

Measuring What Matters:

Further Validation for the Tardive Dyskinesia Impact Scale, a Novel Patient-Reported Outcome Measure in Valbenazine Clinical Trials

Morgan Bron, PharmD, MS; Susan D. Mathias, MPH; Donald E. Stull, PhD; Hui Zhang, PhD; Michelle Turner, MS; Eduardo Dunayevich, MD; Ashok Parameswaran, MD; Dawn Vanderhoef, PhD, DNP, PMHNP-BC, FAANP; M. Mercedes Perez-Rodriguez, MD, PhD; and Christoph U. Correll, MD

Abstract

Importance: Tardive dyskinesia (TD) is a persistent, potentially disabling, medication-induced movement disorder that has been underrecognized. Involuntary movements in TD have a substantial impact beyond movement on individuals with TD. To quantify TD impact and burden, the Tardive Dyskinesia Impact Scale (TDIS), a new, TD-specific, fit-for-purpose patient-reported outcome (PRO) measure, was developed. Objectives were to examine how TDIS contributes to understanding of TD burden and use of clinician-reported outcomes (ClinROs) and other PROs in clinical trials assessing effects of vesicular monoamine transporter 2 inhibitors on TD. TDIS analyses included

assessment of correlations between TDIS and other clinical outcome assessments (PROs and ClinROs), estimation of the minimal clinically importance difference (MCID), and description of the change in TDIS individual items longitudinally and via item response theory.

Observations: In KINECT trials, TDIS followed a similar trajectory to Abnormal Involuntary Movement Scale. An MCID of 4 points in TDIS was considered clinically meaningful. Item-level analyses showed that TDIS is reliable and precise for individual items. Most improved items in longitudinal analyses were self-consciousness (mean change: -1.24), embarrassment (-1.19), unwanted attention (-1.00), and mouth noises

(-1.05), which exceeded the empirically derived item-level threshold for meaningful change (≥ 0.8). TDIS showed moderate correlation with treatment response as measured by Patient's Global Impression of Change ($r = 0.30$) and Clinician's Global Impression of Change ($r = 0.34$) scores.

Conclusions and Relevance: TDIS is the only disease-specific PRO that has been validated in individuals with TD and complements ClinROs and other PROs by providing a comprehensive picture of TD impact beyond movement symptoms and can measure potential benefit of TD treatments.

J Clin Psychiatry 2026;87(2):25nr16047

Author affiliations are listed at the end of this article.

Tardive dyskinesia (TD) is an involuntary, hyperkinetic, persistent, and potentially disabling movement disorder in which individuals experience abnormal involuntary movements. TD is associated with exposure to dopamine receptor-blocking agents (DRBAs), including antipsychotics and gastrointestinal agents such as metoclopramide.^{1,2} TD can manifest during treatment or within weeks after discontinuation or dose reduction of the offending agent.^{3,4} Risk factors for TD include older age, female sex, non-Asian race, comorbid diabetes, substance use disorders, longer and intermittent treatment with and higher doses of DRBAs, treatment-emergent dystonia, Parkinsonism, and akathisia.⁵

The burden of TD has been underrecognized, as TD was often observed in individuals with long-term schizophrenia or those receiving first-generation antipsychotics; the notion that TD only occurs in individuals receiving chronic antipsychotic treatment was based on outdated case reports from the 1950s–1960s from mental health institutions.^{1,6–8} It

was assumed that these individuals were often not aware of their involuntary movements.⁶ Recent studies have shown that many individuals are aware of their symptoms, underscoring the need for diagnosis and treatment.⁹

Before the first US Food and Drug Administration (FDA) drug approved specifically for TD in 2017, only off-label treatments were available, and they were notoriously ineffective. Additionally, the advent of second-generation antipsychotics was expected to reduce TD incidence.¹ However, TD has been observed with use of first- and second-generation antipsychotics, even at low doses.^{9–12} Anyone receiving a DRBA for ≥ 3 months in adults or ≥ 1 month in older adults (≥ 60 years) is at risk for TD.¹³

Importantly, the expanded use of antipsychotics in bipolar disorder and major depressive disorder (MDD), and other off-label indications, may increase TD prevalence and a need to better recognize and understand the impact of TD on health-related quality of life (HRQOL).^{10,14} The predicted global prevalence of TD

Clinical Points

- The validated, tardive dyskinesia–specific Tardive Dyskinesia Impact Scale (TDIS) was recently developed to help patients quantify symptom burden and provide a benchmark for treatment effect. However, how it relates to other clinician- and patient-reported outcomes is not well documented. Additionally, a minimal clinically important difference has not been estimated for TDIS.
- Results of the analyses showed that TDIS followed a similar trajectory to other clinician- and patient-reported outcomes that have been measured in tardive dyskinesia clinical trials.
- A change of 4 points in the TDIS total score is considered clinically meaningful. TDIS is meant to be used alongside other assessments to provide a comprehensive picture of the impact of tardive dyskinesia.

was based on people with schizophrenia and related disorders only.

Based on the current understanding of TD as a distinct condition from other movement disorders and psychiatric conditions, treatment for TD is recommended by both the American Psychiatric Association (APA) and the American Academy of Neurology (AAN).^{15–17} With evidence demonstrating clinically meaningful improvements in TD symptoms regardless of underlying psychiatric diagnosis or patient characteristics, the APA guidelines and recommended updates to the AAN guideline indicate vesicular monoamine transporter 2 (VMAT2) inhibitors as first-line treatment for TD.^{15–18} Valbenazine and deutetrabenazine are reversible VMAT2 inhibitors and are the only medications approved by the FDA for managing TD.^{19,20}

Improvements beyond movement may have an impact on an individual's functional and emotional well-being. As such, the Tardive Dyskinesia Impact Scale (TDIS), a novel, TD-specific, fit-for-purpose, patient-reported outcome (PRO), was developed with patient and caregiver input.²¹ The psychometric validation of the TDIS and its factor structure were previously described in the Farber et al 2024 publication.²¹ Analyses evaluating the measurement properties of the TDIS demonstrated strong reliability (Cronbach α = 0.88–0.90; intraclass correlation coefficient = 0.83) and evidence of construct validity, including limited overlap with clinician-rated movement severity (Abnormal Involuntary Movement Scale [AIMS]) and known-groups differences consistent with both patient- and clinician-reported global impressions items and with symptom bother. The final TDIS consists of 11 questions in 2 domains, with 8 items assessing physical impairment and 3 items assessing socioemotional distress over the previous 7 days.²¹ Each question is scored from 0 (no impact) to 4 (most impact). Total scores range from 0 to 44, with higher scores indicating greater impairment and disability.²¹ TDIS is

a well-supported measure for assessing the impact of TD symptoms on functioning and well-being of adults with moderate to severe TD lasting ≥ 3 months and a diagnosis of schizophrenia, schizoaffective disorder, or mood disorder.

The primary objective of this review was to examine how TDIS uniquely contributes to the understanding of the impact of TD and how it complements existing clinician-reported outcomes (ClinROs) and other PROs. The secondary objective was to describe how TDIS can be used to evaluate the effectiveness of available treatments, particularly VMAT2 inhibitors.

METHODS

As valbenazine is the only medication that has been studied with TDIS, data from valbenazine TD clinical trials KINECT 3²² with its extension study^{23–25} and KINECT 4²⁶ were used in the analyses. Analyses included assessment of correlations between TDIS and ClinROs and other PROs, estimation of the minimal clinically important difference (MCID) for TDIS, also known as the meaningful change threshold, and evaluation of individual items of TDIS longitudinally and via item response theory (IRT).

