Minimal Clinically Important Differences for Burke-Fahn-Marsden Dystonia Rating Scale and 36-Item Short-Form Health Survey

Dávid Pintér, MD,1* József Janszky, MD, DSc,1,2 and Norbert Kovács, MD, DSc1,2

1Department of Neurology, Medical School, University of Pécs, Pécs, Hungary
2MTA-PTE Clinical Neuroscience MR Research Group, Pécs, Hungary

ABSTRACT: Background: Although an increasing number of trials are reported on the treatment of generalized or segmental isolated dystonia, the minimal clinically important difference thresholds for the most frequently reported outcome measures are still undetermined. Objectives: To estimate the minimal clinically important difference for the Burke-Fahn-Marsden Dystonia Rating Scale and the 36-Item Short-Form Health Survey in generalized or segmental dystonia. Methods: A total of 898 paired examinations of 198 consecutive patients, aged >18 years, with idiopathic and inherited (torsin family 1 member A positive) segmental and generalized isolated dystonia were analyzed. To calculate the minimal clinically important difference thresholds, both anchor- and distribution-based methods were used simultaneously. Results: Any improvement >16.6% or worsening larger than 21.5% on the Burke-Fahn-Marsden Dystonia Rating Scale indicates a minimal, yet clinically relevant, change. Threshold values for the Burke-Fahn-Marsden Dystonia Disability Scale were 0.5 points for both decline and improvement. Cut-off scores for the Physical Component Summary, the Mental Component Summary, and the Global (Total or Overall) Score of the 36-Item Short-Form Health Survey were 5.5 and 5.5, 6.5 and 7.5, and 7.5 and 8.5 points for clinically meaningful improvement and deterioration, respectively. Conclusions: The minimal clinically important difference represents the smallest change in an outcome measure that is meaningful to patients. Our estimates for the Burke-Fahn-Marsden Dystonia Rating Scale and the 36-Item Short-Form Health Survey may allow more reliable judgment of the clinical relevance of different treatments for segmental and generalized isolated dystonia.

Key Words: 36-item Short-Form Health Survey; Burke-Fahn-Marsden Dystonia Rating Scale; deep brain stimulation; dystonia; minimal clinically important difference

There is increasing research into health-related quality of life (HRQoL) in dystonia.1 Currently, available data strongly suggest that patients with dystonia generally experience lower levels of HRQoL than do healthy individuals.1 Several factors seem to determine the level of disability related to dystonia, including, but not limited to, the objective severity of dystonia and the presence of nonmotor symptoms, such as anxiety and depression, sleep disturbances, and pain.2 Given that treatments for dystonia aim to relieve these problems in a holistic manner, measuring the changes in HRQoL can be an adequate way to evaluate the effectiveness of therapeutic interventions. The Burke-Fahn-Marsden Dystonia Rating Scale (BFMD-RS; frequently reported also as the Burke-Fahn-Marsden Motor Scale), Burke-Fahn-Marsden Dystonia Disability Scale (BFMD-DS), and 36-Item Short Form Health Survey (SF-36) are among the most frequently reported outcome measures in such trials.3,4
The BFMD-RS is currently the only instrument that is recommended by the International and Parkinson Movement Disorder Society for evaluating the severity of generalized dystonia. Other available scales, such as the Global Dystonia Rating Scale and the Unified Dystonia Rating Scale, are only suggested by the task force for such a purpose. A PubMed search, using two terms (“dystonia” AND “Burke-Fahn-Marsden Dystonia Rating Scale”) for All Fields, conducted on November 20, 2019, found 209 publications. A considerable part of these items was DBS interventions reporting changes in the scores of the BFMD-RS and the BFMD-DS as their outcomes. According to the results of a recent analysis, the SF-36 is the most commonly used tool in studies reporting aspects of HRQoL in dystonia. Consequently, the BFMD-RS, BFMD-DS, and SF-36 are widely used in clinical research for dystonia; however, their minimal clinically important difference (MCID) thresholds have not yet been established. Although the first randomized and controlled trials on DBS for dystonia used the arbitrary >25% improvement in BFMD-RS scores as the indicator of clinical relevance, the accuracy and feasibility of this approach have never been tested before.

In clinical practice, a discrepancy may exist between levels of improvement or worsening in the objective severity of disease and the changes perceived by the patient. For example, the level of symptomatic improvement reported by the patient may be superior to that captured by clinicians. Therefore, the sole use of threshold values established by only objective estimations for detecting clinically relevant changes in the severity of patients’ symptoms may lead to distortions during evaluation of the efficacy of a treatment. However, the use of MCID threshold values, which reflect the smallest changes in an outcome measure that are meaningful to patients, may serve as a more feasible approach for revealing clinically important changes and contribute thereby to a more reliable translation of clinical outcomes into clinical practice.

