Short Communication – Parkinsonism and Related Disorders

Minimal Clinically Important Change in the Toronto Western Spasmodic Torticollis Rating Scale

Alberto J. Espay, MD, MSc[a]; Richard Trosch, MD[b]; Gustavo Suarez, MD[c]; Jonathan Johnson, MS[d]; Dominic Marchese, PharmD[c1]; Cynthia Comella, MD[e]

a. James J and Joan A Gardner Center for Parkinson's disease and Movement Disorders, Department of Neurology, University of Cincinnati, Cincinnati, Ohio, USA.

b. Oakland University William Beaumont School of Medicine, Farmington Hills, Michigan, USA.

c. Ipsen Biopharmaceuticals, Basking Ridge, New Jersey, USA.

d. OptumInsight, Eden Prairie, Minnesota, USA.

e. Department of Neurology, Rush University Medical Center, Chicago, Illinois.

[1]Present address: Westerville, Ohio, USA

Abstract count: 155; **Text word count**: 1774. **Title character count**: 94

Corresponding author

Alberto J. Espay, MD, MSc

University of Cincinnati, College of Medicine

260 Stetson Street, Suite 2300

PO Box 670525

Cincinnati, OH 45267-0525 United States

Phone: (513) 558-4050

Fax: (513) 558-7015

alberto.espay@uc.edu

Running title: MCIC for TWSTRS

Financial disclosures: AJE, RT, & CC report fees for consultancy and speaker fees from Ipsen Biopharmaceuticals, and participated in the ANCHOR-CD trial. DM was employed by Ipsen Biopharmaceuticals at the time of the study. JJ is employed by OptumInsight, which conducted the statistical analyses (funded by Ipsen Biopharmaceuticals). GS is employed by Ipsen Biopharmaceuticals.

Key Words: Cervical dystonia, Toronto Western Spasmodic Torticollis Rating Scale, minimal clinically important change, botulinum toxin, abobotulinumtoxinA.

Conflict of interest: AJE and CC report fees for consultancy and speaker fees from Ipsen Biopharmaceuticals, and were investigators in the ANCHOR-CD trial. DM was employed by Ipsen Biopharmaceuticals at the time of the study and the MCIC analysis; he has no commercial or financial relationships that could be construed as a potential conflict of interest. JJ is employed by OptumInsight, which conducted the analyses (funded by Ipsen Biopharmaceuticals). GS is employed by Ipsen Biopharmaceuticals.

Study funding source: This work was sponsored by Ipsen Biopharmaceuticals**.**

ABSTRACT

Objectives: To characterize the minimal clinically important change (MCIC) after treatment in cervical dystonia patients using the Toronto Western Spasmodic Torticollis Rating Scale (TWSTRS).

Methods: Changes in the TWSTRS from an observational study of abobotulinumtoxinA in the routine management of cervical dystonia (NCT01314365) were analyzed using the Patient Global Impression of Change (PGIC) as anchor.

Results: For the overall population (N=304, baseline TWSTRS-Total score 43.4±19.4), the MCIC for the TWSTRS Total score was -11.9 (95%CI: -13.9, -10.0; p<0.0001). However, thresholds ranged from -3.2 to -18.0 dependent on baseline severity. TWSTRS-Total scores improved linearly by 3 points for every one-point PGIC increase. There was similar linearity between the graded PGIC categories and TWSTRS subscale scores (severity, disability, and pain).

Conclusions: A 3-point change is the minimal clinically important change after treatment using TWSTRS as endpoint with higher cutoffs for greater baseline disease severity. For an average trial population (TWSTRS-total: 40-45), a 12-point decrease is clinically meaningful.

Introduction

Cervical dystonia (CD), the most common form of focal dystonia, is characterized by sustained abnormal muscle contractions causing twisting and turning of the head and neck leading to disabling and sometimes painful postures. Depending on the muscles involved, patients may exhibit torticollis (rotation), laterocollis (tilting), anterocollis (flexion), retrocollis (extension), or combinations thereof Botulinum neurotoxin (BoNT) is recommended as first line treatment for CD.