TDIS was included as an exploratory endpoint in valbenazine clinical trials KINECT 3 and KINECT 4. KINECT 3 was a randomized, placebo-controlled, double-blind trial in which adults with schizophrenia or mood disorder and TD received valbenazine 40 mg or 80 mg daily or placebo for 6 weeks and entered a 42-week extension period, followed by a 4-week washout period.^{23,24} KINECT 4 was a long-term, open-label, 52-week trial assessing the safety and efficacy of valbenazine for the treatment of TD, in which adults with schizophrenia or mood disorder received valbenazine 40 mg or 80 mg daily (dose escalated at week 4) for 48 weeks followed by a 4-week washout period.²⁶ The primary efficacy endpoint in both trials was change in the total AIMS score (sum of items 1–7) from baseline over the study period. The AIMS includes 12 items, of which 3 take into account global judgment.²⁷ While the first 7 AIMS items rate movement severity across body regions, item 8 assesses overall TD severity, item 9 assesses incapacitation due to TD, and item 10 assesses TD awareness and distress. Higher scores on AIMS items 8 and 9 indicate greater disease severity and incapacitation (0 = none to 4 = severe) via clinician assessment, and higher scores on AIMS item 10 indicate greater disease awareness/distress (0 = no awareness to 4 = aware, severe distress) as reported by the patient. The AIMS was scored at week 6 in KINECT 3 and at week 8 in KINECT 4 by 2 central AIMS video raters who were blinded to treatment and time point. Secondary efficacy endpoints included Patient Global Impression of Change

(PGIC) and the Clinical Global Impression of the Patient's Change Specific to TD (CGIC-TD) and scores on all these measures in the long-term extension phases. The PGIC is a single-item PRO that evaluates change in disease on a Likert rating scale.²⁸ The CGIC-TD assesses improvement due to treatment from a clinician's perspective using a Likert rating scale, similar to the PGIC.²⁸ In the KINECT trials, the PGIC and CGIC-TD were scored as 1 = very much improved; 2 = much improved; 3 = minimally improved; 4 = no change; 5 = minimally worse; 6 = much worse; 7 = very much worsened.^{22–26}

To understand how the patient-perceived TD-related impacts correlated with clinician-rated TD severity assessments and the effect of treatment, the relationship between TDIS, AIMS total score, AIMS item 8 (overall severity), and AIMS item 10 (awareness/distress) was examined. AIMS items 8 and 10 scores (assessed by site raters) were only included from KINECT 4 as these items were not assessed in KINECT 3. Participants from both trials were included if they had TD for ≥ 3 months before the study; moderate or severe TD based on their AIMS item 8 score (score of 3 or 4); and a diagnosis of schizophrenia, schizoaffective disorder, or mood disorder. The AIMS total score, AIMS item 8, AIMS item 10, and TDIS total score were assessed at baseline and every 4 weeks through week 52 to evaluate change over time.

The MCID represents the smallest change in score on an outcome measure that patients and clinicians perceive as clinically meaningful.^{29,30} The MCID was estimated using data from participants in KINECT 3 (baseline to week 6) and KINECT 4 (baseline to week 8) who received ≥ 1 dose of valbenazine or placebo and had ≥ 1 postbaseline AIMS assessment. Both anchor- and distribution-based approaches were used; for the anchor-based approach, single-item PGIC and CGIC-TD scores equal to 3 (minimally improved) were used as anchors to define meaningful change.^{28,31,32} Distribution-based approaches included the standardized response mean (SRM) and the standard error of measurement (SEM).³¹ SRM assesses change in relation to sample variation and is independent of sample size. SRM values of 0.2, 0.5, and 0.8 represent small, moderate, and large change, respectively.³¹ SEM assesses the precision of the instrument by incorporating the baseline standard deviation and the reliability coefficient; thresholds from 1.00 to 2.77 SEM have been used to define clinically meaningful differences.³¹

Individual items on TDIS were analyzed using IRT, based on baseline pooled data from participants across all treatment arms. IRT can be used to design or evaluate clinical outcome assessments (COAs) as discussed in the US FDA COA draft guidance document.³³ IRT approaches have been reviewed in published literature.^{34,35} The purpose of IRT was to understand how well each TDIS

item can differentiate between individuals with varying levels of the characteristic being measured (eg, walking, embarrassment, pain). Separate models were used for the 2 domains of TDIS, physical impairment and socioemotional distress.²¹

IRT measures the difficulty, discrimination, and precision of each question in a questionnaire or outcome measure. Discrimination and difficulty parameters were estimated for each item. "Discrimination" shows how well an item differentiates between individuals with different levels of a characteristic. Moderate to very high slopes are preferred, as they indicate greater differentiation between the characteristic levels. Slopes >1.7 indicate very high discrimination, 1.35–1.69 high discrimination, 0.65–1.34 moderate discrimination, and <0.65 low discrimination.³⁶ "Difficulty" indicates how much of the characteristic a person needs to have to agree with or select a response option. Difficulty values are indicated by thresholds and ideally range from -3 to $+3$. A value of -3 indicates lower difficulty (easier endorsement), 0 moderate difficulty, and $+3$ higher difficulty (harder endorsement). A wide range of item thresholds is desirable to accurately capture the spectrum of symptom severity. Furthermore, the precision of individual items is assessed using item information curves, which illustrate the measurement precision of each item across levels of the trait. A steep item information curve indicates greater precision, and a flat slope indicates less precision at a specific level of a characteristic. Finally, overall precision of TDIS was assessed via test information curves (TIC), which shows the reliability of the outcome across the full range of a trait. A broad curve indicates that the items effectively assess the trait across a wide range of severity levels.

Each TDIS item was also evaluated on its ability to detect meaningful change over time. Mean TDIS scores for individual items and for each domain were evaluated at baseline through week 48, with data pooled across treatment arms to focus on the measurement properties of TDIS rather than treatment effects. A subgroup analysis of participants receiving valbenazine was conducted to determine whether item-level improvements seen in the pooled population were also evident with active treatment. A ≥ 0.8 -point change from baseline to week 48 on each item was considered a potentially meaningful improvement based on approximately the 0.5 standard deviation of the in-sample baseline scores across items, ranging from 0.52 to 0.72.³⁷ A distribution-based threshold was selected post hoc to aid interpretation of item-level changes. The 0.8 value was chosen as a conservative estimate of meaningful change. For clinical decision-making for an individual patient, a ≥ 1 -point change on a single item may be considered meaningful as TDIS items are scored in whole-point increments. This threshold is different from the MCID for the total TDIS score because the thresholds for

individual TDIS items and for the TDIS total score were derived independently. The MCID for the TDIS total score is not a sum of thresholds for individual items.

RESULTS

TDIS Trajectory Aligns with Other PROs and ClinROs

TDIS scores in KINECT 3 and KINECT 4 followed a similar trajectory to other PROs and ClinROs over 48 weeks and during the washout period (weeks 48–52). In KINECT 3, mean baseline TDIS scores were 14.6 for the 40 mg/day group and 15.7 for the 80 mg/day group and decreased (improved) over time by -7.8 and -9.5 points, respectively, after 48 weeks of treatment with valbenzazine (range: 0–44; Figure 1). In KINECT 4, mean TDIS scores at baseline were 16.5 and decreased (improved) by -11.0 points from baseline to week 48.

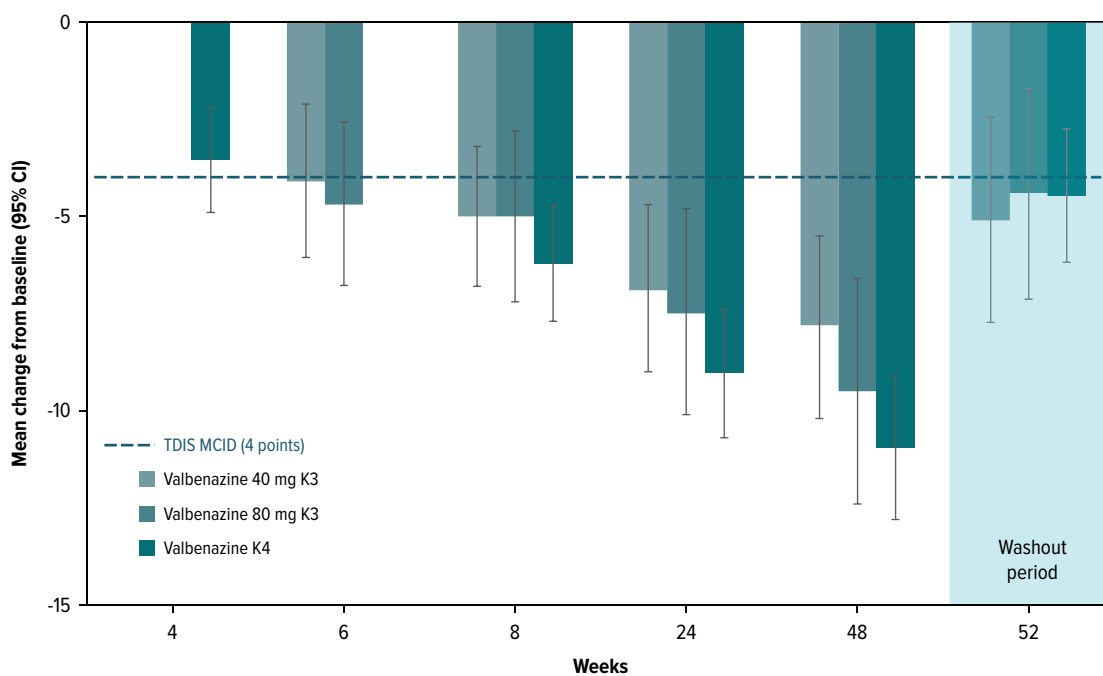
Likewise, total AIMS score and AIMS items 8 and 10 improved over time with valbenzazine treatment. In KINECT 3, the mean baseline total AIMS score was 9.6 for the 40 mg/day group and 10.4 for the 80 mg/day

