily living to a low of 0.659 for physician's assessment of symptom distress. Overall functional impairment CC was 0.707. (Table 3)

DISCUSSION

The CTI was developed to address an unmet need for a comprehensive instrument to adequately document the recognized movements associated with the various TD phenomenology and also provide an assessment of the impact of TD symptoms on the patient's life. Recent consensus panel recommendations indicate that these factors are important to understand how urgently TD symptoms should be addressed and take into consideration symptoms' impact on social interactions, physical impairment, vocational function, and psychological distress.[30] To address this unmet need, the newly developed CTI encompasses 5-item TD movement domains of stereotypy, dystonia, akathisia, myoclonus, and vocalizations. It includes an accompanied listing of the symptoms and manifestations known to commonly occur in each domain and the frequency with which they occur, and, importantly, it incorporates an assessment of functional complaints most often associated with different involuntary movements.

The results of this study indicate that expert clinicians using the CTI had good to excellent reliability in identifying total anatomic symptoms and frequency and very good reliability in identifying individual anatomic symptoms. Likewise, summary functional impairment scores also displayed excellent reliability, and individual impairment scores displayed good to excellent reliability. Test-retest (intrarater) agreement was also fair to very good. An exception was related to the clinician's assessment of subjects' "bother" related to symptoms, which fell into the "questionable" reliability interpretation range. This result impacted the total functional impairments score, which also fell near the "questionable" range. This finding may be accounted for by difficulty of clinician's attempting to rate a patient's subjective experience of impairments without having an opportunity to query a patient during a face-to-face encounter. The interrater reliability scores were higher than the intrarater scores, a result that was also attributed to the limitations of relying on vignettes.

To date, clinical trials of new antipsychotic medications and interventions for TD have relied primarily on the AIMS. The AIMS was initially developed by the National Institute of Mental Health. Although widely accepted and standardized, the AIMS has significant limitations, as it has not kept up with the expanding understanding of TD phenomenology. It was developed for use in a research setting[27]; however, it has noted difficulties of translation into clinical practice.[26] As it does not address current and specific movements descriptively, the AIMS may have limitations in documenting a TD examination

in clinical practice or following the course of changing movements over time. The AIMS measure of "severity" is not defined; instead, it is at the discretion of the rater to determine if severity will be scored by the observed movement amplitude, speed, frequency, functional impact, or some combination of these features. Most importantly, the AIMS does not address the impact of TD on the patient's life including social, physical, vocational, and psychological functioning, hindering the assessment of the need for a therapeutic intervention.[26] The CTI addresses these limitations; however, it is recognized that a process to revise the AIMS is also now underway via an expert panel using a Delphi process.[26]

A feature of the CTI considered in development was the need for wording to be "user-friendly" for less experienced raters. The Steering Committee structured the CTI in a manner that uses descriptive language of specific movements anatomically presented (as opposed to domains) toward this aim. The CTI was designed to include guidance for a brief examination for use in clinical practice to assess tardive movements and to facilitate the rapid recording of a TD examination. Another key feature is that documentation of the anatomic location and frequency of movements can be used to follow the course of TD and responses to therapeutic interventions. This would require documentation of which specific movements are present, along with their observed frequency in addition to their associated subjective symptoms to measure functional impact, which the CTI facilitates. It is important to note that the proportion of patients displaying specific involuntary movements described herein includes only those seen on this collection of videos; it does not imply a population prevalence. Healthy controls and other movement disorders were not included, as these non-afflicted subjects were not necessary to generate interrater reliability in assessing TD movements. This scale is not a phenomenological scale and does not ask the rater to identify movements by phenomenological categories (eg, myoclonus, stereotypy, dystonia, akathisia), which are often subject to debate and opinion. However, to document the movements seen in TD of various phenomenologies, each is listed in descriptive terms (eg, blinking, repetitive foot tapping, hand wringing), which we believe is an objective and improved method compared with the AIMS. We took this approach because not even very experienced movement disorder specialists will agree on the phenomenology of a specific movement and inexperienced clinicians may not be familiar with the known TD phenomenology.

To be practical, our Steering Committee decided the spectrum of movements listed in the CTI should be comprehensive, but not exhaustive. As an example, tardive pain, because of its rarity, was not included, but may be added when present under the option of "other." The CTI does not rate tardive chorea, as it is rare (approximately 3%)[34]; most TD movements are repetitive and patterned, which excludes chorea and athetosis. By limiting the rater to choosing different movements occurring within the spectrum of recognized TD phenomena, misdiagnosis is less likely. Additionally, with repeated use, this feature may serve as a heuristic goal, familiarizing untrained raters with the spectrum of possible TD involuntary movements. The CTI also includes many TD involuntary movements not assessed by prior rating scales, including abnormal respirations, oculogyric deviations, and inner restlessness. Electronic medical record–compatible and iOS app versions of the scale are in development, which should speed the completion of the scale.