All trials that led to the approval of currently available BoNTs have utilized the Toronto Western Spasmodic Torticollis Rating Scale (TWSRS) as the primary efficacy measure. The TWSTRS is a comprehensive scale designed to assess objective physical (severity subscale) and subjective findings (disability and pain subscales) [1]. The reliability and validity of the TWSTRS have been well established, and the severity scores rated by physicians positively correlate with patient's self-reported improvement in disability and pain after treatment with BoNT [2]. However, the minimal clinically important change (MCIC) has not yet been established. The MCIC is defined as the smallest change or difference in scores of a measure perceived by patients as beneficial or harmful [3] Ascertaining this number can help assess whether a statistically significant treatment effect is sufficiently large enough to be interpreted as clinically significant and could help personalize treatment in clinical practice.

ANCHOR-CD was a prospective, open-label, observational registry designed to evaluate the efficacy and safety of a BoNT-A formulation (abobotulinumtoxinA, Dysport® Ipsen Biopharmaceuticals) for the long-term treatment (≈1 year) of adult idiopathic CD in routine clinical practice in the United States [4]. All patients enrolled into this registry underwent regular comprehensive assessments, including the clinician-rated TWSTRS scale and the patient-rated global impression of change (PGIC). Thus, the study design provides a rich dataset upon which the MCIC can be examined by determining the relationship between the patient perception of treatment response and the

corresponding reduction in TWSTRS scores, with the main objective of calculating the MCIC on the TWSTRS-Total score and subscores according to patients' self-reported PGIC scores.

Methods

ANCHOR-CD study design

ANCHOR-CD was a prospective, open-label, observational study (NCT01314365) of abobotulinumtoxinA in the routine treatment of CD, described in detail elsewhere [4]. Briefly, adult (aged ≥18 years) patients diagnosed with primary idiopathic CD were enrolled at 41 US sites. Patients could be BoNT-naïve or previously treated with BoNT if ≥12 weeks had elapsed since the last injection. The decision to prescribe abobotulinumtoxinA was to be made before and independently from the decision to enroll the patient in the registry.

AbobotulinumtoxinA was administered over 4 treatment cycles. The sites, number of injections, and doses were determined by each investigator in accordance with their standards of clinical practice and in line with the United States Prescribing Information for Dysport®. In-office assessments of TWSTRS were made at baseline and at Week 4 following the injection. The PGIC was assessed at Week 4 using a 7-point Likert scale, ranging from +3 (very much improved) to -3 (very much worse).

Analyses

Data from the first treatment cycle were used. *Post-hoc* analyses were performed in the modified intent to treat (mITT) population which included all treated patients who had a documented TWSTRS severity score at Cycle 1 baseline, and at Week 4; and, a documented PGIC response value at Week 4. Due to sample size imbalances in the worsening categories, we modified the PGIC categories such that scores of -1 to -3 were combined into a "worsened" group resulting in 5 categories: "very much improved" (+3), "much improved" (+2), "minimally improved" (+1), "no change" (0) and "worsened" (-1 to -3). Mean±SD TWSTRS-Total score changes from baseline to the Week 4 assessments and their

95% confidence intervals (CIs) in each of the different categories of the PGIC were calculated and were compared with 'no change' using paired t-tests.

Ordinary least squares regression analyses were used to measure the slope of TWSTRS-Total scores across modified PGIC categories. The change in TWSTRS-Total scores served as the dependent variable and PGIC as the independent variable (assuming a linear relationship). The MCIC was calculated using patients' report of a minimal improvement (+1) in their PGIC, for the total population and for quartiles of baseline TWSTRS-Total score (0-28.5, 28.5-41.0, 41.0-52.0, 52.0+).

Results

As described previously, 347 of 350 enrolled patients (86 men, 261 women) received at least one dose of abobotulinumtoxinA treatment (ITT population) [4]. Of these, 304 patients had documented TWSTRS severity and PGIC scores at baseline and at the end of Cycle 1- Week 4 (mITT population). For the mITT population, mean TWSTRS-Total score was 43.4±19.4 at baseline; most patients (67.2%) had a complex form of CD, with 25.6% reported as exhibiting pure torticollis. Overall, 72.6% of patients had received previous BoNT treatment for CD.