group, which improved by -3.0 and -4.8 points, respectively, by week 48 (range: 0–28).²³ In KINECT 4, the mean baseline total AIMS score was 14.2 for the 40 mg/day group and 15.0 for the 80 mg/day, which improved by -10.2 and -11.0 points, respectively, over 48 weeks.²⁶ Improvement in AIMS items 8 and 10 scores was also observed; at baseline, the AIMS items 8 and 10 mean scores were 3.2 and 2.7, respectively, which decreased to 1 (minimal movements with no distress) at week 48 (range: 0–4). During the washout period (weeks 48–52), the mean AIMS total score, AIMS item 8, AIMS item 10, and TDIS total score all increased, indicating worsening severity when treatment was removed.^{23,26}

Analyses were conducted to evaluate the relationship between TDIS and COAs after observing the trends in KINECT 3 and KINECT 4. In KINECT 4, at baseline, a low/moderate correlation was observed between TDIS and AIMS total score ($r = 0.31$), AIMS item 8 ($r = 0.22$), and AIMS item 10 ($r = 0.31$), suggesting little overlap between the assessments, each measuring specific aspects of TD severity and impact/burden. Additionally, a comparison between TDIS and PGIC and CGIC-TD scores in both KINECT 3 and KINECT 4 was performed. In KINECT 3, individuals who reported improvement in

Figure 1.

Change from Baseline to Week 52 in TDIS Total Score in Individuals Who Received Valbenzazine 40 mg or 80 mg and Completed KINECT 3 or KINECT 4



Abbreviations: K3 = KINECT 3, K4 = KINECT 4, MCID = minimal clinically important difference, N/A = not applicable, TDIS = Tardive Dyskinesia Impact Scale.

their symptoms (PGIC score 1–3) also had decreased TDIS total scores (ie, less TD impact).²¹ The correlation between PGIC and change in TDIS score at week 6 was moderate ($r = 0.30$). In KINECT 4, a moderate correlation was also seen between change in TDIS score from baseline to week 8 with PGIC ($r = 0.30$) and CGIC-TD ($r = 0.34$).

Although TDIS and AIMS each evaluate distinct aspects of TD, TDIS demonstrated a comparable pattern of improvement to AIMS total score, TD severity (AIMS item 8), and awareness/distress (AIMS item 10) over 48 weeks of treatment and worsening in all measures when treatment was withdrawn in KINECT 4. In KINECT 3, TDIS and total AIMS score followed a similar trajectory with valbenazine treatment. PGIC and CGIC-TD scores also followed a similar pattern observed in TDIS scores in both KINECT 3 and KINECT 4.

Minimal Clinically Important Difference for TDIS

An MCID analysis of AIMS was performed in both the valbenazine and deutetrabenazine regulatory trials, yielding a convergent MCID of 2 points on the AIMS as being indicative of minimal improvement on the CGIC-TD.^{38,39}

For the MCID analyses of TDIS (range: 0–44 points), PGIC and CGIC-TD scores were available at week 6 for 201 individuals in KINECT 3 and at week 8 for 148 individuals in KINECT 4. In individuals reporting minimal improvement on the PGIC (score equal to 3), the mean change in TDIS total score was –3.9 points at week 6 in KINECT 3 and –3.8 points at week 8 in KINECT 4 (Table 1). For those reporting minimal improvement on the CGIC-TD (score equal to 3), the mean change in TDIS total score was –3.1 points in KINECT 3 and –4.0 points in KINECT 4. In KINECT 3, 41% and 43% of participants were minimally improved on the PGIC and CGIC-TD, respectively; in KINECT 4, the corresponding proportions were 32% and 38%.

A large proportion of the KINECT 3 population had improved symptoms with treatment, as indicated by scores ≤ 3 (minimally to very much improved) in the PGIC (74%; $n = 149$) and CGIC-TD (71%; $n = 143$) measures. Similarly, much of the KINECT 4 population had PGIC (86%; $n = 128$) and CGIC-TD (86%; $n = 127$) scores of ≤ 3 . For individuals with CGIC-TD and PGIC of ≤ 3 , the improvement in TDIS total score ranged from –3.1 to –10.2 points at weeks 6 and 8, respectively. The improvement in TDIS score was highest for individuals with PGIC and CGIC-TD scores of 1 (very much improved), ranging from –7.7 to –10.2.

For the distribution-based analysis, the SEM was –1.66 for KINECT 3 and –1.68 for KINECT 4, falling within the suggested range for defining clinically meaningful differences.³¹ The SRM was –0.58 for

Table 1.

Change in Total TDIS Scores From Baseline to Week 6 (KINECT 3) and Week 8 (KINECT 4) by PGIC and CGIC-TD Score

Measure	KINECT 3		KINECT 4	
	n	Mean change in TDIS from baseline	n	Mean change in TDIS from baseline
PGIC				
Very much improved (1)	22	–7.7	29	–9.5
Much improved (2)	44	–5.8	52	–6.4
Minimally improved (3) ^a	83	–3.9	47	–3.8
No change (4)	46	–1.5	15	–1.4
CGIC-TD				
Very much improved (1)	10	–10.2	21	–8.6
Much improved (2)	46	–5.9	50	–8.0
Minimally improved (3) ^a	87	–3.1	56	–4.0
No change (4)	58	–2.9	19	–1.1

^aThe MCID for TDIS was defined as the mean change corresponding to the “minimally improved” category (PGIC or CGIC-TD = 3).

Abbreviations: CGIC-TD = Clinician’s Global Impression of the Patient’s Change Specific to TD, MCID = minimal clinically important difference, PGIC = Patient’s Global Impression of Change, TDIS = Tardive Dyskinesia Impact Scale.

