

It is illegal to post this copyrighted PDF on any website.

Revisiting the Abnormal Involuntary Movement Scale: Proceedings From the Tardive Dyskinesia Assessment Workshop

John M. Kane, MD^{a,b,*}; Christoph U. Correll, MD^{a,b,c}; Andrew A. Nierenberg, MD^d;
Stanley N. Caroff, MD^e; and Martha Sajatovic, MD^f; on behalf of the Tardive Dyskinesia Assessment Working Group⁹

ABSTRACT

Objective: To provide an historic overview of the Abnormal Involuntary Movement Scale (AIMS) in clinical trials of tardive dyskinesia (TD), with current recommendations for analyzing and interpreting AIMS data.

Participants: Seven psychiatrists and 1 neurologist were selected by the workshop sponsor based on each individual's clinical expertise and research experience.

Evidence: Using PubMed entries from January 1970 to August 2017, participants selected studies that used the AIMS to evaluate TD treatments. The selections were intended to be representative rather than prescriptive or exhaustive, and no specific recommendations for TD treatment are implied.

Consensus Process: The Working Group met in October 2016 to discuss the AIMS as an assessment tool, outline the challenges of translating clinical trial results into everyday clinical practice, and propose different methods for reporting AIMS data in clinically relevant terms. Recommendations for selecting TD studies for review, analyzing and interpreting AIMS data, and synthesizing discussions among the participants were initiated during the onsite workshop and continued remotely throughout development of this report. Disagreements were resolved via group e-mails and teleconferences. Consensus was based on final approval of this report by all workshop participants.

Conclusions: For both research and clinical practice, the AIMS is a valid measure for assessing TD and the effects of treatment, but alternative analyses of AIMS data (eg, effect size, minimal clinically important difference, response analyses, category shifts) may provide broader evidence of clinical effectiveness. No single analysis of AIMS data can be considered the standard of clinical efficacy; multiple analytic approaches are recommended.

J Clin Psychiatry 2018;79(3):17cs11959

To cite: Kane JM, Correll CU, Nierenberg AA, et al. Revisiting the Abnormal Involuntary Movement Scale: proceedings from the Tardive Dyskinesia Assessment Workshop. *J Clin Psychiatry*. 2018;79(3):17cs11959.

To share: <https://doi.org/10.4088/JCP.17cs11959>

© Copyright 2018 Physicians Postgraduate Press, Inc.

^aDepartment of Psychiatry, The Zucker Hillside Hospital, Glen Oaks, New York

^bDepartment of Psychiatry, The Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, New York

^cCenter for Psychiatric Neuroscience, The Feinstein Institute for Medical Research, Manhasset, New York

^dDepartment of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts

^eDepartment of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania

^fDepartment of Psychiatry, University Hospitals Cleveland Medical Center, Cleveland, Ohio

⁹Members of the Tardive Dyskinesia Assessment Working Group are listed at the end of the article.

*Corresponding author: John M. Kane, MD, The Zucker Hillside Hospital—Department of Psychiatry, The Donald and Barbara Zucker School of Medicine at Hofstra/Northwell—Department of Psychiatry, 75-59 263rd St, Glen Oaks, NY 11004 (jkane2@northwell.edu).

Tardive dyskinesia (TD) is a chronic disorder characterized by involuntary stereotyped, choreic, athetoid, and/or dystonic movements in 1 or more areas of the body, including the orofacial region (eg, tongue thrusting, lip smacking and/or pursing, grimacing), extremities (eg, stereotypic piano-playing movements, flexion/extension of the ankles or toes), and torso (eg, choreoathetoid movements, pelvic rocking).^{1,2} This disorder can result from exposure to dopamine receptor blocking agents (DRBAs) such as antipsychotics and drugs used to treat gastrointestinal disorders (eg, metoclopramide).² Given the difficulty in treating TD, prevention, close monitoring, and earliest possible diagnosis are critical in optimizing patient outcomes.

The term *tardive dyskinesia* was coined in 1964 by Faurbye et al³ in an article that described patients who had developed chronic involuntary movements several months after starting antipsychotic treatment. In the following decades, the link between antipsychotics and TD became widely accepted, with attempts to find effective treatments for TD beginning in the early 1970s.⁴ With development of the newer second-generation (atypical) antipsychotics, it was hoped that the risk for medication-induced TD would diminish.^{4,5} However, as shown in several recent studies,^{6–8} TD continues to be a problem in patients who require any type of antipsychotic treatment. In a 2017 meta-analysis of antipsychotic clinical trials conducted by Carbon et al,⁷ mean probable TD prevalences of 30.0% and 20.7% were found in patients exposed to first- and second-generation antipsychotics, respectively. However, in the subgroup of patients with no lifetime exposure to first-generation antipsychotics, the TD prevalence with second-generation antipsychotics was 7.2%. A contributing factor to the ongoing problem of TD may be the expanding use of atypical antipsychotics in psychiatric indications beyond schizophrenia, including bipolar disorder and refractory major depressive disorder.

A report from the 1992 American Psychiatric Association (APA) TD Task Force cited antipsychotic discontinuation as “the most logical ‘treatment’” for managing TD, with dose reduction suggested if discontinuation is unfeasible.⁹ The 2010 APA Practice Guideline for the Treatment of Patients with Schizophrenia recommends switching to an antipsychotic with a lower risk for TD,¹⁰ although no

- The Abnormal Involuntary Movement Scale (AIMS) can be used to measure the severity of abnormal movements in tardive dyskinesia (TD), but diagnosis requires an assessment of medication history and a clinical evaluation of symptoms. Development of standardized guidelines for the screening, diagnosis, and treatment of patients with TD is warranted.
- Many TD clinical trials have used the AIMS as an efficacy outcome, but the studies vary widely in design and conduct. Interpreting trial results and applying them to clinical practice can be challenging. Presenting AIMS data through different types of analyses, such as minimal clinically important difference or response rates, may provide a broader and more clinically relevant perspective on study results.
- Ongoing education and AIMS training may be necessary for improving the diagnosis and treatment of TD in clinical settings.

antipsychotic medication is completely risk-free. However, these approaches may not be viable options for patients who require long-term antipsychotic therapy and are psychiatrically stable on their current treatment regimen. Moreover, there is extremely limited information on whether or how often TD resolves and on how long it takes to do so after discontinuation or dose reduction of DRBAs.¹¹

Two reversible and selective vesicular monoamine transporter 2 inhibitors, valbenazine and deutetrabenazine, are now approved by the US Food and Drug Administration (FDA) for the treatment of TD.^{12,13} A number of other potential treatments have been tried, but as reported by the American Association of Neurology, there was insufficient or limited evidence for many of these drugs.¹⁴ In the historical absence of approved TD medications, some medications, such as tetrabenazine, were used off-label based on promising data from open-label or single-center trials.¹⁵

The need for effective treatment of TD is underscored by the negative impact of this disorder that stigmatizes patients and contributes to social isolation.¹⁶ In some cases, TD can also be physically debilitating and have a serious negative impact on daily functioning and quality of life.² Even “mild” forms of TD can be highly distressing, especially when noticeable abnormal movements lead to negative consequences, such as loss of vocational opportunities or isolation from family and friends.¹⁶ Although possibly confounded by the severity of the underlying psychiatric illness, which can also affect outcomes, the presence of TD in patients with schizophrenia has been associated with increased mortality, poorer treatment outcomes, lower productivity, and reduced quality of life.^{17–19}

Given the ongoing risk of TD in patients requiring antipsychotic medications or other DRBAs and the negative impact of TD on quality of life, the availability of novel treatments and the current resurgence of interest in TD are encouraging and potentially transformative for individuals affected by TD. However, available treatments do not eliminate

the need for careful assessment of involuntary movements and preventative efforts. Monitoring and recognition of TD are critical skills for clinicians prescribing DRBAs. As part of this effort, use of the Abnormal Involuntary Movement Scale (AIMS) in TD studies and the challenges of translating AIMS study results into clinical practice need to be addressed.