Limitations

This study is limited by several factors. First, assessment of the CTI relied on video examinations, many of which previously existed from the clinical practices of two movements disorder specialists (R.M.T. and J.H.F.). These video tapes were created to document TD symptoms in clinical practice, and although they were edited to produce a standardized examination, as they were not initially filmed for research purposes, it is possible that they may have missed documenting the presence or absence of certain movements. For a minority of the existing videos, if a movement could not be assessed in a video (eg, assessment of foot tapping when an existing video did not clearly show the subject's feet), raters were instructed to assume that the movement was absent. This could have overestimated agreement, increasing our ICC values. However, unlike many symptoms in psychiatry, abnormal movements are objectively observable and usually not open to interpretation. If raters observe tongue protrusion or eye blinking, they easily come to consensus that the patient stuck their tongue out and blinked. Therefore, the CTI does not require agreement concerning the classification of the phenomenon observed (eg, dystonic vs chorea vs stereotypy), instead relying on a simple description of the movement (eg, blinking, chewing, grunting), which is much less subject to interpretation bias. Second, since this study was conducted using videos and not during live encounters, it is not clear that raters' agreement would be as highly correlated if assessing a live patient in a clinical setting. An actual clinical encounter would facilitate patient assessment and present fewer challenges to performing a rating. However, remote video assessment of TD by central raters has become standard in clinical trials. Also, because a considerable number of clinical encounters are now via telemedicine, we believed it is important to show the applicability of a scale to a video examination. Regardless, it has previously been established that raters of videotaped TD assessments can achieve very good agreement examining symptoms in this manner, particularly as videotapes produce a reliable record of orofacial dyskinesia, one of the most common TD symptoms. [35] The videos used included examples of tardive oculogyric movements, respirations, akathisia, dystonia, blepharospasm, and

other "non-classical" movements that accurately reflect the spectrum of TD movements seen in clinical practice. Videos reflected the bias inherent in who was referred to clinician's clinics, the clinician's decision to create a video, and the subject's willingness to allow this. The study subjects represent a diverse population of TD patients, and the CTI scale demonstrated very good interrater reliability.

Third, not all functional impact was documented in the subjects' medical records or assessable during the videotaped encounters (eg, impacts on social and employment functioning). Clinician's assessment of functional disability was not based on movements depicted in videos but relied wholly on the patient's description of their experience extracted verbatim from their medical records. However, subjects' perception of "bother" related to symptoms was also not clearly documented in videos or in some of the medical records supporting the vignette creation, which may have led to raters having only "fair" agreement for this parameter in intrarater assessments. It is possible without clear assessment in a live encounter, including asking a patient how much they were bothered by their symptoms, that raters interpreted the degree of "bother" differently in retests. Fourth, the Steering Committee did not ultimately include items assessing and documenting TD-related pain, as it was believed to be a rare phenomenon, and the intent for the CTI was to be comprehensive, but not exhaustive. Future versions of the CTI may revisit the addition of pain symptoms to incorporate that less frequent but potentially clinically impactful finding. Fifth, this study incorporated only face validity in its construction with use of experts in TD. Future study should be conducted to understand the generalizability to generalist psychiatrists, psychiatry trainees, and/or other clinicians who are not experts in movement disorders. Sixth, videos used in the study may not have been representative of a non-selected general population; rather, they were drawn from two movement disorder practices and were limited to those individuals that provided consent to have videos used in the study. An attempt was made to include videos representing most movements included on the scale, including less common movements. Therefore, the reported proportion of patients displaying involuntary movements may not reflect a general TD population. This scale was initially conceived as a clinical tool for assessing TD and for the documentation of the tardive examination. However, it is also a quantitative scale and may appropriately be used to assess TD severity in clinical trials with proper validation study, which is planned. Future study should include more representation from general psychiatrists and neurologists and other non-expert health care providers. Finally, as in all movement assessments, there may be significant variations during the course of a day or even over minutes. A full validation study, including the use of the "gold standard" AIMS assessment instrument for comparison to accomplish this, is necessary.