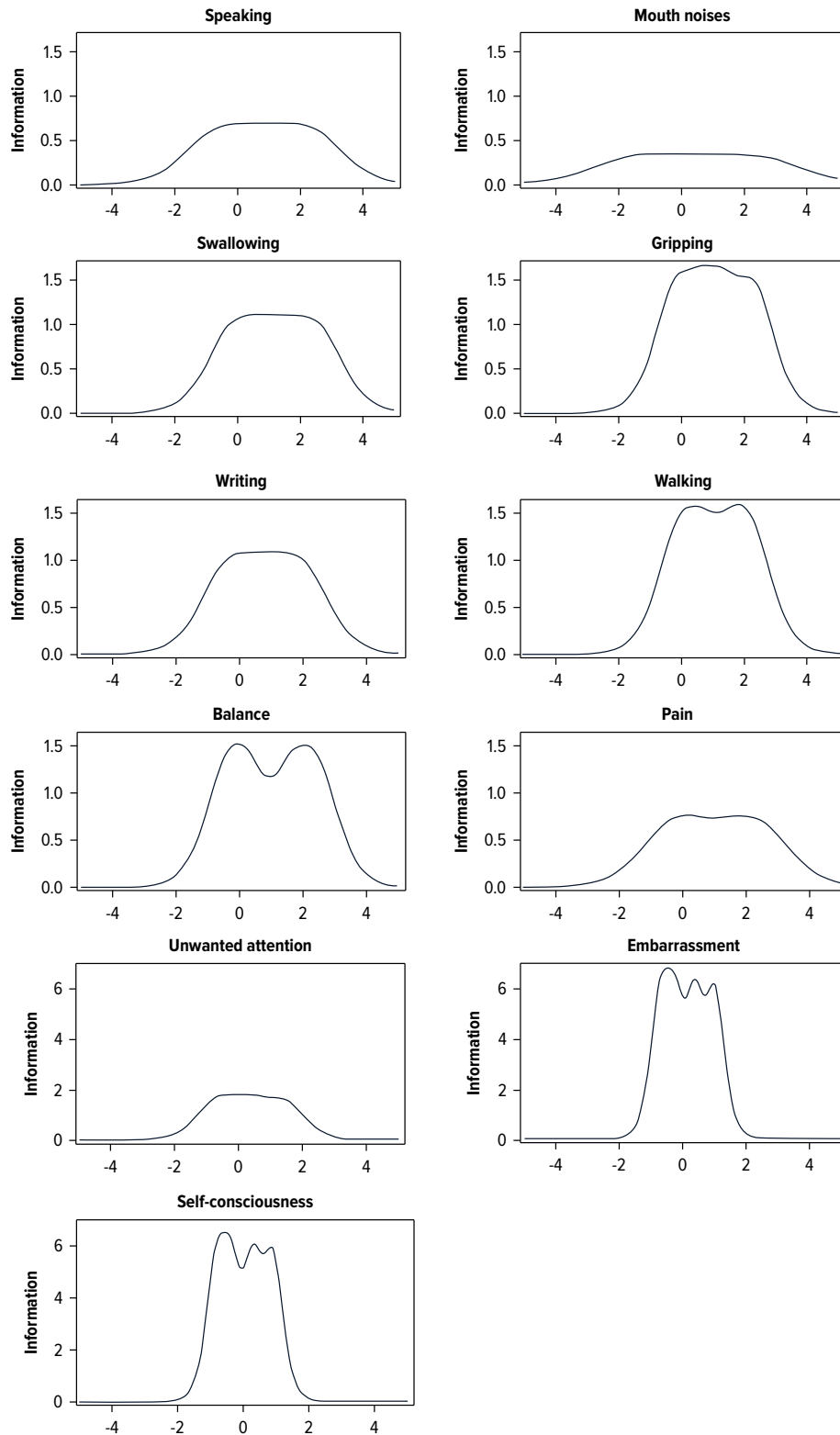
KINECT 3 and –0.74 for KINECT 4, suggesting a moderate to large effect size change.³¹

An MCID estimate of 4 points was derived using in-sample anchor- and distribution-based approaches, anchored by PGIC scores equal to 3 and supported by CGIC-TD scores equal to 3 (minimally improved). This suggests that a 4-point change in TDIS total score (range: 0–44 points) may represent meaningful improvement in TD symptoms. In KINECT 3 and KINECT 4, the mean change in TDIS scores for individuals who completed the studies consistently met or exceeded the MCID at each study week (Figure 1).

TDIS Item Analysis with Item Response Theory and Change over Time

For the IRT analyses of individual TDIS items, combined data from KINECT 3 or KINECT 4 were used, including 388 individuals with TD who completed TDIS at baseline and at least 1 follow-up visit, regardless of treatment assignment. The model for physical distress demonstrated strong measurement properties across a broad range of physical impairment levels, capturing individuals with low, moderate/average, and high impairment. Seven of 8 items in the physical function domain (gripping, walking, balance, swallowing, writing, pain, and speaking) exhibited high discrimination (slope range: 1.51–2.35; Figure 2), effectively differentiating levels of impairment. The remaining item (mouth noises) provided moderate discrimination indicated by a flatter slope (slope = 1.08). No items showed low discrimination. Items varied in difficulty (thresholds range: –1.57 to 2.58), with lower-difficulty items endorsed at lower impairment levels and higher-difficulty items endorsed at higher impairment levels.

Figure 2.
Item Information Curves for Physical and Socioemotional Domains of TDIS^a



^aItem information curves illustrate the measure precision of each item across levels of the trait. A steep curve represents greater precision, while a flat curve indicates less precision. For example, gripping has a steep curve and mouth noises has a flat curve.

Abbreviation: TDIS = Tardive Dyskinesia Impact Scale.

The TIC indicated that measurement precision was highest for individuals with average to above-average physical impairment (TIC peak precision around 0 and +2; Figure 3).

All 3 items in the socioemotional domain exhibited high to very high discrimination, evidenced by the steep slopes (slope range: 2.44–4.89), effectively differentiating between levels of socioemotional distress (Figure 2). The difficulty thresholds ranged from low to high, indicating that response options captured varying severity levels (threshold range: –0.71–1.47). The TIC showed that the socioemotional domain provided the most precise measurement for individuals with average socioemotional distress (TIC peaks around –0.5, +0.5, +1; Figure 3). However, as this domain contains fewer items than the physical function domain (3 vs 8), it may provide less precise estimates for individuals with very low or very high socioemotional distress.

To further explore the sensitivity of individual items to change over time, longitudinal changes in TDIS item scores were evaluated. Altogether, 252 individuals, regardless of treatment assignment (valbenazine or placebo), who completed 1 of the phase 3 trials and had TDIS data at baseline and at least 1 follow-up, were included in the analysis. In both the physical and socioemotional domains of TDIS, the mean scores for all items trended downward, denoting improvement. The physical domain score improved (decreased) by 5.86 points from baseline to week 48; the socioemotional domain improved by 3.44 points. In the physical domain, the most impacted items at baseline were mouth noises (1.80), writing (1.29), and speaking (1.28). The greatest improvement from baseline to week 48 was observed for mouth noises (mean change: –1.05), speaking (–0.88), pain (–0.81), and writing (–0.78); 3 of the items exceeded the definition of meaningful improvement of ≥ 0.8 (Figure 4). In the socioemotional domain, the most impacted items at baseline were self-consciousness (mean score: 1.99), followed by embarrassment (1.89), and unwanted attention (1.74). At week 48, self-consciousness (mean change: –1.24), embarrassment (–1.19), and unwanted attention (–1.00) exceeded the definition of meaningful improvement from baseline.

A separate analysis of individuals receiving valbenazine was conducted to assess whether the patterns of item-level change were consistent and more pronounced in those receiving active therapy. Among the 181 individuals receiving valbenazine, similar patterns of item-level improvement were observed. In the socioemotional domain, the largest mean changes from baseline to Week 48 were seen for self-consciousness (–1.34), embarrassment (–1.28), and unwanted attention (–1.12). In the physical domain, mouth noises (–1.12), speaking (–0.92), and pain (–0.91) showed the

greatest improvements. These items exceeded the threshold for meaningful improvement of ≥ 0.8 (Supplementary Figure 1).

In clinical practice, a ≥ 1 -point change on any individual item is considered meaningful for assessing change in an individual patient; however, a threshold ≥ 0.8 points was applied in these analyses since group-level mean change scores were evaluated.