METHODS

The Tardive Dyskinesia Assessment Workshop was convened on October 13, 2016, in New York, New York, to discuss the application and interpretation of the AIMS as an assessment tool for TD. Workshop participants (ie, the Working Group) were invited by the sponsor, Neurocrine Biosciences, Inc., based on their clinical expertise and research experience. The Working Group included 7 psychiatrists with interests in psychopharmacology and drug safety (J.M.K. [chair], C.U.C., A.A.N., S.N.C., M.S., J. P. McEvoy, MD; A. J. Cutler, MD) and 1 neurologist specializing in movement disorders (M. A. Stacy, MD).

The contents of this report represent proceedings from the workshop and from subsequent communications, which were conducted via teleconferences, group e-mails, and shared comments on manuscript drafts that were distributed to all workshop participants for feedback. Disagreements were resolved via e-mail or teleconference as needed. Consensus was based on the final approval of this report by all workshop participants. The goals of the report were to (1) provide clinicians with an historical overview of how the AIMS has been used to evaluate TD in clinical studies, (2) outline some of the challenges of translating clinical trial results into clinical practice, (3) discuss various approaches to analyzing and interpreting AIMS data, and (4) provide consensus statements on these areas. The discussion for each of these goals is organized into the 4 main sections below.

THE ABNORMAL INVOLUNTARY MOVEMENT SCALE

Structure and Scoring

The original AIMS, which was developed by the National Institute of Mental Health for research purposes, includes a total of 12 items.²⁰ The first 7 items are used to measure the severity of abnormal movements in the orofacial region (4 items: facial muscles, lips, jaw, tongue), upper extremities (1 item), lower extremities (1 item), and trunk (1 item). Items 8–12 are related to clinician global judgment of severity, patient awareness, incapacitation due to the abnormal movements, and dental status (1 item each). A later version of the AIMS contains 14 items, which includes 2 additional items for edentulousness and the disappearance of abnormal movements during sleep (Table 1).

Directions for scoring are limited in the original AIMS, and a simple rating scale is provided for scoring items 1–7: 0 = none; 1 = minimal, may be extreme normal; 2 = mild; 3 = moderate; 4 = severe. Because of the simplicity of this scale, it is generally agreed that the AIMS can be easily

It is illegal to post this copyrighted PDF on any website.

Table 1. Abnormal Involuntary Movement Scale (AIMS) Items and Scoring

	Original (Guy 1976 ²⁰)	Variations
Total number of items	12 items Items 1–7: severity of abnormal movement by body region (face, lips, jaw, tongue, upper extremities, lower extremities, trunk) Items 8–10: global judgments of overall severity, patient incapacitation due to abnormal movements, and patient awareness of abnormal movements Items 11–12: current dental problems and denture use	14 items Items 1–12: same as original Item 13: presence of edentia Item 14: disappearance of abnormal movements with sleep
Examination procedures	12 examination steps including activation maneuver	Additional instruction for patient to remove shoes and socks ²¹ Supplemental details for each of the 12 examination steps of the original AIMS ²²
Rating scale for items 1–7	Simple scale applied to all 7 items 0 = none 1 = minimal, may be extreme normal 2 = mild 3 = moderate 4 = severe	Additional criteria for assessing quality, frequency, and amplitude of abnormal movements ^{22,23} Additional criteria for scoring abnormal movements in specific body areas, including fingers, tongue, lips, jaw, and toes ^{22,23}
Scoring of items 1–7	Rate according to the highest severity observed Subtract 1 point if abnormal movement occurred only with an activation maneuver	Except for upper extremities, do not subtract 1 point for activation maneuver ²³ Do not subtract 1 point for any activation maneuver ²²
Scoring of item 8 (global severity)	No directions provided	Based on comparison of current patient to other patients ²³ Equals the highest single score from items 1–7 ²²
Total score	No mention of total score	Equals the sum of scores from items 1–7 ²³
Criteria for tardive dyskinesia	No diagnostic criteria provided for clinical or research purposes	Criteria developed for research purposes Schooler-Kane criteria ²⁴ ≥ 3 months of cumulative exposure to a neuroleptic drug Score ≥ 3 (moderate or severe) in ≥ 1 AIMS item or score ≥ 2 (mild or worse) in ≥ 2 AIMS items Absence of other conditions that might produce abnormal involuntary movements Glazer-Morgenstern-Doucette criteria ²⁵ AIMS total score ≥ 3 with score ≥ 2 (mild or worse) in ≥ 1 AIMS item

administered in both research and clinical settings. However, it has also been noted that the lack of detailed instruction and descriptors could be challenging for inexperienced raters and therefore contribute to high interrater variability.^{22,23} To mitigate these potential problems and make the scale more specific to TD, detailed instructions have been developed that include quality (eg, choreic or athetoid), frequency, amplitude, and location of abnormal movements as factors in scoring.^{22,23}

Supplementary instructions for administering the AIMS have been developed and published by several research groups. These instructions include the removal of shoes and socks for examination, not subtracting 1 point for abnormal movements that occur only with activation maneuvers, methods for examining and scoring specific body areas, and methods for scoring item 8 (global judgment of severity) (Table 1). However, one development that is central to understanding clinical trial results is the calculation of a total score. The original AIMS does not include mention of a total score, but it has become a convention to sum the individual scores from items 1–7. Thus, although there is a generally accepted range for the AIMS total score (0 to 28), the range itself is not linear because each constituent item is scored separately. In other words, an AIMS total score of 7 could represent a score of 1 (minimal) on each item or a different combination of item scores—for example, a score of

3 (moderate) on 1 item, score of 4 (severe) on another item, and score of 0 (none) on the remaining items. The AIMS total score may be a useful index for measuring the overall effects of treatment in a TD clinical trial, but, as discussed later, its applicability in clinical practice may be limited.

Instructions and refinements for the global and distress measures (items 8–10) and for dental pathology (items 11–12) have not been standardized. In addition, items 9 and 10 have not correlated with responses on the anatomic measures of severity of TD (items 1–7) and may not be reliable given the lack of awareness and insight reported by some patients.^{26,27} Given the critical importance of the subjective and social impact of TD, further research to develop reliable instruments for this measure is necessary, as indicated in the recommendations at the end of this report. The following discussion refers only to the AIMS items related to anatomic severity (items 1–7).

AIMS in TD Studies

The original AIMS is strictly an instrument for measuring the anatomic distribution and severity of abnormal movements and does not provide criteria for diagnosing TD. The diagnosis of TD continues to be based on the patient's clinical presentation, evaluation to rule out other diagnostic possibilities, and medication history, as summarized in the *Diagnostic and Statistical Manual of Mental Disorders*,

It is illegal to post this copyrighted PDF on any website.

