Journal Articles

2019

Abnormal involuntary movement scale in tardive dyskinesia: Minimal clinically important difference

M. Stacy
M. Sajatovic
J. M. Kane
Zucker School of Medicine at Hofstra/Northwell, jkane2@northwell.edu
A. J. Cutler
G. S. Liang

See next page for additional authors

Follow this and additional works at: https://academicworksmedicine.hofstra.edu/articles

Part of the Psychiatry Commons

Recommended Citation
Stacy M, Sajatovic M, Kane JM, Cutler AJ, Liang GS, O'Brien CF, Correll CU. Abnormal involuntary movement scale in tardive dyskinesia: Minimal clinically important difference. . 2019 Jan 01; 34(8):Article 5738 [ p.]. Available from: https://academicworksmedicine.hofstra.edu/articles/5738. Free full text article.

This Article is brought to you for free and open access by Donald and Barbara Zucker School of Medicine Academic Works. It has been accepted for inclusion in Journal Articles by an authorized administrator of Donald and Barbara Zucker School of Medicine Academic Works. For more information, please contact academicworks@hofstra.edu.
Abnormal Involuntary Movement Scale in Tardive Dyskinesia: Minimal Clinically Important Difference

Mark Stacy, MD,1* Martha Sajatovic, MD,2,3 John M. Kane, MD,4,5 Andrew J. Cutler, MD,6 Grace S. Liang, MD,7 Christopher F. O’Brien, MD,7 and Christoph U. Correll, MD4,5,8

1Brody School of Medicine, East Carolina University, Greenville, North Carolina, USA
2Case Western Reserve University School of Medicine, Cleveland, Ohio, USA
3University Hospitals Cleveland Medical Center, Cleveland, Ohio, USA
4The Zucker Hillside Hospital, Glen Oaks, New York, USA
5The Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Hempstead, New York, USA
6SUNY Upstate Medical University, Syracuse, NY, USA
7Neurocrine Biosciences, Inc., San Diego, California, USA
8Charité Universitätsmedizin, Department of Child and Adolescent Psychiatry, Berlin, Germany

ABSTRACT: Background: A minimal clinically important difference has not been established for the Abnormal Involuntary Movement Scale in patients with tardive dyskinesia. Valbenazine is a vesicular monoamine transporter 2 inhibitor approved for the treatment of tardive dyskinesias in adults. Efficacy in randomized, double-blind, placebo-controlled trials was defined as the change from baseline in Abnormal Involuntary Movement Scale total score (sum of items 1-7).

Objectives: To estimate an minimal clinically important difference for the Abnormal Involuntary Movement Scale using valbenazine trial data and an anchor-based method.

Methods: Data were pooled from three 6-week double-blind, placebo-controlled trials: KINECT (NCT01688037), KINECT 2 (NCT01733121), and KINECT 3 (NCT02274558). Valbenazine doses were pooled for analyses as follows: "low dose," which includes 40 or 50 mg/day; and "high dose," which includes 75 or 80 mg/day. Mean changes from baseline in Abnormal Involuntary Movement Scale total score were analyzed in all participants (valbenazine- and placebo-treated) with a Clinical Global Impression of Change-Tardive Dyskinesia or Patient Global Impression of Change score of 1 (very much improved) to 3 (minimally improved).

Results: The least squares mean improvement from baseline to week 6 in Abnormal Involuntary Movement Scale total score was significantly greater with valbenazine (low dose: −2.4; high dose: −3.2; both, P < 0.001) versus placebo (−0.7). An minimal clinically important difference of 2 points was estimated based on least squares mean changes in Abnormal Involuntary Movement Scale total score in participants with a Clinical Global Impression of Change-Tardive Dyskinesia score ≤3 at week 6 (mean change: −2.2; median change: −2) or Patient Global Impression of Change score ≤3 at week 6 (mean change: −2.0; median change: −2).

Conclusions: Results from an anchor-based method indicate that a 2-point decrease in Abnormal Involuntary Movement Scale total score may be considered clinically important. © 2019 The Authors. Movement Disorders published by Wiley Periodicals, Inc. on behalf of International Parkinson and Movement Disorder Society.