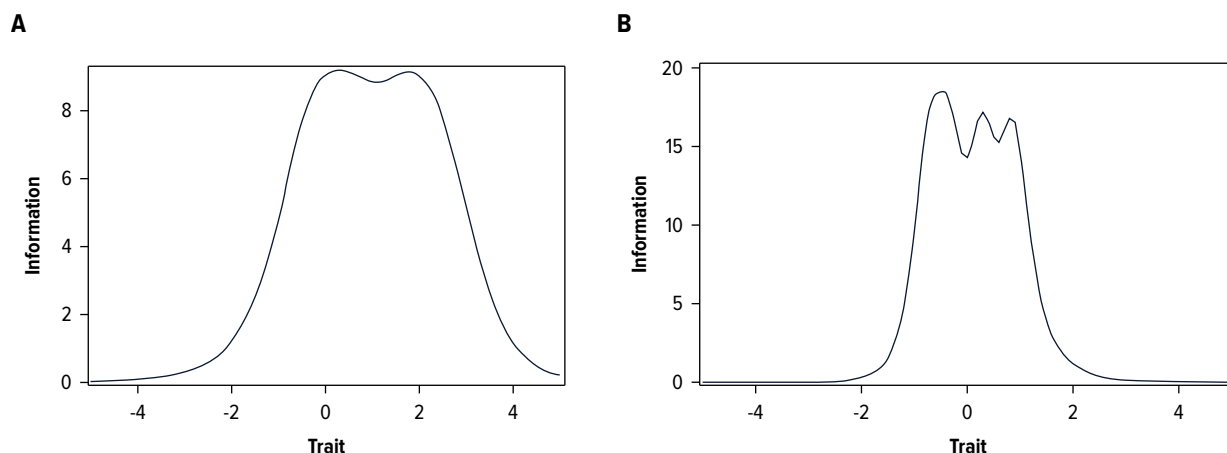
DISCUSSION

Dyskinesia associated with TD can be burdensome on an individual's and caregiver's life.^{40–43} In the prospective, observational, multicenter RE-KINECT study conducted at 37 outpatient psychiatry centers in the US, which surveyed 204 individuals with suspected TD during a 2-week period, $>40\%$ of individuals reported TD movements had a profound impact on aspects of daily living such as talking, eating, being productive, and socializing.⁹ A survey of 269 individuals with self-reported TD showed similar results, with $>80\%$ reporting being bothered by TD symptoms.⁴⁰ Individuals who were employed ($n = 193$) reported decreased presenteeism (mean, 68.4% impairment) and increased absenteeism (mean, 29.1% work time missed) in the last 7 days due to TD. Individuals with TD reported worse well-being, HRQOL, and social withdrawal when compared with individuals with schizophrenia, bipolar disorder, or MDD without TD and with the age- and sex-adjusted general population.⁴⁴

Caregivers of individuals with TD also reported impacts on their own abilities to function and on their HRQOL; when asked to rate the impact of caring for individuals with TD on their own life on a scale from 0 (no impact at all) to 10 (impacted as bad as you can imagine), the median score was 6.⁴⁵ In a separate survey of 162 caregivers of individuals with TD, caregivers reported 46.4% activity impairment and 49.5% work impairment (for those who were employed, $n = 136$) in the past 7 days based on data from the Work Productivity and Activity Impairment questionnaire.⁴¹

Additionally, the impact of TD on daily activities does not necessarily align with the severity of uncontrollable movements. In a survey in which 80% of individuals with TD reported being bothered by TD symptoms, only 40% reported TD symptoms to be severe or very severe in the previous week.⁴⁰ These survey findings are supported by exploratory analyses of the RE-KINECT data where self-rated impact on daily activities, such as self-care, socializing, and being productive, was significantly associated with HRQOL decline ($P < .01$), while self-rated symptom severity was also associated, though at a lower level of significance ($P < .05$). Similarly, functional impairment was significantly associated with patient-rated impact ($P < .001$), whereas no significant association

Figure 3.
Test Information Curve for TDIS Physical (A) and Socioemotional (B) Domains^a



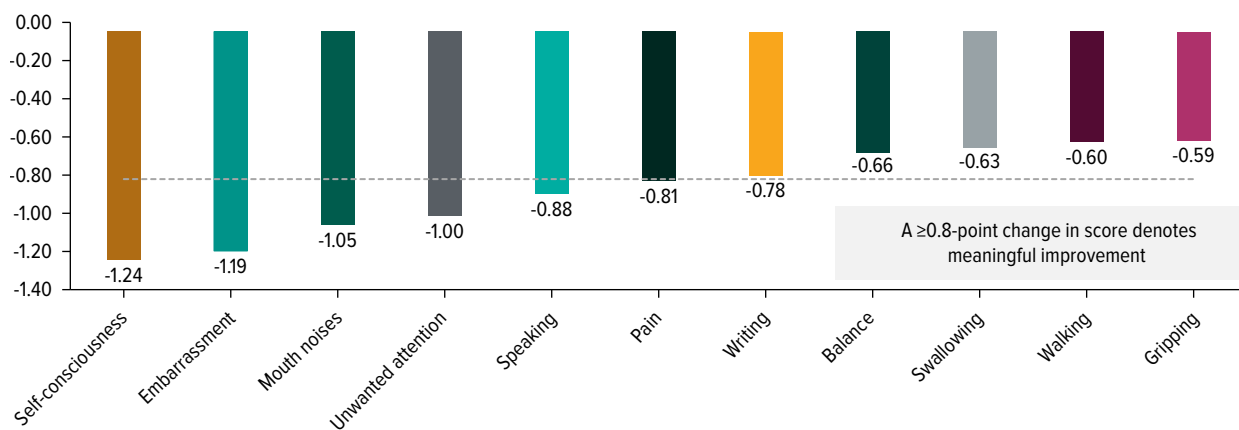
^aTest information curves illustrate the reliability of the TDIS measure. A broad curve indicates that the items effectively assess the trait (specifically physical and socioemotional domains for TDIS) across a range of severity levels.
Abbreviation: TDIS = Tardive Dyskinesia Impact Scale.

was observed with symptom severity ($P = .560$).⁴³ The findings from the survey and RE-KINECT study highlight the fact that TD severity does not always correlate with impact on daily functioning, so that other factors need to be considered. Some individuals who have mild symptoms may still find that their symptoms are debilitating, bothersome, or embarrassing. These results provide additional rationale for developing a TD-specific measure that is self-reported by the patient and includes items beyond movement symptoms.

PROs are becoming increasingly used in many disease states to better understand the individual's

perspective of their symptoms and experience during their illness or related to treatment.^{46,47} Recognizing the value of PROs and ClinROs, the FDA released guidance documents on their use in clinical trials for new medications.⁴⁸ PROs can be categorized into generic or disease-specific. Generic PROs are appropriate for use with individuals in any disease state or population and allow for comparisons across different diseases and populations. Disease-specific PROs assess symptoms, functional impact, or other aspects that are relevant to individuals with a particular health condition.^{46,47,49} Currently, TDIS is the only disease-specific PRO that has been designed for and validated in TD and specifically

Figure 4.
Change in TDIS Item Mean Scores from Baseline to Week 48 (n = 252)



Abbreviation: TDIS = Tardive Dyskinesia Impact Scale.

assesses symptoms and functional impacts that are directly relevant to TD.

An example of a generic PRO that has been used in trials with VMAT2 inhibitors is the PGIC. Treatment with valbenazine has been associated with improvement in PGIC in phase 3 trials, with 88% of individuals reporting “much improved” or “very much improved” at 12 months (48 weeks).⁵⁰ Deutetrabenazine has also been associated with improvement in PGIC in an open-label extension study, with 61% of individuals reporting “much improved” or “very much improved” at 13.5 months (54 weeks).⁵¹

Other generic PROs have been used in clinical trials assessing individuals with TD receiving valbenazine, such as EQ-5D-5L⁵² and Sheehan Disability Scale (SDS).⁵³ EQ-5D-5L measures an individual’s current health state, and SDS measures the severity of functional impairment.^{52,53} Valbenazine was associated with improvement in these measures in a phase 4, double-blind, placebo-controlled withdrawal study of individuals with TD receiving valbenazine (NCT03891862).⁴² In individuals (n = 59) receiving valbenazine from baseline through week 16, the EQ-5D-5L utility index improved by 0.17 (range: -0.573 to 1), the EQ-5D-5L visual analog scale by 6.4 points (range: 0–100), and the total SDS by 9.1 points (range: 0–30).⁴²

ClinROs, as defined by the FDA, are measurements based on observation of an individual’s health condition by a trained healthcare professional.⁵⁴ ClinROs used in TD include the AIMS, CGIC-TD, Impact-TD and Clinician’s Tardive Inventory (CTI).^{28,55–57} AIMS was a primary endpoint and CGIC was a secondary endpoint in phase 3 clinical trials for valbenazine and deutetrabenazine.^{22–26,51,58} CGIC-TD and AIMS do not measure specific domains in detail of the functional or social/emotional impact of TD. Impact-TD is a recently developed ClinRO that assesses individuals’ TD-related HRQOL. It includes 4 functional domains (social, psychological/psychiatric, physical, and vocational/educational/recreational), with each domain rated from 0 (no impact) to 3 (severe impact).⁵⁵ The CTI assesses the presence, frequency, and amplitude of abnormal movements in 6 anatomic domains and level of functional impairment in 5 domains, with higher values indicating greater frequency of symptoms (range: 0–30) and more severe functional impairment (range: 0–3).^{56,57}