Fifth Edition (*DSM-5*).²⁸ However, standardized diagnostic criteria for TD have been developed for research purposes, with the most notable examples being the Schooler-Kane²⁴ and Glazer-Morgenstern-Doucette²⁵ criteria (Table 1). These criteria, based on a priori thresholds of severity, are primarily used to estimate the prevalence or incidence of TD in general or clinical populations and to qualify and monitor patients entered into a clinical study. As noted by Schooler and Kane, a more definitive clinical diagnosis of TD requires a history of treatment with antipsychotics (or other DRBAs) and persistence of abnormal movements after discontinuing antipsychotic treatment.²⁴ If the AIMS is included as part of an overall clinical diagnosis of TD, its utility in institutional settings may differ from its use in a more general psychiatric population.

In clinical trials, the AIMS is often used as a safety assessment to monitor the emergence of TD in subjects who are receiving an antipsychotic medication for schizophrenia or other psychiatric disorder. The AIMS is also used as an efficacy measure in clinical trials that focus on improvements in TD. These studies generally rely on the AIMS total score (sum of items 1–7) as an overall index of TD severity. In placebo-controlled trials that include statistical testing, results such as a mean change from baseline in the AIMS total score can help clinicians decide whether to try a new treatment in practice.

However, there are differences across clinical trials that need to be considered when interpreting AIMS results. Such differences are summarized in a sample of TD studies that were selected to illustrate a range of study designs, rating methods for the AIMS, and different types of AIMS results (Table 2). This selection was based on results from a PubMed search that included a simple search string (“Abnormal Involuntary Movement Scale” AND “tardive”) and a single set of search dates (from January 1, 1970, to August 31, 2017). The selected studies were intended to be representative rather than prescriptive or exhaustive, and no specific recommendations for TD treatment are implied.

One factor to consider when interpreting AIMS results is the rating method. For example, as in the single-center trial of *Ginkgo biloba* by Zhang et al,³³ each subject may be assessed by the same investigator throughout treatment. Although this investigator may have been blinded to treatment, he or she would have known how long the patient was receiving treatment and could have been more alert to and/or expectant of improvements as the study progressed. In addition, because of different training backgrounds or personal experiences, investigators within a study site might not have applied the same approach for rating the severity of TD movements. Therefore, variability among investigators could be high unless interrater reliability was specifically tested and confirmed. This potential for individual bias and less than optimal interrater variability may be minimized in studies that use central (offsite) video raters. For example, in the multicenter trials of valbenazine and deutetrabenazine,^{29–32} each subject’s AIMS examination was video recorded in a standardized manner at all study

visits. Central raters who viewed these videos were blinded to treatment. Moreover, they did not know which study visit they were viewing since the sequence of the videos was scrambled, which may have minimized the potential for inflated scores at baseline and overly reduced scores at the end of the study. Finally, the valbenazine studies included well-defined anchors for scoring each AIMS item, and all AIMS scoring required a consensus between 2 central raters who watched the videos together. These types of controls are expected to become the new standard for evaluating TD therapies.

APPLYING THE AIMS TO CLINICAL PRACTICE

The AIMS can be used in both research and clinical settings and administered by any health care professional with appropriate training.^{39,40} Similar to its use in clinical trials, the AIMS examination and rating scores can be used in clinical practice to document the emergence of medication-induced TD and monitor changes in TD severity over time. To this end, formal guidelines have been developed that propose administration of the AIMS at regular intervals in patients receiving antipsychotics in clinical settings (eg, every 3–12 months depending on risk factors).⁴¹ However, additional review of these recommendations may be necessary since patients could develop signs of TD that would be missed within these time intervals. A more conservative approach in clinical practice may be for all patients and their caregivers to be informed about regular self-examination and to be questioned and briefly examined for abnormal movements at every clinic visit.

Clinical trials often use mean changes in the AIMS total score to evaluate whether a medication has demonstrable unwanted effects on the incidence of abnormal involuntary movements. However, this can be problematic since newly emergent cases with TD may be obscured in many patients, with no change in the total AIMS score from baseline. In addition, different mathematical approaches (eg, arithmetic mean, geometric mean, median) and analyses (eg, analysis of covariance, mixed-effects model for repeated measures, nonlinear machine learning algorithms) could be used to present score changes. Therefore, categorical, case-based reporting is important and should always accompany mean score reporting. Additionally, the severity of TD cases needs to be reported, which has been absent in most of the recently meta-analyzed prevalence studies of TD in patients receiving antipsychotics.⁷

Similarly, clinical trials often use mean changes in the AIMS total score to evaluate whether a medication has demonstrable efficacy for TD improvement. When interpreting such results, it is important to remember that the total score is not a linear scale, but rather, the sum of 7 individual item scores. Each AIMS item may have face validity for rating the severity of a particular abnormal movement (ie, from 0 = none to 4 = severe), but the total score does not have ideal clinimetric properties for rating overall severity as discussed above. Clinical researchers should be

You are prohibited from making this PDF publicly available.

It is illegal to post this copyrighted PDF on any website.**Table 2. AIMS Outcomes in Tardive Dyskinesia Studies**