In the mITT analysis, the average mean decrease (from baseline to Week 4) in TWSTRS-Total scores for PGIC categories ranged from -5.89 ($p<0.01$) in the "worsened" group to -19.63 in the "very much improved" group (**Figure 1a**). Assuming an equal distance between the modified PGIC categories, the relationship between TWSTRS-Total scores and improved/no change PGIC categories was found to be linear in the mITT population. The least squares regression slope for PGIC categories was 2.9±0.51 ($p<0.0001$), which suggests that TWSTRS-Total scores improved by 2.9 points for every 1 point increase in PGIC rating. There was similar linearity between the graded PGIC categories and baseline Total and subscale TWSTRS scores (severity, disability, and pain), regardless of treatment status (BoNT previously treated and naïve populations) (**Figure 1b**).

[figure 1 about here]

Overall, for the total mITT population (mean baseline TWSTRS-Total score of 43.4±19.4), the mean (95% CI) change in TWSTRS-Total scores in those patients who rated minimally improved (+1) on the PGIC at Week 4 was -11.9 (-13.9, -10.0) ($p<0.0001$). Further analysis showed a range of MCIC dependent on baseline severity based on TWSTRS-Total scores, when assessing mean change in TWSTRS based on baseline quartiles (**Table 1**). The mean (95% CI) MCIC ranged from -3.18 (-7.6, 1.3) for the lowest baseline TWSTRS quartile to -18.0 (-20.3, -15.7) for the highest TWSTRS quartile.

[Table 1 about here]

Discussion

To our knowledge, this is the first study to characterize a MCIC for the TWSTRS. In our analyses, the MCIC ranged from -3.2 for patients with a baseline TWSTRS-Total score of ≤28.5 up to -18.0 for patients with a baseline TWSTRS-Total score of >52. For an average trial population with a baseline TWSTRS-total score between 40 and 45, we suggest that a decrease of approximately 12 points should be considered clinically meaningful. Our findings also show that the patient perception of change is linearly associated with the TWSTRS scale. Furthermore, TWSTRS-Total scores improved by 3 points for every 1-point increase in PGIC, regardless of disease severity or prior treatment status.

Previous clinical trials have predefined 'response' to BoNT treatment as a reduction of at least 30% from baseline TWSTRS-Total scores [5, 6]. While such definitions have been based on clinical experience rather than scale clinimetrics, our results showing the relationship between baseline score and change from baseline support the validity of this approach. Furthermore, our results provide the first quantitative TWSTRS cutoffs for assessment of response based on quartiles of

baseline scores, which may better assist with future clinical trial design and interpretation of the clinical relevance of any therapeutic response. For example, in the recent randomized placebo-controlled trial of abobotulinumtoxinA compared to its new liquid formulation, a treatment difference of 1.5 points between the two active groups may not be clinically relevant (whereas the differences of 14.0 points and 12.5 points for the respective active treatments vs. placebo are above the MCIC) [7]. The apparent linear relationship between TWSTRS-Total scores and PGIC categories is also of interest for future clinical trial design and appears to be supported by data from another registry of abobotulinumtoxinA, INTEREST-IN CD1, in which 73.6% of patients treated after one cycle of abobotulinumtoxinA were reported to have ≥25% improvement in TWSTRS-Total scores and 69.8% of patients self-reported an improvement in PGIC [8].

This analysis should be considered a first step in understanding a MCIC for TWSTRS in CD. We show that there is a clear relationship between baseline severity and the magnitude of improvement required for a patient to appreciate a clinically important change. While it is tempting to develop a "one size fits all" linear model that would be generalizable to the full CD population, recent experiences with the Unified Parkinson Disease Rating Scale in Parkinson disease, have shown that the MCIC can vary dependent on many factors, including study methodology and the efficacy of the intervention [9-12]. Indeed, there is currently no consensus on the best approach for determining the MCIC, and experts recommend that estimation of MCIC for a specific measure should be based on multiple approaches and triangulation of methods [3]. This should also include use of other patient reported outcomes as anchor for the analyses. For example, ANCHOR-CD also included quality of life ratings using the CDIP-58, and treatment satisfaction using the Treatment Satisfaction Questionnaire for Medication [4]. We focused on PGIC ratings for this first analysis because it was collected as part of the primary composite endpoint.