CONCLUSIONS

The CTI is a new integrated instrument developed by an expert Steering Committee of neurologists and psychiatrists. This scale was initially conceived as a clinical tool for assessing TD and for the documentation of the tardive examination. However, it is also a quantitative scale and may appropriately be used to assess TD severity in clinical trials with proper validation study, which is planned. The instrument has shown face validity and good reliability in assessing the presence of TD signs by description of movements, frequency, and functional impacts. This study was limited in that only movements observable on video could be rated; however, good reliability in this setting suggests the CTI could be applicable to telemedicine encounters. Future studies are needed to determine validity versus currently utilized scales and generalizability to routine clinical care and to understand minimally clinically important differences in longitudinal patient follow-up in a clinical setting and its potential validity as a clinical rating instrument.

Article Information

Published Online: January 24, 2024. https://doi.org/10.4088/JCP.23m14886

Submitted: April 4, 2023; accepted October 16, 2023.

To Cite: Trosch RM, Comella CL, Caroff SN, et al. The Clinician's Tardive Inventory (CTI): a new clinical tool for documenting and rating tardive dyskinesia. *J Clin Psychiatry.* 2024;85(1):23m14886.

Author Affiliations: Oakland University, William Beaumont Medical School, Rochester, Michigan (Trosch); Rush University, Chicago, Illinois (Comella); University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania (Caroff); Houston Methodist Neurological Institute, Houston, Texas (Ondo); Weill Cornell Medical School, New York, New York, (Ondo); EPI-Q Inc, Oakbrook, Illinois (Shillington, LaChappelle); University of South Florida, Tampa, Florida (Hauser); The Zucker Hillside Hospital, Department of Psychiatry, Northwell Health, Glen Oaks, New York (Correll); Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Department of Psychiatry and Molecular Medicine, Hempstead, New York (Correll); Charité Universitätsmedizin Berlin, Department of Child and Adolescent Psychiatry, Berlin, Germany (Correll); Brown University, Butler Hospital, Providence, Rhode Island (Friedman).

Corresponding Author: Alicia C. Shillington, PhD, MPH, EPI-Q Inc, 230 E. Ohio St, Ste 410 1064, Chicago, IL 60611 (alicia.shillington@epi-q.com).

Relevant Financial Disclosures: Dr Trosch has received consulting fees in the last 24 months from EPI-Q, Amneal, Ipsen, Supernus, and Neurocrine Biosciences; has received speaker fees from Amneal, Ipsen, Supernus, and Neurocrine Biosciences; and has received funds from Neurocrine Biosciences. **Dr Comella** has received consulting fees from Acorda, Allergan, Lundbeck, Merz, Acadia, Ipsen, Jazz, Adamas, Neurocrine Biosciences, Revance, Sunovion, AEON Biopharma, and EPI-Q; receives royalties from Cambridge, Wolters Kluwer, and EPI-Q; and serves on the editorial board of *Clinical Neuropharmacology* and *Sleep Medicine*. **Dr Caroff** receives royalties from Wolters Kluwer and American Psychiatric Press, Inc; receives research grant support from Eagle Pharmaceuticals; and receives consultant fees from EPI-Q and Neuroscience Education Institute. **Dr Ondo** receives research grants from Biogen, Lundbeck, Sun, Restless Legs Syndrome Foundation, Parkinson's Study Group, Sun, Dystonia Coalition (NIH), Biogen, Cerevel, SCION, Harmony, Sunovion, and Revance; receives speaker fees from Teva, Acadia, Acorda, Neurocrine Biosciences, USWorldMeds, Allergan, Kyowa Kirin, and Sunovion; receives consulting fees from Takeda, Merz, Jazz, XWPharma,, Emalex, Supernus, Amneal, and EPI-Q; and receives royalties from *Restless Legs Syndrome*, *Movement Disorders in Clinical Practice*, and *UpToDate*. **Dr Shillington** is an employee and shareholder of EPI-Q Inc. EPI-Q received funding to support the conduct of this study from Neurocrine Biosciences. **Dr LaChappelle** is an employee of EPI-Q Inc. EPI-Q received funding to support the conduct of this study from Neurocrine Biosciences. **Dr Hauser** has received consulting fees in the last 24 months from Abbvie, Acorda, Adamas, Alterity, Amneal, Cerevance, Curium,