Study, Design, and Treatment ^a	Participants	AIMS Scoring and Analyses	AIMS Total Score Outcomes ^b
Hauser et al, 2017 ²⁹ KINECT 3 study Multicenter RDBPC, 6 weeks Valbenazine 40 mg qd (n = 70) Valbenazine 80 mg qd (n = 79) Placebo (n = 76)	TD per <i>DSM</i> criteria Moderate or severe TD per qualitative assessment by a blinded, external reviewer using a video of the subject's AIMS assessment at screening Stable-dose psychiatric medications allowed	Scored by consensus between 2 central AIMS video raters who were blinded to treatment and study visit Cohen <i>d</i> used to estimate ES	Mean (SD) score at BL PBO: 9.9 (4.3) Tx 40 mg: 9.7 (4.1) Tx 80 mg: 10.4 (3.6) LS mean change from BL at wk 6 PBO: -0.1 Tx 40 mg: -1.9; <i>P</i> = .002; ES, <i>d</i> = 0.52 Tx 80 mg: -3.2; <i>P</i> < .001 (statistically significant per testing procedure); ES, <i>d</i> = 0.90 Response (≥ 50% improvement) at wk 6 PBO: 8.7% Tx 40 mg: 23.8%; <i>P</i> = .0200 Tx 80 mg: 40.0%; <i>P</i> < .001
Fernandez et al, 2017 ³⁰ ARM-TD study Multicenter RDBPC, 12 weeks Deutetrabenazine 12–48 mg bid (n = 58) Placebo (n = 59)	Moderate or severe TD with symptoms that are bothersome or cause functional impairment Stable-dose psychiatric medications allowed AIMS total score ≥ 6 (for post hoc subgroup analyses)	Scored by consensus between 2 central video raters who were blinded to treatment and study visit	Mean (SD) score at BL PBO: 9.6 (3.8) Tx: 9.6 (4.1) LS mean change from BL at wk 12 (ITT) PBO: -1.6 Tx: -3.0; <i>P</i> = .019 LS mean change from BL at wk 12 (BL AIMS ≥ 6 subgroup) PBO: -1.9 Tx: -3.4; <i>P</i> = .027
Anderson et al, 2017 ³¹ AIM-TD study Multicenter RDBPC, 12 weeks Deutetrabenazine 12 mg/d (n = 75) Deutetrabenazine 24 mg/d (n = 74) Deutetrabenazine 36 mg/d (n = 75) Placebo (n = 74)	Clinical diagnosis of moderate or severe TD with symptoms that are bothersome or cause functional impairment AIMS total score ≥ 6 Stable-dose psychiatric medications allowed	Scored by consensus between 2 central video raters who were blinded to treatment and study visit; raters reviewed videos in pairs	Mean (SD) score at BL (mITT) PBO: 9.5 (2.7) Tx 12 mg: 9.6 (2.4) Tx 24 mg: 9.4 (2.9) Tx 36 mg: 10.1 (3.2) LS mean change from BL at wk 12 (mITT) PBO: -1.4 Tx 12 mg: -2.1, <i>P</i> = .217 Tx 24 mg: -3.2, <i>P</i> = .003 Tx 36 mg: -3.3, <i>P</i> = .001 Response (≥ 50% improvement) at wk 12 PBO: 12% Tx 12 mg: 13% Tx 24 mg: 35%; <i>P</i> = .005 Tx 36 mg: 33%; <i>P</i> = .007
O'Brien et al, 2015 ³² KINECT 2 study Multicenter RDBPC, 6 weeks Valbenazine 25–75 mg qd (n = 45) Placebo (n = 44)	TD per <i>DSM</i> criteria Moderate or severe TD per qualitative assessment by a blinded, external reviewer using a video of the subject's AIMS assessment at screening Stable-dose psychiatric medications allowed	Scored by consensus between 2 central AIMS video raters who were blinded to treatment and study visit	Mean (SD) score at BL PBO: 7.9 (4.5) Tx: 8.0 (3.5) LS mean change from BL (SEM) at wk 6 PBO: -0.2 (1.1) Tx: -2.6 (1.2); <i>P</i> = .0005 Response (≥ 50% improvement) at wk 6 PBO: 18.2% Tx: 48.9%; <i>P</i> = .002
Zhang et al, 2011 ³³ Single-site RDBPC, 12 weeks <i>Ginkgo biloba</i> 240 mg tid (n = 78) Placebo (n = 79)	TD per Schooler-Kane criteria ²⁴ Stable-dose antipsychotics allowed, including clozapine (n = 128) Anticholinergics also allowed	Scored by trained clinical psychiatrists who were blinded to treatment Each subject assessed by the same psychiatrist Method for estimating ES not reported	Mean (SD) score at BL PBO: 6.9 (3.6) Tx: 7.0 (2.9) Mean (SD) score at wk 12 PBO: 7.0 (3.3) Tx: 4.9 (2.2); <i>P</i> < .0001; ES = 0.77 Response (≥ 30% improvement) at wk 12 PBO: 5.1%; Tx: 51.3%; <i>P</i> < .001
Woods et al, 2008 ³⁴ Single-site RDBPC, 12 weeks OL, 12 weeks Levetiracetam 500–3,000 mg bid (n = 25) Placebo (n = 25)	TD per Glazer-Morgenstern-Doucette criteria ²⁴ Stable-dose antipsychotics and anticholinergics allowed	AIMS scored by 2 raters at the single study site who had established interrater reliability in an earlier study AIMS total score not defined; presumed to be sum of items 1–7	Mean (SD) score at BL PBO: 8.0 (3.1) Tx: 9.4 (3.4) % improvement from BL at wk 12 PBO: 18.7% Tx: 43.5%; <i>P</i> < .05 % improvement from BL at wk 24 Switched from PBO to Tx: 39.1% Received continuous Tx: 57.7%

(continued)

You are prohibited from making this PDF publicly available.

Table 2 (continued).

Study, Design, and Treatment ^a	Participants	AIMS Scoring and Analyses	AIMS Total Score Outcomes ^b
Bai et al, 2003 ³⁵ Single-site RDBPC, 12 weeks Risperidone 2–6 mg (n = 22) Placebo (n = 20)	Severe abnormal involuntary movements in ≥ 1 of 7 body areas (face, lips, jaw, tongue, upper extremities, lower extremities, trunk) for ≥ 3 months 4-week washout of antipsychotics required Concomitant use of anticholinergics required Concomitant use of benzodiazepines allowed but not required	AIMS scored by 3 investigators using the Munetz and Benjamin procedure ²² with an interrater reliability of 0.94 Analyses based on AIMS item scores also reported but not included here	Mean (SD) score at BL PBO: 16.4 (4.3) Tx: 15.4 (5.0) Mean (SD) score at wk 12 PBO: 15.4 (5.7) Tx: 9.9 (4.4); <i>P</i> = .002 Response (≥ 3–4 point improvement) at wk 12 PBO: 30% Tx: 68%; <i>P</i> = .029
Adler et al, 1999 ³⁶ Multicenter RDBPC, up to 2 years <i>d</i> -Vitamin E 1,600 IU/d (n = 73) Placebo (n = 85)	TD per Schooner-Kane criteria ²⁴ with dyskinetic movements for ≥ 3 months and cumulative antipsychotic exposure ≥ 3 months Stable doses of antipsychotics and other psychotropic medications allowed Anticholinergic medications discontinued during treatment	AIMS scored by a research assistant at all study visits and by the investigator at BL and select study visits Analysis based on mean AIMS score (not defined)	Mean (SD) score at BL PBO: 9.8 (3.2) Tx: 10.8 (4.2) Mean (SD) score at wk 52 PBO: 9.8 (3.2) Tx: 9.5 (4.6); <i>P</i> = .36
Ondo et al, 1999 ³⁷ Prospective case series Tetrabenazine 25–150 mg tid (n = 20) Mean treatment duration: 20.3 weeks Mean dose: 57.9 mg/d	TD based on typical clinical appearance that was temporarily related to DRBA use and absence of other potential etiologies 30-day washout of DRBA required	AIMS videos scored by investigator not involved with patient management and blinded to sequence (pre- or post-treatment) Analyses based on AIMS “subjective” score (sum of items 8–10) also reported but not included here	Mean (SD) score at BL Tx: 17.9 (4.4) Mean (SD) score at EoT Tx: 8.2 (5.3); <i>P</i> < .001 vs BL % improvement from BL at EoT Tx: 60.4%
Stewart et al, 1982 ³⁸ Single-site RDBPC, 6 weeks OL, 6 weeks Baclofen 10–30 mg tid (n = 14) Placebo (n = 19)	DRBA-induced TD per referring physician (psychiatrist or neurologist) TD for ≥ 6 months prior to entry Continued stable-dose DRBA required	Methods for scoring not reported AIMS total score not defined; presumed to be sum of items 1–7 Analyses limited to patients who completed the DBPC period (baclofen, n = 13; PBO, n = 17) and OL period (baclofen, n = 16)	Score (SD) at BL PBO: 16.6 (5.5) Tx: 17.4 (6.5) Score (SD) at wk 6 PBO: 12.4 (6.8) Tx: 11.0 (7.4); <i>P</i> < .05 vs BL; NS vs PBO Response (≥ 25% improvement) at wk 6 PBO: 47% Tx: 67% Response at wk 12 Of 9 DBPC PBO nonresponders: 7 became responders Of 7 DBPC Tx responders: 2 became nonresponders

^aN values represent the number of intent-to-treat subjects.