While these tools are valuable for understanding disease severity and impacts on HRQOL from a clinician’s viewpoint, clinicians may not have the same perception of the impact and severity of TD as patients and their caregivers. For example, in an analysis comparing HRQOL and functional impairment between individuals with possible TD vs those without abnormal involuntary movements, those who rated their TD impact as severe reported significantly worse HRQOL (EQ-5D-5L score $P < .001$) and functional impairment (SDS score $P < .001$) than those with clinician-rated severity alone (EQ-5D-5L and SDS scores $P > .05$).⁴³ Similarly, severity of abnormal

involuntary movements rated by caregivers (rated as “none,” “some,” or “a lot” for head/face neck/trunk, upper extremities, lower extremities) was more strongly associated with patient ratings of severity ($P < .01$ for all locations) compared with clinician ratings of severity ($P < .05$ for neck/trunk and upper and lower extremities).⁴⁵ Additionally, caregivers and patients reported greater severity and more affected body regions vs clinicians.⁴⁵ This discrepancy may be due to the subjective nature of TD severity and impact; while clinicians may consider the severity as mild, patients may report that TD symptoms have a substantial impact on their socioemotional functioning or activities of daily living.^{9,43} Additionally, ClinROs such as AIMS measure different aspects of TD compared with PROs. By using ClinROs and PROs together, a more complete picture of the burden and impact of TD from the patient’s perspective can be captured.²¹

TDIS, a TD-specific PRO, can help individuals quantify the impact of symptoms on daily life and well-being and provide a benchmark for the effect of treatment by the clinician. TDIS uniquely complements existing COAs, as it shows a similar pattern of change with treatment compared with AIMS, PGIC, and CGIC-TD, while measuring distinct and clinically relevant aspects of TD that extend beyond movement symptoms. An MCID of 4 points in TDIS total score (range: 0–44) was established using anchor- and distribution-based methods to help guide interpretation of TDIS.

The IRT analyses demonstrated that the individual items in TDIS are reliable and precise in measuring physical, social, and emotional impacts of TD and that each question captures unique aspects of TD impact. The longitudinal item analysis response provided clarity on the most impacted items (embarrassment, unwanted attention, self-consciousness, mouth noises, and speaking) and captured meaningful change of these items over the 48-week study period as measured by the within-sample benchmarks with valbenazine treatment. These results align with the results of the qualitative studies completed during the development of the TDIS, in which greater than 50% of participants endorsed unwanted social attention, difficulty speaking, and social isolation as the most common impacts of TD.²¹ Based on distribution-based analyses, the initial estimate of meaningful change for individual items ranged from 0.52 to 0.72 points; a conservative threshold of 0.8 points was selected post hoc as a preliminary reference point to guide interpretation of item-level changes. For clinical use, a ≥ 1 -point change in single TDIS items can be considered meaningful given the whole-point scoring structure. However, caution is warranted, as the meaningful change threshold for total TDIS score was estimated separately using an anchor-based approach and is not equivalent to the sum of individual item thresholds. Anchor-based methods may be applied in the future to further refine thresholds at the item level.

Although TDIS has been psychometrically validated in patients with TD, there are some limitations to consider. The figures presented herein do not include placebo data, reflecting the focus of these analyses on descriptive patterns of change in the active treatment valbenazine group rather than comparative treatment effects. Future research incorporating control groups will help further contextualize observed changes. TDIS has only been studied in patients with TD receiving valbenazine and has not been assessed with other VMAT2 inhibitors or in untreated patients with TD. In addition, some patients may be unaware of TD movements, such as those with schizophrenia, which could affect how they perceive their symptoms.

TDIS is not meant to replace any PRO or ClinRO but rather be used alongside other COAs (eg, AIMS and CGIC-TD) to provide a holistic view of patient burden and improvement. Indeed, the use of TDIS, along with other PROs and ClinROs, characterized functional and socioemotional impacts of TD and their change over time along with the changes in the motor manifestations of TD in valbenazine phase 3 and long-term extension trials.^{22–26} Given the relevance of TD, using PROs like TDIS can provide a comprehensive picture of the impact of TD when used in conjunction with ClinROs and are key to understanding the underappreciated impact and burden of TD from the patient's perspective and the potential benefit of TD treatments.

Article Information

Published Online: April 29, 2026. <https://doi.org/10.4088/JCP.25nr16047>
© 2026 Physicians Postgraduate Press, Inc.

Submitted: July 25, 2025; accepted January 30, 2026.

To Cite: Bron M, Mathias SD, Stull DE, et al. Measuring what matters: further validation for the Tardive Dyskinesia Impact Scale, a novel patient-reported outcome measure in valbenazine clinical trials. *J Clin Psychiatry* 2026;87(2):25nr16047.

Author Affiliations: Neurocrine Biosciences, Inc., San Diego, California (Bron, Zhang, Dunayevich, Parameswaran, Vanderhoef); Health Outcomes Solutions, Palm Beach Gardens, Florida (Mathias, Turner); IQVIA, Durham, North Carolina (Stull); Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York (Perez-Rodriguez); The Zucker Hillside Hospital; Glen Oaks, New York (Correll); The Donald and Barbara Zucker School of Medicine at Hofstra/Northwell; Hempstead, New York (Correll); Charité Universitätsmedizin, Department of Child and Adolescent Psychiatry; Berlin, Germany (Correll); German Center for Mental Health (DZPG), partner site Berlin, Berlin, Germany (Correll).

Corresponding Author: Morgan Bron, PharmD, MS, 12780 El Camino Real, San Diego, CA 92130 (mbron@neurocrine.com).

Relevant Financial Relationships: Drs Bron, Zhang, Dunayevich, Parameswaran, and Vanderhoef are employees and stockholders of Neurocrine Biosciences, Inc., and have no other disclosures. Ms Mathias is an employee of and Ms Turner is a consultant of Health Outcomes Solutions. Dr Stull is an employee of IQVIA. Dr Perez-Rodriguez is a consultant of Neurocrine Biosciences, Inc., and Mitsubishi Tanabe Pharma Corporation and has received grant funding from Neurocrine Biosciences, Inc. Dr Correll has been a consultant and/or advisor to or has received honoraria from AbbVie, Alkermes, Allergan, Angelini, Aristo, Autobahn, Boehringer-Ingelheim, Bristol-Myers Squibb, Cardio Diagnostics, Cerevel, CNX Therapeutics, Compass Pathways, Darnitsa, Delpor, Denovo, Draig, Eli Lilly, EuMentis Therapeutics, Gedeon Richter, GH, Hikma, Holmusk, Intra-Cellular Therapies, Jamjoom Pharma, Janssen/J&J, Karuna, LB Pharma, Lundbeck, MedInCell, MedLink, Merck, Mindpax, Mitsubishi Tanabe Pharma, MapLight, Mylan, Neumora Therapeutics, Neuraxpharm, Neurocrine Biosciences, Inc., Neurelis, Newron, Noven, Novo Nordisk, Otsuka, PPD Biotech, Recordati, Relmada, Response Pharmaceuticals, Reviva, Rovi, Saladax, Sanofi, Seqirus,

Servier, Sumitomo Pharma America, Sunovion, Sun Pharma, Supernus, Tabuk, Takeda, Teva, Terran, Tolmar, Vertex, Viatrix, and Xenon Pharmaceuticals. He provided expert testimony for Janssen, Lundbeck, and Otsuka. He served on a Data Safety Monitoring Board for Compass Pathways, Intra-Cellular Therapies, Relmada, Reviva, and Rovi. He has received grant support from Boehringer-Ingelheim, Janssen, and Takeda. He received royalties from UpToDate and is also a stock option holder of Cardio Diagnostics, Kuleon Bioscience, LB Pharma, MedLink Global, Mindpax, Quantic, Terran.

Funding/Support: Neurocrine Biosciences, Inc., provided funding for this study and development of the manuscript.