Key Words: AIMS; clinical trial; MCID; tardive dyskinesia; valbenazine

Correction added on July 17, 2019, after first online publication: The images for Figures 3 and 4 have been revised.

*The copyright line for this article was changed on July 17, 2019, after original online publication

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

*Correspondence to: Dr. Mark Stacy, Brody School of Medicine, East Carolina University, 600 Moye Boulevard, Greenville, NC 27834, USA; E-mail: stacyma17@ecu.edu

Funding agency: The study and post-hoc analyses were supported by Neurocrine Biosciences, Inc.

Relevant conflicts of interest/financial disclosures: Drs. Stacy, Sajatovic, Kane, Cutler, and Correll have served as advisors/consultants to Neurocrine Biosciences, Inc. Dr. Liang is a full-time employee and Dr. O’Brien is a former employee of Neurocrine Biosciences, Inc.; both are shareholders in the company.

Full financial disclosures and author roles may be found in the online version of this article.

Received: 21 August 2018; Revised: 20 May 2019; Accepted: 27 May 2019

Published online 24 June 2019 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.27769

Movement Disorders, Vol. 34, No. 8, 2019 1203
Tardive dyskinesia (TD) is a hyperkinetic movement disorder that is associated with exposure to an antipsychotic or other dopamine receptor blocking agent (DRBA), such as metoclopramide. Despite the development and widespread use of second-generation antipsychotics, TD remains a relevant potential risk of DRBAs. Two medications, valbenazine and deutetrabenazine, are now approved for the treatment of TD in adults. The placebo-controlled clinical trials of these drugs had some differences in study design (e.g., treatment duration, eligibility criteria), but both used the Abnormal Involuntary Movement Scale (AIMS) to measure the presence, severity, and changes in TD. Results from the valbenazine and deutetrabenazine trials showed that both compounds had measurable and statistically significant benefits as assessed by mean changes in AIMS total score.

Although the mean change in AIMS total score is currently standard for evaluating efficacy in clinical trials, the implications of this outcome for everyday practice are unclear. Along with other analytical approaches (e.g., Cohen’s effect size, number needed to treat [NNT]), one way to estimate clinical relevance of recent TD trial results would be to identify a minimal clinically important difference (MCID) for the AIMS total score. Two approaches are generally used to estimate MCIDs: distribution-based, which relies on a standard deviation (SD) or standard error of the measurement; and anchor-based, which uses an external measure (e.g., 7-point global assessment scale) as an independent criterion for improvement. An MCID for the AIMS has not been established in patients with TD, possibly because of the lack of large, well-controlled, and prospectively designed studies in this population. With the completion of three randomized controlled trials with valbenazine, a data set is now available that includes AIMS results for >350 study participants. Moreover, this data set is now available that includes AIMS data set is now available that includes AIMS results for >350 study participants. Moreover, this data set is now available that includes AIMS results for >350 study participants. Moreover, this data set is now available that includes AIMS results for >350 study participants. Moreover, this data set is now available that includes AIMS results for >350 study participants. Moreover, this data set is now available that includes AIMS results for >350 study participants.

Statistical Analyses

All analyses were conducted in the pooled intent-to-treat (ITT) population, defined as participants who received

![Figure 1: Valbenazine studies. Valbenazine dose groups were pooled as follows: “low dose” (50 mg/day and 100/50 mg/day [KINET], 50 mg/day [KINET 2], and 40/40 mg [KINET 3]); and “high dose” (75 mg/day [KINET 2], 80 mg/day [KINET 3]). Participants who received valbenazine 25 mg/day in KINET 2 were not included in the pooled analyses. Participants randomized to valbenazine 80 mg/day in KINET 3 received 40 mg/day for 1 week.](image-url)
TABLE 1. Baseline characteristics (pooled ITT population)