^bAIMS total score defined as the sum of scores from AIMS items 1–7, as defined in the study unless indicated otherwise.

Abbreviations: AIMS = Abnormal Involuntary Movement Scale, bid = twice daily, BL = baseline, DBPC = double-blind placebo-controlled, EoT = end of treatment, DRBA = dopamine receptor blocking agent, *DSM* = *Diagnostic and Statistical Manual of Mental Disorders*, ES = effect size, LS = least squares, NS = nonsignificant, ITT = intent-to-treat, mITT = modified intent-to-treat, OL = open-label, PBO = placebo, qd = once daily, RDBPC = randomized double-blind placebo-controlled, SD = standard deviation, SEM = standard error of the mean, TD = tardive dyskinesia, tid = 3 times daily, Tx = treatment.

encouraged to provide reports that include individual item scores in addition to the AIMS total score. This procedure will facilitate the generalizability of research findings, which clinicians can appreciate and compare with their own office-based assessments.

In randomized, controlled clinical trials, the AIMS may best be scored by blinded central raters who are viewing standardized video recordings or by 2-way live examinations and following protocol-defined rating procedures (eg, descriptive anchors for each severity level). In contrast, AIMS scoring in clinical settings usually involves face-to-face interactions between the clinician and patient, both of whom are aware of what treatment has been prescribed and how long the patient has been treated. In a community mental health care setting or group practice, several clinicians may be responsible for the same patient, and differences in their

individual experiences could affect how TD is assessed. For example, a clinician with limited AIMS training who has seen only a few cases of TD may have a different concept of what constitutes a “severe” abnormal movement than an AIMS-trained clinician who has seen hundreds of cases over many years. At this time, no single approach can be recommended for improving interrater reliability within clinical settings; each institution or practice needs to develop their own protocols for screening, diagnosis, and monitoring. However, renewed education and training in use of the AIMS are necessary to ensure reliability of ratings. Such methods could include instructional videos, in-house training by an experienced clinician, and/or participation in continuing medical education activities. Moreover, as shown in the work by Lane et al,²³ interrater reliability may be improved by implementing specific scoring criteria for AIMS items 1–7.

It is illegal to post this copyrighted PDF on any website.

As an instrument that measures the frequency, amplitude, distribution, and/or persistence of abnormal movements, the AIMS can be administered to any patient regardless of psychiatric diagnosis. From a clinical and patient or caregiver perspective, however, the patient's diagnosis (eg, schizophrenia or mood disorder), level of functioning, and psychiatric stability may be important factors in determining overall functional significance of TD in terms of impact on quality of life. For example, an individual with stable bipolar disorder and high psychosocial functioning could have a rating of 2 (mild) in a single AIMS item, such as the tongue. In a clinical trial in which efficacy is being evaluated and averaged within a group of participants (rather than in an individual patient), this rating would equal a total score of 2 and might be considered a "low" overall score and would not even meet research diagnostic criteria (eg, Schooler-Kane) for inclusion in the study. In a clinical setting, however, the same individual may complain of having a fairly disabling tongue dyskinesia with considerable disruption to social and work activities, and the practicing physician would have to make the diagnosis of TD and could consider this patient as having a significant or even serious case of TD. Another individual with unstable schizophrenia and minimal social interaction could have the same rating (ie, score of 2 on the AIMS tongue item only), but in this case, the TD may not be considered functionally significant by the physician because the movements may be overshadowed by other urgent psychosocial needs.

In addition to the natural variability of TD, which can fluctuate during the day and over time,⁴² TD varies widely from patient to patient, in terms of both clinical presentation and psychosocial impact. Consequently, applying clinical study results to this heterogeneous population can be challenging. Establishing a statistically significant change in the mean AIMS total score in a clinical trial is a crucial initial step for demonstrating the efficacy of a TD treatment. However, as discussed in the following section, additional types of AIMS analyses are both possible and necessary for ascertaining the potential benefits of various TD medications.

TYPES OF AIMS ANALYSES

The Working Group discussed the different methods that could be used to analyze AIMS total and item scores in a clinically meaningful way. Results of the discussion are presented below, along with a general caveat that the clinical relevance of any specific analysis may be driven by what the individual clinician wants to know and the specific therapeutic needs of the patient and the patient's caregiver. From an epistemic standpoint, it should be noted that application of these analyses to certain types of clinical trial data may be inappropriate. For example, "clinical relevance" is not germane to a proof-of-concept trial that was only designed to detect a possible drug effect. Therefore, application of the analyses described below may need to be limited to data from larger and well-controlled studies that were specifically designed to establish efficacy.

Treatment Effect Sizes

Treatment effect sizes provide a way to standardize mean score changes by incorporating placebo effects, sample sizes, and standard deviations. Such standardization allows the results of 1 assessment (eg, AIMS) to be quantitatively compared with the results of another assessment (eg, Unified Dyskinesia Rating Scale). For the AIMS, treatment effect sizes could be estimated for the total score and/or individual item scores, with the interpretation of effect sizes initially following general conventions. For example, a Cohen $d=0.5$ may indicate a moderate or medium treatment effect. However, as Jacob Cohen himself cautioned, "the terms 'small,' 'medium,' and 'large' are relative, not only to each other, but to the area of behavioral science or even more particularly to the specific content and research method being employed in any given investigation."⁴³ Therefore, more research in the field is needed to better understand what constitutes a clinically meaningful treatment effect size for TD beyond a mathematical or statistical metric.

Minimal Clinically Important Difference

The minimal clinically important difference (MCID) is the mean score change for an assessment of interest (eg, AIMS) in subjects who experienced a defined level of clinical benefit. An MCID is unlikely to influence clinical decisions unless it represents a minimal level of acceptable improvement. To that end, MCIDs are often based on a clinician- or patient-rated anchor scale (eg, the Clinical Global Impression of Change-Tardive Dyskinesia [CGI-TD] or Patient Global Impression of Change [PGIC]), the standard deviation or standard error of the mean for the assessment of interest, and/or by expert consensus (eg, Delphi method).

The Working Group agreed that no MCID for the AIMS has been established in TD. As a test case, data were pooled from placebo-controlled trials of valbenazine and analyzed using a minimal global response (CGI-TD rating of "minimally improved" or better, score ≤ 3) and a more robust response (CGI-TD rating of "much improved" or better, score ≤ 2) as anchors. Among subjects who met either of these CGI-TD criteria (regardless of treatment), the mean changes from baseline in AIMS total score were -2.2 and -3.4 , respectively. These results suggest that in adults with TD, the MCID for AIMS total score may be 2 or 3 points.⁴⁴ A similar approach could be taken in which CGI-TD responses are correlated with the percent change from baseline in AIMS total score. Further MCID analyses from current TD trials are warranted, and a manuscript from the Working Group is currently in development.

Response Analyses

Response analyses are valuable for identifying the percentage of individual subjects in a clinical trial who achieved a specific threshold of improvement, although the utility of a threshold depends on its application. For example, a low threshold ($\geq 10\%$ or $\geq 20\%$ improvement) may be sufficient in a proof-of-concept study to establish possible

treatment effect, but it may be insufficient for establishing a clinically meaningful response or making treatment decisions.