Strengths of our analyses include the fact that they are based on the findings of a routine practice registry study that enrolled a broad range of patients with varying disease severities and clinical presentations. While this increases the generalizability of the findings, limitations of the study include the lack of a placebo control and inter-rater standardization (e.g. through TWSTRS training) that would normally be required for a randomized, controlled trial. Another strength, however, is that we used the patient self-rated impression of change and not the usual clinician rated measure, which is relevant because the fundamental question for meaningful outcomes of any therapeutic intervention is whether the patients themselves feel improved [9]. It also avoids the issue of potential rater bias where an investigator might rate the patient as having apparent improvement when the patient may have perceived little has changed. However, due to small sample sizes in the worsening PGIC categories, we collapsed these categories into a combined "worsened" category that has not been validated. In addition, for the relatively small subset of patients (n=20) in the lowest quartile group who suggested "minimal improvement" the range of change in TWSTRS includes 'worsening'. Here, it should be noted that these quartile analyses were performed to provide additional sensitivity around the MCIC estimated across the mITT population and to conceptually connect the patient's impression of *minimal* improvement to change on a clinical scale.

In summary, we present the first characterization of the minimal clinically important difference for the TWSTRS. Such estimates are important for regulatory agencies, clinicians, and patients in properly assessing the clinical relevance of treatment responses.

Acknowledgements

The ANCHOR-CD study was supported by Ipsen Biopharmaceuticals. We thank the study investigators and patients who participated in the original trial. Medical writing assistance for this report was provided by Anita Chadha-Patel of ACP Clinical Communications Ltd and was funded by Ipsen Biopharmaceuticals.

JJ and DM take full responsibility for the integrity of the data and the accuracy of the statistical analyses.

Financial Disclosures

AJE has received grant support from the NIH, Great Lakes Neurotechnologies and the Michael J Fox Foundation; personal compensation as a consultant/scientific advisory board member for Abbvie, TEVA, Impax, Acadia, Cynapsus/Sunovion, Lundbeck, and USWorldMeds; publishing royalties from Lippincott Williams & Wilkins, Cambridge University Press, and Springer; and honoraria from Abbvie, US WorldMeds, Lundbeck, Acadia, the American Academy of Neurology, and the Movement Disorders Society.

DM was employed by Ipsen Biopharmaceuticals at the time of the study and the MCIC analysis. He has no commercial or financial relationships that could be construed as a potential conflict of interest.

RT reports consultancy for Ipsen Pharma.

JJ is employed by OptumInsight, which conducted the analyses (funded by Ipsen Biopharmaceuticals).

GS is employed by Ipsen Biopharmaceuticals

CC serves on the editorial board of Clinical Neuropharmacology, Sleep Medicine and Continuum. She receives research support from the NIH R01NS074343, U54NS065701, Dystonia Medical Research Foundation, Merz Pharmaceutical, Revance Therapeutic, Retrophin and Acorda Therapeutic. She receives compensation/honoraria for services as a consultant or an advisory committee member: Acorda Therapeutics, Allergan, Inc; Lundbeck Ltd.; Medtronic Inc.; Merz Pharmaceuticals; Acadia Pharmaceuticals; Jazz Pharmaceuticals, Neurocrine Biosciences Inc., Revance Therapeutic. She receives royalties from Cambridge, Wolters Kluwer. She receives research support from the Parkinson's Disease Foundation.

References

1. Comella CL, Stebbins GT, Goetz CG, Chmura TA, Bressman SB, Lang AE. Teaching tape for the motor section of the Toronto Western Spasmodic Torticollis Scale. Mov. Disord. 12 (1997) 570-575.

2. Jost WH, Hefter H, Stenner A, Reichel G. Rating scales for cervical dystonia: a critical evaluation of tools for outcome assessment of botulinum toxin therapy. J. Neural Transm. 120 (2013) 487-496.

3. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J. Clin. Epidemiol. 61(2) (2008) 102-109.