Enterin, Inhibikase, Jazz, Kyowa Kirin, Merck, Merz, Neurocrine Biosciences, Orion, Pharmather, Sage, Scion, Sunovion, Sio Gene Therapies, Supernus, Tolmar, Vivifi Biotech, Ovid, Pharma Two B, Global Kinetics, PD Neurotechnology, BlueRock, Neuroderm, Lotus, and EPI-Q Inc. Dr Hauser has received speaker fees from Acorda, Adamas, Amneal, Cerevel, Kyowa Kirin, Neurocrine Biosciences, Supernus, and Inhibikase. Institutional Research Funding: Dr Hauser's institution has received funds from Enterin, Roche, Biogen, Pharma Two B, Revance, Impax, Covance/Sun, Global Kinetics, Neuraly, Centogene, NeuroDerm, Abbvie, Cerevance, Cerevel, UCB, Lundbeck, Sanofi, Micheal J. Fox Foundation, AEON Biopharma, and the Parkinson Foundation. Dr Hauser holds stock in Revance Therapeutics and stock options in Axial Therapeutics, Inhibikase, and Enterin. He has received fees from USF licensing of a PD diary. Dr Hauser acknowledges a Center of Excellence grant from the Parkinson Foundation. **Dr Friedman** receives royalties from Cambridge University Press, Karger Press, and Medlink and has received consulting fees for EPI-Q. **Dr Correll** has received consultant fees from AbbVie, Acadia, Alkermes, Allergan, Angelini, Aristo, Boehringer-Ingelheim, Cardio Diagnostics, Cerevel, CNX Therapeutics, Compass Pathways, Darnitsa, Gedeon Richter, Hikma, Holmusk, IntraCellular Therapies, Janssen/Johnson & Johnson, Karuna, LB Pharma, Lundbeck, MedAvante-ProPhase, MedInCell, Merck, Mindpax, Mitsubishi Tanabe Pharma, Mylan, Neurocrine Biosciences, Newron, Noven, Otsuka, Pharmabrain, PPD Biotech, Recordati, Relmada, Reviva, Rovi, Seqirus, SK Life Science, Sunovion, Sun Pharma, Supernus, Takeda, Teva, and Viatris; has provided expert testimony for Janssen and Otsuka; has served on a Data Safety Monitoring Board for Compass, Lundbeck, Relmada, Reviva, Rovi, Supernus, and Teva; has received research support from Janssen and Takeda; receives royalties from UpToDate; and has stock options for Diagnostics, Mindpax, LB Pharma, and Quantic.

Funding/Support: Financial support in the form of an investigator-initiated study grant was received from Neurocrine Biosciences, Inc.

Role of the Funders/Sponsors: Neurocrine played no role in the design, conduct, or interpretation of the study.