|                                | Placebo (n = 158) | Valbenazine Low Dose* (n = 114) | Valbenazine High Dose** (n = 101) |
|--------------------------------|------------------|---------------------------------|-----------------------------------|
| Age, mean (SD), years          | 55.8 (10.1)      | 54.9 (9.1)                      | 56.2 (10.4)                       |
| Male, n (%)                    | 89 (56.3)        | 72 (63.2)                       | 55 (54.5)                         |
| Race, n (%)                    |                  |                                 |                                   |
| White                           | 86 (54.4)        | 64 (56.1)                       | 62 (61.4)                         |
| Black or African American       | 63 (39.9)        | 44 (38.6)                       | 36 (35.6)                         |
| Psychiatric diagnosis group, n (%) |            |                                 |                                   |
| Schizophrenia/schizoaffective disorder | 116 (73.4)  | 90 (78.9)                       | 61 (60.4)                         |
| Mood disorder                   | 42 (26.6)        | 24 (21.1)                       | 40 (39.6)                         |
| Concomitant use of antipsychotics, n (%) |        |                                 |                                   |
| Any antipsychotic               | 130 (82.3)       | 102 (89.5)                      | 77 (76.2)                         |
| Atypical only                   | 102 (78.5)       | 77 (75.5)                       | 63 (61.8)                         |
| Typical only or both            | 28 (18.1)        | 25 (24.5)                       | 14 (13.8)                         |
| BPRS score at screening, mean (SD) | 30.5 (7.6)    | 31.6 (7.9)                      | 28.9 (6.8)                        |
| AIMS total score at baseline    |                  |                                 |                                   |
| Mean (SD)                       | 8.9 (4.4)        | 9.0 (4.2)                       | 9.5 (3.6)                         |
| Median (minimum, maximum)       | 8 (1, 26)        | 9 (0, 20)                       | 9 (3, 20)                         |

*Includes participants who received valbenazine 40 or 50 mg/day.
**Includes participants who received valbenazine 75 or 80 mg/day.
Similar to results for CGI-TD score ≤3, analyses based on PGIC score ≤3 yielded an MCID estimation of 2 points, with a median 20% improvement from baseline (Supporting Information Appendix; Supporting Information Table S1). Analyses based on PGIC score ≤2 also yielded an MCID estimation of 2 points (compared to 3 points for CGI-TD score ≤3), with a median 30% total score improvement from baseline.

**Discussion**

Although the AIMS total score is the current standard for determining efficacy in TD clinical trials, translating this outcome into clinical practice can be challenging.\textsuperscript{16} To address that challenge, the TD Workshop participants discussed different ways to analyze AIMS data and identified the MCID as one possible approach.\textsuperscript{17} Based on both clinician- and patient-rated anchors of minimal improvement (CGI-TD and PGIC score ≤3 at week 6), mean and median changes in AIMS total score (sum of items 1-7) suggested an MCID of 2 points in adults with TD. Analyses based on more rigorous definitions of global improvement (CGI-TD and PGIC score ≤2 at week 6) suggested a clinician-based MCID of 3 points and a patient-based MCID of 2 to 3 points. Clinically, these proposed MCIDs may be useful for interpreting the effects of treatment on TD. However, it may be worth noting that the MCID of 2 points is consistent with the distribution-based approach that uses 0.5 times the baseline SD as a threshold for clinically meaningful change.\textsuperscript{18} In the pooled data set, the SD of the mean AIMS total score at baseline in all participants was 4.2, which would correspond to an MCID of 2 points.

The current results were consistent with preliminary MCID analyses, which only included CGI-TD anchors.\textsuperscript{19,20} PGIC anchors were added to the current analyses to address the need for more patient-reported outcomes in TD studies. Given that patients with TD can be unaware of their movements,\textsuperscript{21} these PGIC-based results should be interpreted with some caution. However, consistent with the CGI-TD results, MCID estimates based on patient-reported improvements suggest that a 2- to 3-point decrease in AIMS total score may be considered clinically meaningful. It should also be noted that both anchor-based methods (CGI-TD and PGIC) included placebo responders to lessen the risk of the MCID being specific to valbenazine treatment. Additional MCID analyses based on data from other TD clinical trials (e.g., deutetrabenazine) would help to further

**FIG. 2.** AIMS total score mean change from baseline to week 6. ***$P < 0.001$ versus placebo.

**FIG. 3.** CGI-TD response at week 6. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$ versus placebo. CI, confidence interval.
establish whether an AIMS MCID of 2 to 3 points is applicable to different TD therapies.