4. Trosch RM, Espay AJ, Truong D, Gil R, Singer C, LeWitt PA, Lew MF, Tagliati M, Adler CH, Chen JJ, Marchese D, Comella CL. Multicenter observational study of abobotulinumtoxinA neurotoxin in cervical dystonia: The ANCHOR-CD registry. J. Neurol. Sci. 376 (2017) 84-90.

5. Truong D, Duane DD, Jankovic J, Singer C, Seeberger LC, Comella CL, Lew MF, Rodnitzky RL, Danisi FO, Sutton JP, Charles PD, Hauser RA, Sheean GL. Efficacy and safety of botulinum type A toxin (Dysport) in cervical dystonia: results of the first US randomized, double-blind, placebo-controlled study. Mov. Disord. 2005;20(7):783-791.

6. Truong D, Brodsky M, Lew M, Brashear A, Jankovic J, Molho E, Orlova O, Timerbaeva S for the Global Dysport Cervical Dystonia Study Group. Long-term efficacy and safety of botulinum toxin type A (Dysport) in cervical dystonia. Parkinsonism Relat. Disord. 2010;16(5):316-323.

7. Poewe W, Burbaud P, Castelnovo G, Jost WH, Ceballos-Baumann AO, Banach M, Potulska-Chromik A, Ferreira JJ, Bihari K, Ehler E, Bares M, Dzyak LA, Belova AN, Pham E, Liu WJ, Picaut P. Efficacy and safety of abobotulinumtoxinA liquid formulation in cervical dystonia: A randomized-controlled trial. Mov. Disord. 2016;31(11):1649-1657.

8. Misra VP, Ehler E, Zakine B, Maisonobe P, Simonetta-Moreau M. Factors influencing response to Botulinum toxin type A in patients with idiopathic cervical dystonia: results from an international observational study. BMJ Open. 2012;2(3).

9. Hauser RA, Gordon MF, Mizuno Y, Poewe W, Barone P, Schapira AH, Rascol O, Debieuvre C, Fräßdorf M. Minimal clinically important difference in Parkinson's disease as assessed in pivotal trials of pramipexole extended release. Parkinsons Dis. (2014) 467131.

10. Schrag A, Sampaio C, Counsell N, Poewe W. Minimal clinically important change on the unified Parkinson's disease rating scale. Mov. Disord. 21(8) (2006) 1200-1207.

11. Shulman LM, Gruber-Baldini AL, Anderson KE, Fishman PS, Reich SG, Weiner WJ. The clinically important difference on the unified Parkinson's disease rating scale. Arch. Neurol. 67(1) (2010) 64-70.

12. Hauser RA, Auinger P. Determination of minimal clinically important change in early and advanced Parkinson's disease. Mov. Disord. 26(5) (2011) 813-818.

Table legends:

Figure 1: Relationship between PGIC and changes in: (a) TWSTRS-Total scores and (b) TWSTRS Subscale scores at 4 weeks post-abobotulinumtoxinA treatment

Table 1: Minimal Clinically Important Change in TWSTRS at 4 weeks post-abobotulinumtoxinA treatment according to baseline quartiles of total TWSTRS scores