TD studies have historically defined AIMS response as a $\geq 30\%$ decrease (improvement) from baseline in total score,^{33,45,46} but a more rigorous definition of response ($\geq 50\%$ decrease from baseline) has been used in recent clinical trials.^{29,32} Both benchmarks may be meaningful to clinicians—one because of its historical context and the other because of its stringency. Correlating the percent change in AIMS total score with global anchors of response (eg, CGI-TD response, as described above) would provide additional information about which levels of AIMS response are most clinically meaningful. Because individual patients may have different treatment goals, and percent improvement from baseline is highly influenced by where patients start out, presenting a full range of response criteria (eg, $\geq 10\%$ to $\geq 90\%$) could also be informative for clinicians. Since patients and caregivers often inquire about the “odds of getting better,” clinical trial reports could also include odds ratios and numbers needed to treat (NNTs) for response analyses.

Percent Change From Baseline

The percent change from baseline in the AIMS total score (or in individual AIMS item scores) can be an additionally informative way to present the magnitude of improvement. In contrast to response analyses, which do not include subjects who failed to meet a specific threshold (eg, a subject with 49% improvement cannot be counted in the $\geq 50\%$ response group), percent change captures the experiences of all subjects in a clinical trial and is an intuitively understandable analysis for clinicians, patients, and caregivers. However, as previously mentioned, this type of analysis is highly influenced by baseline severity.

Complete Response or Symptomatic Remission

Clinicians and patients may be particularly interested in the likelihood of substantially reducing, or even eradicating, the signs and symptoms of TD. Response analyses, whether defined as a $\geq 30\%$, $\geq 50\%$, or other reduction in AIMS total score, can show how many subjects in a clinical trial met an overall threshold of TD improvement, but they cannot provide adequate information about symptom resolution, or where patients “end up.” For example, a subject with an AIMS total score of 12 (eg, score = 3 in 4 different items) could have a 50% reduction in total score after treatment, but the subject could still be experiencing moderate symptoms in some areas of the body (eg, score = 3 in 2 items and score = 0 in 2 items). In contrast, a complete response analysis—possibly defined as a score of 0 or 1 on all 7 movement-related items of the AIMS—could show how many subjects had no symptoms or minimal severity in each region of the body after treatment. As with the response analyses, odds ratios and NNTs for complete response would help to make results more accessible to clinicians and patients. Complete response may not be an appropriate analysis for all clinical

trials or applicable to all types of patients, but it could provide a useful data point for assessing treatment options and for informing the clinical decision-making process. It should be noted, however, that a “complete response” measured by AIMS ratings refers to a diminution in the objective severity of observable abnormal movements, but the AIMS *alone* cannot distinguish complete suppression or masking of symptoms from true reversal or remission of TD itself. Such recovery may require that the patient no longer meets the clinical criteria for TD.

Category Shifts

In contrast to percent improvements, which do not take baseline scores into account, category shifts incorporate baseline severity as part of the analysis. As such, shift analyses may be particularly useful when addressing a heterogeneous disorder such as TD. A shift could be defined as a 2-point reduction from baseline in any AIMS item in which the baseline score was ≥ 2 , as was done in a recent study of deutetrabenazine.⁴⁷ Another approach would be to analyze categorical shifts based on an “average” item score, calculated as the AIMS total score at baseline divided by the number of AIMS items that have a score ≥ 1 . The shift could then be defined as an average item score of ≥ 3 at baseline and a score of ≤ 2 after treatment. A final category shift analysis could target syndromal remission (ie, a shift below the syndromal definition of TD), which would translate to no more than a single score of 2 according to Schooler-Kane criteria.²⁴

Functional Remission and Recovery

All of the approaches described above are based on a symptomatic assessment of TD using the AIMS. However, as stated in previous sections of this report, patients do experience varying levels of impairment in subjective well-being, quality of life, and functionality due to TD. Scales that measure such impairment are currently missing but urgently needed, and to achieve complete recovery from TD, such measures are needed to complement AIMS score assessments. Moreover, results on functional measures should meet a minimum threshold of improvement before concluding that a patient has achieved both symptomatic and functional remission from TD and before recovery can be diagnosed. The Working Group recommends a minimum duration of 3 months for sustained symptomatic and functional remission.

SUMMARY AND FUTURE DIRECTIONS

The Working Group agreed that no single analysis can be considered the most clinically relevant method for analyzing AIMS data; therefore, it is important that different types of analyses be conducted and presented, with clear statements as to what was conducted a priori versus post hoc. This multifaceted analytic approach would help achieve the following: (1) confirm the robustness of data from a clinical trial, (2) confirm efficacy across patient subgroups within a clinical trial, and (3) demonstrate that AIMS results

It is illegal to post this copyrighted PDF on any website.

from different or replicate clinical trials are consistent and reinforcing. More discussion is needed between researchers and clinicians to ascertain if there are mutually acceptable standards of clinical relevance that could be used to interpret AIMS results across different clinical trials. With the recent publication of large and well-controlled clinical trials that showed substantial and statistically significant AIMS improvements with valbenazine and deutetrabenazine,^{29–32} one line of inquiry would be to assess whether research diagnostic criteria such as Schooler-Kane²⁴ are adequately sensitive. In addition, TD researchers and clinicians may need to start thinking about optimal ways to adopt new technologies, such as wearable sensors, computerized video-based ratings, and smartphone-based monitoring of abnormal movements and other symptoms.^{48–51}

The Working Group also agreed that more research is needed to better understand the relationship between the AIMS and other assessment tools, including patient-reported measures, caregiver/informant measures, functional scales, and quality of life questionnaires. Future research should also include the use of a broader array of descriptive statistics from the AIMS in epidemiologic studies of TD, particularly in at-risk populations (eg, elderly, women) and in different populations of patients requiring antipsychotic treatment (eg, schizophrenia, schizoaffective disorder, bipolar disorder, refractory major depressive disorder).

CONSENSUS STATEMENTS

1. Diagnosis of TD is based on medication history and presentation of symptoms. The AIMS alone is not a diagnostic tool, but it can be included in a comprehensive assessment of the patient.
2. The AIMS can be used to assess or monitor TD in clinical trials and clinical practice, including the emergence of treatment-related abnormal movements. Clinical trials should implement a method that ensures interrater reliability (eg, consensus of blinded central video raters). Clinical practices are encouraged to implement a standard protocol for TD screening, diagnosis, and assessment that includes a minimal requirement for AIMS training (eg, completion of an instructional video).
3. The AIMS can be administered by any trained health care provider to any patient regardless of psychiatric diagnosis.
4. Although AIMS scores provide reliable information about the severity of abnormal movements, both in research and in practice, they may not be sufficient for indicating the clinical severity and functional impact of TD. Factors such as the patient's psychiatric status, insight into mental and physical illness, subjective well-being, quality of life, and social or occupational burden need to be considered as part of a comprehensive clinical assessment. Such evaluation may require the use of additional scales. However, patient awareness and impact of TD

should not alter the motor examination for AIMS items 1–7.