An additional goal of the current analysis was to include percentage-based MCIDs for the AIMS total score. Participants with a CGI-TD or PGIC score ≤ 3 at week 6 had approximately 20% improvement from baseline in AIMS total score. Participants with a CGI-TD or PGIC score ≤ 2 had approximately 30% to 40% improvement from baseline in AIMS total score. These percentages are consistent with earlier TD studies that defined response as a ≥ 30% improvement in AIMS total score.16 They are also consistent with results from the companion piece to this article, which presents a full range of AIMS total score responses (≥ 10% to ≥ 90% improvement from baseline to week 6).17 In that analysis, the percentage of participants who achieved a ≥ 20%, ≥ 30%, or ≥ 40% AIMS total score response was significantly higher with valbenazine high dose versus placebo. These results were clinically meaningful, as indicated by ORs for response (OR ≥ 4 for valbenazine vs. placebo) and NNTs (of 3 or 4). In the valbenazine clinical trials, AIMS response was defined a priori as ≥ 50% total score improvement,9,10 which is more stringent than the 20% to 40% MCID-based results in the current analysis. Therefore, a greater percentage of patients experienced a clinical benefit in the valbenazine clinical trials than the published ≥ 50% response analyses would imply.

A number of limitations should be noted. First, all analyses were conducted post hoc. None of the valbenazine trials were designed for estimation of an MCID, and the pooled valbenazine dose groups included participants who received slightly different low doses (40 and 50 mg/day) and high doses (75 and 80 mg/day). Second, results of the analyses may not be generalizable to all patients with TD. The trials primarily included psychiatric patients who were exposed to antipsychotic medications, and MCIDs may be different in nonpsychiatric patients who were exposed to an antiemetic (e.g., metoclopramide) or other DRBA. Study participants were also required to be psychiatrically stable, which may not always be true in real-life settings. In addition, participants in the valbenazine studies were required to have moderate or severe TD based on the

![FIG. 4. Estimation of AIMS MCID. Based on all participants who met CGI-TD response criteria regardless of treatment (valbenazine or placebo).](image-url)
participants agreed that the AIMS MCID can be an
clinicians or study participants.
recall bias, and inter-rater agreement was not tested for
times the SD). In addition, anchors can be susceptible to
tent with a commonly used distribution method (i.e., 0.5
Moreover, MCIDs from both anchor types were consis-
tances were conducted without considering the AIMS total
score at baseline. Nor did they consider AIMS items scores
variability of dyskinetic movements.22,23 The MCID ana-
lyses were conducted without considering the AIMS total
score may be considered an MCID if minimal
improvement is the treatment goal; a 3-point decrease
be the MCID if more robust improvement is desired.
Much more research is needed to understand the impact
of TD on patients and caregivers, including the bench-
marks of physical, functional, and social improvements
that constitute a truly meaningful clinical difference.

Limitations of the AIMS itself should also be consid-
ered. Given that the AIMS total score is the current
“gold standard” for evaluating efficacy in TD clinical
trials, determining an MCID based on this measure is a
reasonable endeavor. However, the AIMS does not cap-
ture the social and functional deficits associated with
TD. In addition, one-time or episodic complications
related to TD, such as a fall related to gait problems,
are not adequately captured by the AIMS. Improve-
ments in these domains must be considered along with
dyskinetic movements when determining whether a
patient is experiencing clinically meaningful improve-
ments. Methodologies for administering and scoring the
AIMS should also be considered. The proposed MCIDs
presented in this report are based on AIMS evaluations
that were scored by consensus between two central video
raters (movement disorder specialists) who were blinded
to treatment and study visit. In clinical settings, the AIMS
is administered and assessed in real time by a physician or
other qualified professional who knows what the patient
is taking and how long he or she has been treated.
Therefore, an MCID based on clinical trial data, as investigat-
in this report, should be considered as more of a guideline
(rather than an imperative) for everyday practice. Given
that the analyses in this report are limited to valbenazine
data, they may not be generalizable to all AIMS results,
including those that have been reported in other TD clinici-
tal trials (e.g., deutetrabenazine). Applying the proposed
MCIDs from this report to other TD trials should also be
done with caution given that differences in study design
(e.g., double-blind vs. open-label, treatment duration, eli-
ghility criteria, and allowance of concomitant medi-
cations) may affect treatment outcomes.