References

1. Bhidayasiri R, Boonyawairoj S. Spectrum of tardive syndromes: clinical recognition and management. *Postgrad Med J*. 2011;87(1024):132–141.
2. Aquino CC, Lang AE. Tardive dyskinesia syndromes: current concepts. *Parkinsonism Relat Disord*. 2014;20(suppl 1):S113–S117.
3. Woerner MG, Correll CU, Alvir JMJ, et al. Incidence of tardive dyskinesia with risperidone or olanzapine in the elderly: results from a 2-year, prospective study in antipsychotic-naïve patients. *Neuropsychopharmacology*. 2011;36(8):1738–1746.
4. Caroff SN, Ungvari GS, Cunningham Owens DG. Historical perspectives on tardive dyskinesia. *J Neurol Sci*. 2018;389:4–9.
5. Woods SW, Morgenstern H, Saksa JR, et al. Incidence of tardive dyskinesia with atypical versus conventional antipsychotic medications: a prospective cohort study. *J Clin Psychiatry*. 2010;71(4):463–474.
6. Carbon M, Hsieh CH, Kane JM, et al. Tardive dyskinesia prevalence in the period of second-generation antipsychotic use: a meta-analysis. *J Clin Psychiatry*. 2017;78(3):e264–e278.
7. Woerner MG, Alvir JM, Saltz BL, et al. Prospective study of tardive dyskinesia in the elderly: rates and risk factors. *Am J Psychiatry*. 1998;155(11):1521–1528.
8. Lerner V, Miodownik C. Motor symptoms of schizophrenia: is tardive dyskinesia a symptom or side effect? a modern treatment. *Curr Psychiatry Rep*. 2011;13(4):295–304.
9. Lerner PP, Miodownik C, Lerner V. Tardive dyskinesia (syndrome): current concept and modern approaches to its management. *Psychiatry Clin Neurosci*. 2015;69(6):321–334.
10. Ward KM, Citrome L; Tardive Dyskinesia-Key Differences in Pathophysiology and Clinical Management. Antipsychotic-related movement disorders: drug-induced Parkinsonism vs tardive dyskinesia—key differences in pathophysiology and clinical management. *Neurol Ther*. 2018;7(2):233–248.
11. Waln O, Jankovic J. An update on tardive dyskinesia: from phenomenology to treatment. *Tremor Other Hyperkinet Mov (N Y)*. 2013;3(0):tre-03-161-4138-1.
12. Factor SA. Management of tardive syndrome: medications and surgical treatments. *Neurotherapeutics*. 2020;17(4):1694–1712.
13. Burke RE, Fahn S, Jankovic J, et al. Tardive dystonia: late-onset and persistent dystonia caused by antipsychotic drugs. *Neurology*. 1982;32(12):1335–1346.
14. Kiriakakis V, Bhatia KP, Quinn NP, et al. The natural history of tardive dystonia: a long-term follow-up study of 107 cases. *Brain*. 1998;121(Pt 11):2053–2066.
15. Skidmore F, Reich SG. Tardive Dystonia. *Curr Treat Options Neurol*. 2005;7(3):231–236.
16. DeLeon ML, Jankovic J. Clinical features and management of tardive dyskinesias, tardive myoclonus, tardive tremor, and tardive tourettism. In: Sethi K, ed. *Drug Induced Movement Disorders*. New York: Marcel Dekker, Inc.; 2004:77–109.
17. Burke RE, Kang UJ, Jankovic J, et al. Tardive akathisia: an analysis of clinical features and response to open therapeutic trials. *Mov Disord*. 1989;4(2):157–175.
18. Stacy M, Cardoso F, Jankovic J. Tardive stereotypy and other movement disorders in tardive dyskinesias. *Neurology*. 1993;43(5):937–941.
19. Tominaga H, Fukuzako H, Izumi K, et al. Tardive myoclonus. *Lancet*. 1987;1(8528):322.
20. Fahn S, Jankovic J. *Principles and Practice of Movement Disorders*. Philadelphia: Churchill Livingstone, Elsevier; 2007.
21. Yagi G, Takamiya M, Kanba S, et al. Mortality rate of schizophrenic patients with tardive dyskinesia during 10 years: a controlled study. *Keio J Med*. 1989;38(1):70–72.
22. Ballesteros J, González-Pinto A, Bulbena A. Tardive dyskinesia associated with higher mortality in psychiatric patients: results of a meta-analysis of seven independent studies. *J Clin Psychopharmacol*. 2000;20(2):188–194.
23. McEvoy J, Gandhi SK, Rizio AA, et al. Effect of tardive dyskinesia on quality of life in patients with bipolar disorder, major depressive disorder, and schizophrenia. *Qual Life Res*. 2019;28(12):3303–3312.
24. Browne S, Roe M, Lane A, et al. Quality of life in schizophrenia: relationship to sociodemographic factors, symptomatology and tardive dyskinesia. *Acta Psychiatr Scand*. 1996;94(2):118–124.
25. Solmi M, Pigato G, Kane JM, et al. Treatment of tardive dyskinesia with VMAT-2 inhibitors: a systematic review and meta-analysis of randomized controlled trials. *Drug Des Devel Ther*. 2018;12:1215–1238.
26. Kane JM, Correll CU, Nierenberg AA, et al; Tardive Dyskinesia Assessment Working Group. Revisiting the abnormal involuntary movement scale: proceedings from the tardive dyskinesia assessment workshop. *J Clin Psychiatry*. 2018;79(3):17cs11959.
27. Abnormal Involuntary Movement Scale. (117-AIMS). In: Guy W, ed. *ECDEU Assessment Manual for Psychopharmacology*. Rockville, MD: National Institute of Mental Health; 1976:534–537.
28. Bhidayasiri R, Kane JM, Frei KP, et al. Expanding phenomenologic heterogeneity of tardive syndromes: Time for an updated assessment tool. *Parkinsonism Relat Disord*. 2020;77:141–145.
29. Jackson R, Brams MN, Citrome L, et al. Assessment of the impact of tardive dyskinesia in clinical practice: consensus panel recommendations. *Neuropsychiatr Dis Treat*. 2021;17:1589–1597.
30. Stacy M, Sajatovic M, Kane JM, et al. Abnormal involuntary movement scale in tardive dyskinesia: minimal clinically important difference. *Mov Disord*. 2019;34(8):1203–1209.
31. Hauser RA, Barkay H, Wilhelm A, et al. Minimal clinically important change in Abnormal Involuntary Movement Scale score in tardive dyskinesia as assessed in pivotal trials of deutetrabenazine. *Parkinsonism Relat Disord*. 2022;97:47–51.
32. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163.
33. Matheson GJ. We need to talk about reliability: making better use of test-retest studies for study design and interpretation. *PeerJ*. 2019;7:e6918.
34. Savitt D, Jankovic J. Tardive syndromes. *J Neurol Sci*. 2018;389:35–42.
35. Barnes TR, Trauer T. Reliability and validity of a tardive dyskinesia videotape rating technique. *Br J Psychiatry*. 1982;140(5):508–515.