Role of the Sponsor: Neurocrine Biosciences, Inc., assisted in the conduct of the study; study design, analysis, and interpretation of the data; and the preparation, review, and approval of the manuscript.

Previous Presentation: Portions of the data have been previously presented as posters including Psych Congress Elevate, May 28–31, 2025, Las Vegas, Nevada; Psych Congress 2024, October 29–November 2, 2024, Boston, Massachusetts; Neuroscience Education Institute Congress 2024, November 7–10, 2024, Colorado Springs, Colorado; ISPOR Europe 2023, November 12–15, 2023, Copenhagen, Denmark.

Acknowledgments: The authors acknowledge the late Ross D. Crosby, PhD, of Sandford Research, Sioux Falls, South Dakota, for his significant contributions to this field. Dr Crosby had no financial disclosures. The authors also acknowledge Sara Gao, MS, of Neurocrine Biosciences, Inc., San Diego, California, and Connie Ly, MScPh, of IQVIA, Durham, North Carolina, for assistance with statistical analyses, and Andi Gundlach, PharmD, and Bridgette Schroeder, PharmD of Cencora, Conshohocken, Pennsylvania, for their aid in medical writing and strategic support. Ms Gao is an employee of Neurocrine Biosciences, and Ms Ly was contracted with Neurocrine Biosciences at the time of the analyses. Drs Gundlach and Schroeder are employees of Cencora, which is contracted with Neurocrine Biosciences.

Supplementary Material: Available at Psychiatrist.com.

References

- Caroff SN. Overcoming barriers to effective management of tardive dyskinesia. *Neuropsychiatr Dis Treat*. 2019;15:785–794.
- Kane JM, Correll CU, Nierenberg AA, et al. Revisiting the Abnormal Involuntary Movement Scale: proceedings from the Tardive Dyskinesia Assessment Workshop. *J Clin Psychiatry*. May/June 2018;79(3):17cs11959.
- Hauser RA, Meyer JM, Factor SA, et al. Differentiating tardive dyskinesia: a video-based review of antipsychotic-induced movement disorders in clinical practice. *CNS Spectr*. 2022;27(2):208–217.
- Zádori D, Veres G, Szalárdy L, et al. Drug-induced movement disorders. *Expert Opin Drug Saf*. 2015;14(6):877–890.
- Solmi M, Pigato G, Kane JM, et al. Clinical risk factors for the development of tardive dyskinesia. *J Neural Sci*. 2018;389:21–27.
- Caroff SN, Ungvari GS, Cunningham ODG. Historical perspectives on tardive dyskinesia. *J Neural Sci*. 2018;389:4–9.
- Kline NS. On the rarity of “irreversible” oral dyskinesias following phenothiazines. *Am J Psychiatr*. 1968;124(8S):48–54.
- Baminiwatta A, Correll CU. Historical developments, hotspots, and trends in tardive dyskinesia research: a scientometric analysis of 54 years of publications. *Front Psychiatry*. 2023;14:1194222.
- Caroff SN, Yeomans K, Lenderking WR, et al. RE-KINECT: a prospective study of the presence and healthcare burden of tardive dyskinesia in clinical practice settings. *J Clin Psychopharmacol*. 2020;40(3):259–268.
- 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.
- Gardea-Resendez M, Taylor-Desir MJ, Romo-Nava F, et al. Clinical phenotype of tardive dyskinesia in bipolar disorder. *J Clin Psychopharmacol*. 2022;42(2):159–162.
- Loughlin AM, Lin N, Abler V, et al. Tardive dyskinesia among patients using antipsychotic medications in customary clinical care in the United States. *PLoS One*. 2019;14(6):e0216044.
- American Psychiatric Association (APA). *Medication-Induced Movement Disorders and Other Adverse Effects of Medication*. Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association Publishing; 2022.
- Alabaku O, Yang A, Tharmarajah S, et al. Global trends in antidepressant, atypical antipsychotic, and benzodiazepine use: a cross-sectional analysis of 64 countries. *PLoS One*. 2023;18(4):e0284389.
- American Psychiatric Association (APA). *Guideline statements and implementation*. The American Psychiatric Association Practice Guideline for the Treatment of Patients With Schizophrenia. American Psychiatric Association Publishing; 2020.
- Bhidayasiri R, Fahn S, Weiner WJ, et al. Evidence-based guideline: treatment of tardive syndromes. *Rep Guidel Develop Subcomm Am Acad Neurology*. 2013;81(5):463–469.