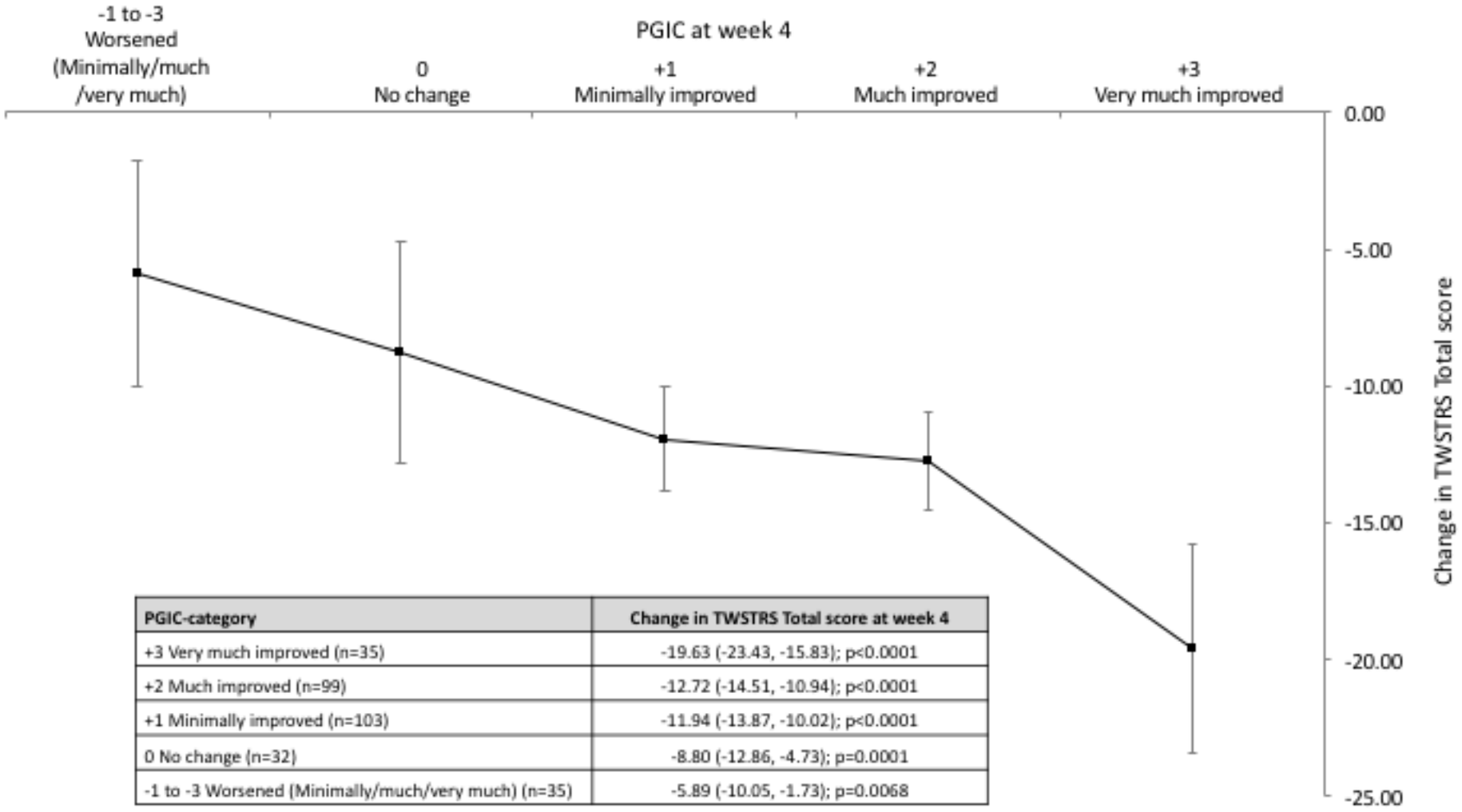
Table(s)

Table 1. Minimal Clinically Important Change in TWSTRS at 4 weeks post-abobotulinumtoxinA treatment according to baseline quartiles of total TWSTRS scores

Baseline TWSTRS Quartiles	Mean (95%CI) Change in TWSTRS-Total score at Week 4	Mean (95%CI) Change in TWSTRS-Total score in patients with PGIC score +1 at Week 4 (MCIC)
0 - 28.5	-5.14 (-7.48, -2.79) N=48	-3.18 (-7.60, 1.25) N=20
28.5 - 41.0	-11.49 (-14.26, -8.72) N=49	-9.60 (-14.33, -4.86) N=21
41.0 - 52.0	-14.01 (-16.26, -11.75) N=57	-13.79 (-16.63, -10.95) N=36
>52.0	-18.37 (-20.13, -16.61) N=48	-18.02 (-20.32, -15.71) N=26

Figure(s)

Click here to download high resolution image



PGIC-category	Change in TWSTRS Total score at week 4
+3 Very much improved (n=35)	-19.63 (-23.43, -15.83); p<0.0001
+2 Much improved (n=99)	-12.72 (-14.51, -10.94); p<0.0001
+1 Minimally improved (n=103)	-11.94 (-13.87, -10.02); p<0.0001
0 No change (n=32)	-8.80 (-12.86, -4.73); p=0.0001
-1 to -3 Worsened (Minimally/much/very much) (n=35)	-5.89 (-10.05, -1.73); p=0.0068