5. The AIMS total score is not a linear scale, and caution should be taken in drawing any conclusions about overall TD severity based on total score alone.
6. Many studies have used the AIMS to evaluate the effects of treatment on TD. However, TD studies vary widely in design and conduct, and any interpretation of results would benefit from understanding how the AIMS was administered and scored.
7. In contemporary TD clinical trials, efficacy is generally based on a mean change from baseline in the AIMS total score, with significance testing versus placebo. Other types of AIMS analyses are possible and warranted to broaden the scope of clinically meaningful study results, although no single type of analysis can be considered a definitive measure of clinical significance. It is recommended that clinical trials provide different types of analyses so that each clinician can find the types of results that are most applicable to his or her practice (eg, treatment effect size, percent score improvement, MCID, response analyses, shift analyses).
8. Professional or regulatory agencies should consider convening a task force to develop standardized guidelines for measuring and reporting efficacy in clinical trials across a broad array of analyses using the AIMS, including new instruments for measuring insight, awareness, and the psychosocial and vocational stigma and impact of TD.
9. Re-education programs targeting psychiatrists in clinical practice should be developed and disseminated to enhance and standardize examination, diagnosis, and treatment guidelines based on the AIMS for patients with TD. Ongoing efforts are being made within the research community to increase awareness of TD and establish standards that can guide screening/diagnosis and treatment. Psychiatric clinics and offices are encouraged to begin identifying training requirements that are appropriate and easy to implement.

Submitted: October 4, 2017; accepted March 16, 2018.

Published online: May 8, 2018.

Tardive Dyskinesia Assessment Working Group: Stanley N. Caroff (Perelman School of Medicine, University of Pennsylvania), Christoph U. Correll (Zucker Hillside Hospital; Hofstra Northwell School of Medicine; Feinstein Institute for Medical Research), Andrew J. Cutler (Meridien Research), John M. Kane (Zucker Hillside Hospital; Hofstra Northwell School of Medicine), Joseph P. McEvoy (Medical College of Georgia, Augusta University), Andrew A. Nierenberg (Massachusetts General Hospital), Martha Sajatovic (University Hospitals Cleveland Medical Center), and Mark Stacy (Brody School of Medicine, East Carolina University).

Potential conflicts of interest: Dr Kane has been a consultant and/or advisor to and/or has received honoraria from Alkermes, Allergan, Bristol-Myers Squibb, IntraCellular Therapies, Janssen, Lundbeck, Minerva, Neurocrine, Otsuka, Pierre Fabre, Reviva, Sunovion, Takeda, and Teva and is a shareholder of LB Pharma, MedAvante, and The Vanguard Research Group. Dr Correll has been a consultant and/or advisor to or has received honoraria from Alkermes,

Allergan, Bristol-Myers Squibb, Gerson Lehrman Group, IntraCellular Therapies, Janssen/J&J, LB Pharma, Lundbeck, Medavante, Medscape, Neurocrine, Otsuka, Pfizer, Sunovion, Takeda, and Teva; provided expert testimony for Bristol-Myers Squibb, Janssen, and Otsuka; served on a Data Safety Monitoring Board for Lundbeck and Pfizer; and received grant support from Takeda. **Dr Nierenberg** has received research grants from Takeda/Lundbeck, PamLab, GlaxoSmithKline, NeuroRX Pharma, Marriott Foundation, National Institutes of Health, Brain & Behavior Research Foundation, Janssen, and the Patient Centered Outcomes Research Institute and has been a consultant and/or attended advisory board meetings for Takeda/Lundbeck, PamLab, Alkermes, PAREXEL, Sunovion, Naurex, Neurocrine, Hoffmann-La Roche/Genentech, Eli Lilly, Pfizer, SLACK, and Physicians Postgraduate Press. **Dr Caroff** has been a consultant and/or attended advisory board meetings for Teva and Neurocrine Biosciences and received grant support from Neurocrine Biosciences. **Dr Sajatovic** has received research grants from Merck, Alkermes, Janssen, Reuter Foundation, Woodruff Foundation, Reinberger Foundation, National Institutes of Health, and Centers for Disease Control and Prevention; been a consultant to Bracket, Prophase, Otsuka, Sunovion, Neurocrine, Supernus, and Health Analytics; received royalties from Springer Press, Johns Hopkins University Press, Oxford Press, and UpToDate; and been compensated for CME activities by the American Physician Institute, MCM Education, CMEology, and Potomac Center for Medical Education.

Funding/support: The Tardive Dyskinesia Assessment Workshop was supported by Neurocrine Biosciences, Inc., San Diego, California. All workshop members (authors and nonauthors) received an honorarium for onsite participation. However, no remuneration was provided to participants for their contributions as an author or reviewer of this publication.

Role of sponsor: The sponsor was involved in planning and organizing the Tardive Dyskinesia Assessment Workshop. The authors are responsible for the content, final approval, and decision to submit this manuscript for publication.

Acknowledgments: Clinical trial data for valbenazine were provided by Neurocrine Biosciences, Inc., San Diego, California. Medical reviews of the manuscript were provided by Christopher O'Brien, MD; Grace Liang, MD; Scott Siegert, PharmD; and Khodayar Farahmand, PharmD, all of whom are full-time employees at Neurocrine Biosciences, Inc. Writing and editorial assistance was provided by Mildred Bahn at Prescott Medical Communications Group, Chicago, Illinois, with support from Neurocrine Biosciences, Inc.