17. Bhidayasiri R, Jitkritsakul O, Friedman JH, et al. Updating the recommendations for treatment of tardive syndromes: a systematic review of new evidence and practical treatment algorithm. *J Neurol Sci.* 2018;389:67–75.
18. 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.
19. Deutetrabenazine [prescribing information]. Teva Neuroscience, Inc; 2024.
20. Valbenazine [prescribing information]. Neurocrine Biosciences, Inc; 2024.
21. Farber RH, Stull DE, Witherspoon B, et al. The Tardive Dyskinesia Impact Scale (TDIS), a novel patient-reported outcome measure in tardive dyskinesia: development and psychometric validation. *J Patient-Rep Outcomes.* 2024;8(1):2.
22. Hauser RA, Factor SA, Marder SR, et al. Kinect 3: a phase 3 randomized, double-blind, placebo-controlled trial of valbenazine for tardive dyskinesia. *Am J Psychiatry.* 2017;174(5):476–484.
23. Factor SA, Remington G, Comella CL, et al. The effects of valbenazine in participants with tardive dyskinesia: results of the 1-Year KINECT 3 Extension Study. *J Clin Psychiatry.* 2017;78(9):1344–1350.
24. Correll CU, Josiassen RC, Liang GS, et al. Efficacy of valbenazine (NBI-98854) in treating subjects with tardive dyskinesia and mood disorder. *Psychopharmacol Bull.* 2017;47(3):53–60.
25. Kane JM, Correll CU, Liang GS, et al. Efficacy of valbenazine (NBI-98854) in treating subjects with tardive dyskinesia and schizophrenia or schizoaffective disorder. *Psychopharmacol Bull.* 2017;47(3):69–76.
26. Marder SR, Singer C, Lindenmayer JP, et al. A phase 3, 1-year, open-label trial of valbenazine in adults with tardive dyskinesia. *J Clin Psychopharmacol.* 2019;39(6):620–627.
27. Abnormal Involuntary Movement Scale (117-AIMS). In: Guy W, ed. *ECDEU Assessment Manual for Psychopharmacology: DHEW publication; no (ADM) 76-338.* National Institute of Mental Health.; 1976:534–537.
28. Clinical Global Impressions (028-CGI). In: Guy W, ed. *ECDEU Assessment Manual for Psychopharmacology: DHEW publication; no (ADM) 76-338.* National Institute of Mental Health; 1976:217–222.
29. Thwin SS, Hermes E, Lew R, et al. Assessment of the minimum clinically important difference in quality of life in schizophrenia measured by the Quality of Well-Being Scale and disease-specific measures. *Psychiatry Res.* 2013;209(3):291–296.
30. Mishra B, Sudheer P, Rajan R, et al. Bridging the gap between statistical significance and clinical relevance: a systematic review of minimum clinically important difference (MCID) thresholds of scales reported in movement disorders research. *Heliyon.* 2024;10(5):e26479.
31. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol.* 2003;56(5):395–407.
32. Hurst H, Bolton J. Assessing the clinical significance of change scores recorded on subjective outcome measures. *J Manip Physiol Ther.* 2004;27(1):26–35.
33. Center for Drug Evaluation and Research. Patient-focused drug development: selecting, developing, or modifying fit-for purpose clinical outcome assessments guidance for industry. Food and Drug Administration staff, and other stakeholders draft guidance; 2022. Accessed February 21, 2025. <https://www.fda.gov/media/159500/download>
34. Cappelleri JC, Jason Lundy J, Hays RD. Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clin Ther.* 2014;36(5):648–662.
35. Nguyen TH, Lee CS, Kim MT. Using item response theory to develop and refine patient-reported outcome measures. *Eur J Cardiovasc Nurs.* 2022;21(5):509–515.
36. Steiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to Their Development and Use (5th edition).* Oxford University Press; 2024:480.
37. Norman GR, Sloan JA, Wywich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care.* 2003;41(5):582–592. doi:10.1097/01.Mlr.0000062554.74615.4c.
38. Stacy MD, Sajatovic M, Kane JM, et al. Abnormal involuntary movement Scale in tardive dyskinesia: minimal clinically important difference. *Movement Disord.* 2019;34(8):1203–1209.
39. 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.
40. Jain R, Ayyagari R, Goldschmidt D, et al. Impact of tardive dyskinesia on physical, psychological, social, and professional domains of patient lives: a survey of patients in the United States. *J Clin Psychiatry.* 2023;84(3):22m14694.
41. Jain R, Ayyagari R, Goldschmidt D, et al. Impact of tardive dyskinesia on patients and caregivers: a survey of caregivers in the United States. *J Patient-Rep Outcomes.* 2023;7(1):122.
42. Serbin M, Nedzesky J, Bron M, et al. Impact of valbenazine on work/school, family, and social life in patients with tardive dyskinesia (N=59): Results from a double-blinded, placebo-controlled, phase 4 randomized withdrawal study presented at: Poster presented at: Academy of Managed Care Pharmacy (AMCP) Nexus 2024 Meeting; October 14–17, 2024. Las Vegas, Nevada.
43. Tanner CM, Caroff SN, Cutler AJ, et al. Impact of possible tardive dyskinesia on physical wellness and social functioning: results from the real-world RE-KINECT study. *J Patient-Rep Outcomes.* 2023;7(1):21.
44. 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.
45. Cutler AJ, Caroff SN, Tanner CM, et al. Caregiver-reported burden in RE-KINECT: data from a prospective real-world tardive dyskinesia screening study. *J Am Psychiatr Nurses Assoc.* 2023;29(5):389–399.
46. Churrua K, Pomare C, Ellis LA, et al. Patient-reported outcome measures (PROMs): a review of generic and condition-specific measures and a discussion of trends and issues. *Health Expect.* 2021;24(4):1015–1024.
47. Slade A, Isa F, Kyte D, et al. Patient reported outcome measures in rare diseases: a narrative review. *Orphanet J Rare Dis.* 2018;13(1):61.
48. Food and Drug Administration (FDA). Guidance for industry. In: *Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims;* 2023. <https://www.fda.gov/media/77832/download>
49. Lin X-J, Lin IM, Fan S-Y. Methodological issues in measuring health-related quality of life. *Tzu Chi Med J.* 2013;25(1):8–12. <https://doi.org/10.1016/j.tcmj.2012.09.002>
50. Correll CU, Citrome L, Singer C, et al. Sustained treatment response and global improvements with long-term valbenazine in patients with tardive dyskinesia. *J Clin Psychopharmacol.* 2024;44(4):353–361.
51. Hauser RA, Barkay H, Fernandez HH, et al. Long-Term deutetrabenazine treatment for tardive dyskinesia is associated with sustained benefits and safety: a 3-year, open-label extension Study. *Front Neurol.* 2022;13:773999.
52. EuroQol. EQ-5D-5L. 2025. Accessed February 4, 2025. <https://euroqol.org/information-and-support/euroqol-instruments/eq-5d-5l/>
53. Sheehan DV. *The Anxiety Disease.* Charles Scribner and Sons; 1983.
54. FDA Center for Drug Evaluation and Research. Clinical outcome assessment (COA) compendium. 2021. Accessed December 10, 2024. <https://www.fda.gov/media/130138/download?attachment>
55. Jackson R, Brams MN, Carozzi NE, et al. Impact-Tardive Dyskinesia (Impact-TD) Scale: a clinical tool to assess the impact of tardive dyskinesia. *J Clin Psychiatry.* 2022;84(1):22cs14563.
56. Trosch RM, Shillington AC, Comella CL, et al. A validation study of the Clinician's Tardive Inventory (CTI). *Parkinsonism Relat Disord.* 2025;135:107812.
57. 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.
58. 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.

See supplementary material for this article at [Psychiatrist.com](https://www.psychiatrist.com)

Supplementary Material

Article Title: Measuring What Matters: Further Validation for the Tardive Dyskinesia Impact Scale, a Novel Patient-Reported Outcome Measure in Valbenazine Clinical Trials

Authors: Morgan Bron, PharmD, MS; Susan D. Mathias, MPH; Donald E. Stull, PhD; Hui Zhang, PhD; Michelle Turner, MS; Eduardo Dunayevich, MD; Ashok Parameswaran, MD; Dawn Vanderhoef, PhD, DNP, PMHNP-BC, FAANP; M. Mercedes Perez-Rodriguez, MD, PhD; Christoph U. Correll, MD

DOI Number: 10.4088/JCP.25nr16047

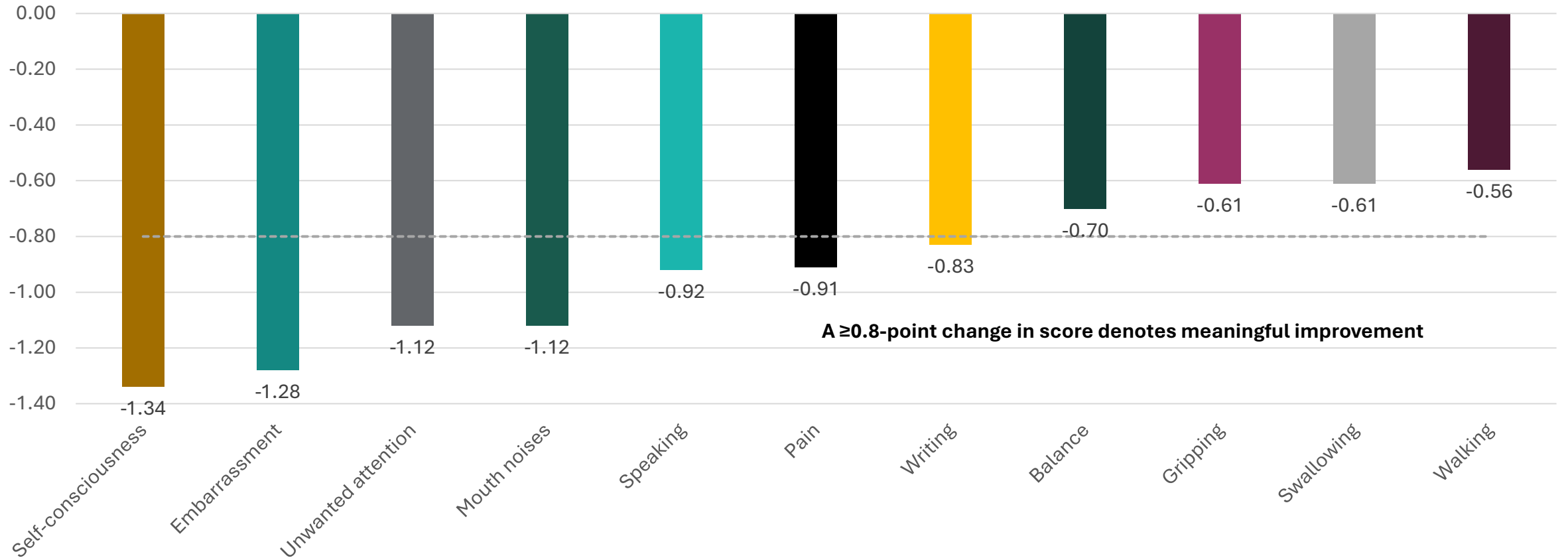
LIST OF SUPPLEMENTARY MATERIAL FOR THE ARTICLE

1. [Figure 1](#) Change in TDIS Item Mean Scores From Baseline to Week 48 for Individuals Receiving Valbenazine (n=181)

DISCLAIMER

This Supplementary Material has been provided by the authors as an enhancement to the published article. It has been approved by peer review; however, it has undergone neither editing nor formatting by in-house editorial staff. The material is presented in the manner supplied by the author.

Supplementary Figure 1. Change in TDIS item mean scores from baseline to week 48 for individuals receiving valbenzazine (n=181)



Abbreviations: TDIS: Tardive Dyskinesia Impact Scale.