Finally, as previously published,15 the limitations of
anchor-based methods should be mentioned. First, differ-
ent anchors may result in different MCIDs, although the
current analysis showed consistency between clinician-
based (CGI-TD) and patient-based (PGIC) anchors.
Moreover, MCIDs from both anchor types were consist-
tent with a commonly used distribution method (i.e., 0.5
times the SD). In addition, anchors can be susceptible to
recall bias, and inter-rater agreement was not tested for
clinicians or study participants.

Despite these various limitations, the TD Workshop
participants agreed that the AIMS MCID can be an
important advancement for clinicians who treat patients
with TD. Taken in conjunction with other types of ana-
lyses (e.g., placebo-corrected mean change, effect size,
treatment response, and NNT), or even added prospecti-
tively to statistical analysis plans, the MCID might help
translate trial data into clinically meaningful information.
Based on both clinician- and patient-rated anchors, the
results of this analysis suggest that a 2-point decrease in
AIMS total score may be considered an MCID if minimal
improvement is the treatment goal; a 3-point decrease
could be below the MCID if more robust improvement is desired.
Much more research is needed to understand the impact
of TD on patients and caregivers, including the benchmarks
of physical, functional, and social improvements that
constitute a truly meaningful clinical difference. ■

Acknowledgments: The manuscript was critically reviewed by Scott
Siegert, PharmD, Khodayar Farahmand, PharmD, and Charles Yonan,
PharmD, all of whom are full-time employees of Neurocrine Biosciences,
Inc. (San Diego, CA). Medical writing and editorial support were provided
by Mildred Bahn, MA, at Prescott Medical Communications Group
(Chicago, IL) with support from Neurocrine Biosciences, Inc.
an OMERACT Rasch Working Group systematic review and critique. J Rheumatol 2016;43:194-202.

14. Jayadevappa R, Cook R, Chhatre S. Minimal important difference to infer changes in health-related quality of life—a systematic review. J Clin Epidemiol 2017;89:188-198.

15. Rai SK, Yazdany J, Fortin PR, Avina-Zubieta JA. Approaches for estimating minimal clinically important differences in systemic lupus erythematosus. Arthritis Res Ther 2015;17:143.

16. Kane JM, Correll CU, Nierenberg AA, Caroff SN, Sajatovic M. Revisiting the Abnormal Involuntary Movement Scale: proceedings from the Tardive Dyskinesia Workshop. J Clin Psychiatry 2018;79:17cs11959.

17. Correll CU, Cutler AJ, Kane JM, McEvoy JP, Liang GS, O’Brien CF. Characterizing treatment effects of valbenazine for tardive dyskinesia: additional results from the KINECT 3 study. J Clin Psychiatry 2018;80:18m12278.

18. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 2003;41:582-592.

19. Sajatovic M, Cutler A, Farahmand K, Burke J, Siegert S, Liang G. Estimation of an MCID for AIMS total score change in tardive dyskinesia. Presented at the 30th Annual Psych Congress, New Orleans, LA, September 16–18, 2017. Available at: https://www.psychcongress.com/posters/estimation-mcid-aims-total-score-change-tardive-dyskinesia. Accessed January 5, 2018.

20. Stacy M, Kurlan R, Burke J, Siegert S, Liang G, O’Brien C. An MCID for AIMS dyskinesia total score change in subjects with tardive dyskinesia [abstract #405]. Mov Disord 2017;32:S158.

21. Emsley R, Niehaus DJ, Oosthuizen PP, Koen L, Chiliza B, Fincham D. Subjective awareness of tardive dyskinesia and insight in schizophrenia. Eur Psychiatry 2011;26:293-296.

22. Hyde TM, Egan MF, Brown RJ, Weinberger DR, Kleinman JE. Diurnal variation in tardive dyskinesia. Psychiatry Res 1995;56:53-57.

23. Stanilla JK, Buchel C, Alarcon J, de Leon J, Simpson GM. Diurnal and weekly variation of tardive dyskinesia measured by digital image processing. Psychopharmacology 1996;124:373-376.

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site.