REFERENCES

- Rana AQ, Chaudry ZM, Blanchet PJ. New and emerging treatments for symptomatic tardive dyskinesia. *Drug Des Devel Ther*. 2013;7:1329–1340.
- Waln O, Jankovic J. An update on tardive dyskinesia: from phenomenology to treatment. *Tremor Other Hyperkinet Mov (N Y)*. 2013;3.
- Faurbye A, Rasch PJ, Petersen PB, et al. Neurological symptoms in pharmacotherapy of psychoses. *Acta Psychiatr Scand*. 1964;40:10–27.
- Cloud LJ, Zutshi D, Factor SA. Tardive dyskinesia: therapeutic options for an increasingly common disorder. *Neurotherapeutics*. 2014;11(1):166–176.
- Tarsy D, Lungu C, Baldessarini RJ. Epidemiology of tardive dyskinesia before and during the era of modern antipsychotic drugs. *Handb Clin Neurol*. 2011;100:601–616.
- 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.
- 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.
- Bakker PR, de Groot IW, van Os J, et al. Long-stay psychiatric patients: a prospective study revealing persistent antipsychotic-induced movement disorder. *PLoS One*. 2011;6(10):e25588.
- American Psychiatric Association. *Tardive Dyskinesia: A Task Force Report of the American Psychiatric Association*. 1st ed. Washington, DC: American Psychiatric Association; 1992.
- Lehman AF, Lieberman JA, Dixon LB, et al. *Practice Guideline for the Treatment of Patients with Schizophrenia*. 2nd ed. Washington, DC: American Psychiatric Association; 2010.
- Zutshi D, Cloud LJ, Factor SA. Tardive syndromes are rarely reversible after discontinuing dopamine receptor blocking agents: experience from a university-based movement disorder clinic. *Tremor Other Hyperkinet Mov (N Y)*. 2014;4:266.
- Ingrezza [package insert]. San Diego, CA: Neurocrine Biosciences, Inc; 2017.
- Austedo [package insert]. North Wales, PA: Teva Pharmaceuticals USA, Inc; 2017.
- Bhidayasiri R, Fahn S, Weiner WJ, et al. Evidence-based guideline: treatment of tardive syndromes: report of the Guideline Development Subcommittee of the American Academy of Neurology. *Neurology*. 2013;81(5):463–469.
- Chen JJ, Ondo WG, Dashtipour K, et al. Tetrabenazine for the treatment of hyperkinetic movement disorders: a review of the literature. *Clin Ther*. 2012;34(7):1487–1504.
- Yassa R, Jones BD. Complications of tardive dyskinesia: a review. *Psychosomatics*. 1985;26(4):305–313.
- 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.
- Ascher-Svanum H, Zhu B, Faries D, et al. Tardive dyskinesia and the 3-year course of schizophrenia: results from a large, prospective, naturalistic study. *J Clin Psychiatry*. 2008;69(10):1580–1588.
- Chong SA, Tay JA, Subramaniam M, et al. Mortality rates among patients with schizophrenia and tardive dyskinesia. *J Clin Psychopharmacol*. 2009;29(1):5–8.
- 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.
- Missouri Department of Mental Health. Health and Safety Forms, Abnormal Involuntary Movement Scale (AIMS). MDMH website. <https://dmh.mo.gov/docs/dd/forms/healthsafety/aims.doc>.
- Munetz MR, Benjamin S. How to examine patients using the Abnormal Involuntary Movement Scale. *Hosp Community Psychiatry*. 1988;39(11):1172–1177.
- Lane RD, Glazer WM, Hansen TE, et al. Assessment of tardive dyskinesia using the Abnormal Involuntary Movement Scale. *J Nerv Ment Dis*. 1985;173(6):353–357.
- Schooler NR, Kane JM. Research diagnoses for tardive dyskinesia. *Arch Gen Psychiatry*. 1982;39(4):486–487.
- Glazer WM, Morgenstern H, Doucette JT. Predicting the long-term risk of tardive dyskinesia in outpatients maintained on neuroleptic medications. *J Clin Psychiatry*. 1993;54(4):133–139.
- Macpherson R, Collis R. Tardive dyskinesia: patients' lack of awareness of movement disorder. *Br J Psychiatry*. 1992;160(01):110–112.
- Arango C, Adami H, Sherr JD, et al. Relationship of awareness of dyskinesia in schizophrenia to insight into mental illness. *Am J Psychiatry*. 1999;156(7):1097–1099.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Fifth Edition. Arlington, VA: American Psychiatric Association; 2013.
- Hauser RA, Factor SA, Marder SR, et al. KINECT 3: a phase 3 randomized, double-blind, placebo-controlled trial of valbenazine (NBI-98854) for tardive dyskinesia. *Am J Psychiatry*. 2017;174(5):476–484.
- Fernandez HH, Factor SA, Hauser RA, et al. Randomized controlled trial of deutetrabenazine for tardive dyskinesia: the ARM-TD study. *Neurology*. 2017;88(21):2003–2010.
- Anderson KE, Stamler D, Davis MD, et al. Deutetrabenazine for treatment of involuntary movements in patients with tardive dyskinesia (AIM-TD): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet Psychiatry*. 2017;4(8):595–604.
- O'Brien CF, Jimenez R, Hauser RA, et al. NBI-98854, a selective monoamine transport inhibitor for the treatment of tardive dyskinesia: A randomized, double-blind, placebo-controlled study. *Mov Disord*. 2015;30(12):1681–1687.
- Zhang WF, Tan YL, Zhang XY, et al. Extract of ginkgo biloba treatment for tardive dyskinesia in schizophrenia: a randomized, double-blind, placebo-controlled trial. *J Clin Psychiatry*. 2011;72(5):615–621.
- Woods SW, Saksa JR, Baker CB, et al. Effects of levetiracetam on tardive dyskinesia: a randomized, double-blind, placebo-controlled study. *J Clin Psychiatry*. 2008;69(4):546–554.
- Bai YM, Yu SC, Lin CC. Risperidone for severe tardive dyskinesia: a 12-week randomized, double-blind, placebo-controlled study. *J Clin Psychiatry*. 2003;64(11):1342–1348.
- Adler LA, Rotrosen J, Edson R, et al. Vitamin E treatment for tardive dyskinesia: Veterans Affairs Cooperative Study #394 Study Group. *Arch Gen Psychiatry*. 1999;56(9):836–841.
- Ondo WG, Hanna PA, Jankovic J. Tetrabenazine treatment for tardive dyskinesia: assessment by randomized videotape protocol. *Am J Psychiatry*. 1999;156(8):1279–1281.
- Stewart RM, Rollins J, Beckham B, et al. Baclofen in tardive dyskinesia patients maintained on neuroleptics. *Clin Neuropharmacol*. 1982;5(4):365–373.
- Menzies V, Farrell SP. Schizophrenia, tardive dyskinesia, and the Abnormal Involuntary Movement Scale (AIMS). *J Am Psychiatr Nurses Assoc*. 2002;8(2):51–56.
- Bark N, Florida D, Gera N, et al. Evaluation of the routine clinical use of the Brief Psychiatric Rating Scale (BPRS) and the Abnormal Involuntary Movement Scale (AIMS). *J Psychiatr Pract*. 2011;17(4):300–303.
- Marder SR, Essock SM, Miller AL, et al. Physical health monitoring of patients with schizophrenia. *Am J Psychiatry*. 2004;161(8):1334–1349.
- Whall AL, Booth DE, Kosinski J, et al. Tardive dyskinesic movements over time. *Appl Nurs Res*.

It is illegal to post this copyrighted PDF on any website.

- 1989;2(3):128–134.
43. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: L. Erlbaum Associates; 1988.
 44. Stacy M, Kurlan R, Burke J, et al. An MCID for AIMS dyskinesia total score change in subjects with tardive dyskinesia (abstract #405). *Mov Disord*. 2017;32(suppl 2):S158.
 45. Hayashi T, Yokota N, Takahashi T, et al. Benefits of trazodone and mianserin for patients with late-life chronic schizophrenia and tardive dyskinesia: an add-on, double-blind, placebo-controlled study. *Int Clin Psychopharmacol*. 1997;12(4):199–205.
 46. Shamir E, Barak Y, Shalman I, et al. Melatonin treatment for tardive dyskinesia: a double-blind, placebo-controlled, crossover study. *Arch Gen Psychiatry*. 2001;58(11):1049–1052.
 47. Fernandez HH, Factor SA, Jimenez-Shahed J, et al. Deutetrabenazine treatment effect in each component of the total Abnormal Involuntary Movement Scale (AIMS) (abstract #916). *Mov Disord*. 2016;31(suppl 2):S294.
 48. Espay AJ, Bonato P, Nahab FB, et al. Technology in Parkinson's disease: challenges and opportunities. *Mov Disord*. 2016;31(9):1272–1282.
 49. LeMoyné R, Tomycz N, Mastroianni T, McCandless C, Cozza M, Peduto D. Implementation of a smartphone wireless accelerometer platform for establishing deep brain stimulation treatment efficacy of essential tremor with machine learning. *Conf Proc IEEE Eng Med Biol Soc*. 2015;2015:6772–6775.
 50. Peterson DA, Littlewort GC, Bartlett MS, et al. Objective, computerized video-based rating of blepharospasm severity. *Neurology*. 2016;87(20):2146–2153.
 51. Roy SH, Cole BT, Gilmore LD, et al. High-resolution tracking of motor disorders in Parkinson's disease during unconstrained activity. *Mov Disord*. 2013;28(8):1080–1087.

You are prohibited from making this PDF publicly available.