Because no clinimetrically verified threshold values for detecting minimal but clinically relevant changes are available for the BFMD-RS, BFMD-DS, and SF-36 in generalized and segmental isolated (primary) dystonia, we aimed to determine these hallmarks.

Materials and Methods

The study protocol was similar to the procedure Makkos and colleagues used to establish MCID estimates for the Unified Dyskinesia Rating Scale. A consecutive series of patients with idiopathic and inherited (torsin family 1 member A [TOR1A] positive) segmental and generalized isolated dystonia aged >18 years were enrolled with the ethical approval of the Regional and Institutional Ethical Committee (3617.316-24987/KK41) in the Department of Neurology, Pécs, Hungary, between 2013 and 2019. None of the included patients had structural brain abnormalities capable of producing dystonia, hallmarks of neurodegeneration (eg, brain iron accumulation) according to MRI, or other known causes of acquired dystonia, including neuroleptic use. The diagnosis of dystonia was confirmed by a neurologist specialized in the diagnosis and treatment of movement disorders.

In addition to demographic, treatment, and disease-related data, the BFMD-RS, BFMD-DS, and the SF-36 were obtained at baseline. With respect to the SF-36, the Physical Component Summary (PCS) and Mental Component Summary (MCS) scores were also calculated in addition to the scores of the eight subscales (physical functioning, social functioning, role limitations because of physical problems, role limitations because of emotional problems, mental health, energy and fatigue, bodily pain, and general health). Although it is not recommended to calculate a single measure of HRQoL based on the individual SF-36 domains, the SF-36 Global Score, which has previously been called Total or Overall Score, has been increasingly reported during the past 20 years. Therefore, we also generated such a global measure by the arithmetic averaging of the scores of the eight subscales. Neurocognitive performance was also measured to detect major neurocognitive disorder (Montreal Cognitive Assessment [MoCA] score < 20.5), which served as an exclusion criterion.

| Characteristic | Mean ± SD or Count (%) |
|----------------|------------------------|
| Age (years)* | 46.1 ± 16.2 |
| Males | 87 (43.9) |
| Handedness | 176 (88.9) |
| Right | 22 (11.1) |
| Left | 12.5 ± 3.0 |
| Education (years) | 10.6 ± 10.4 |
| Duration of disease (years) | 136 (68.7) |
| Distribution of dystonia b | 62 (31.3) |
| Segmental | 39.3 ± 26.2 |
| Generalized | 8.6 ± 6.6 |
| BFMD-DS | 43.5 ± 20.5 |
| SF-36 PCS | 59.9 ± 27.1 |
| SF-36 MCS | 52.0 ± 22.2 |
| SF-36 Global Score c | 46.1 ± 16.2 |

*Every patient was at least 18 years old at baseline.
**Patients only with isolated dystonia were included.
This score was calculated by the arithmetic averaging of the scores of the eight subscales.
SD, standard deviation.
### TABLE 2. Mean changes in scores of the SF-36, BFMD-RS, and BFMD-DS compared to PGI-I scores

| Outcome Measure | PGI-I | No. of Paired Visits | Change (Follow-up vs. Baseline) | ROC Analysis |
|-----------------|-------|----------------------|---------------------------------|-------------|
|                 |       |                      | Mean | Standard Deviation | 95th CI | Effect Size (Cohen’s d) | Optimal Cutoff | Sensitivity (%) | Specificity (%) | LR+ (%) | LR- (%) |
| SF-36 Physical Functioning | A little better | 154 | 3.8 | 13.4 | −1.3 | 8.9 | 0.25 | 3.5 | 60.2 | 57.7 | 1.3 | 0.7 |
| Role Limitations due to Physical Problems | The same | 235 | 0.3 | 7.4 | −2.2 | 3.1 | 0.02 | N/A | N/A | N/A | N/A | N/A |
| SF-36 Physical Functioning | A little worse | 128 | −2.9 | 10.2 | −7.2 | 0.4 | −0.20 | −3.5 | 67.4 | 60.1 | 1.4 | 0.7 |
| SF-36 Role Limitations due to Bodily Pain | The same | 235 | 0.3 | 16.1 | −3.2 | 4.7 | 0.05 | N/A | N/A | N/A | N/A | N/A |
| SF-36 Physical Functioning | A little worse | 128 | −4.6 | 25.2 | −10.6 | 0.4 | −0.22 | −5.0 | 69.8 | 71.1 | 1.5 | 0.7 |
| SF-36 Emotional Problems | The same | 235 | 0.5 | 23.6 | −5.5 | 6.5 | 0.03 | N/A | N/A | N/A | N/A | N/A |
| SF-36 Physical Functioning | A little worse | 128 | −3.1 | 11.5 | −1.6 | −7.8 | −0.25 | −3.5 | 67.8 | 69.9 | 1.7 | 0.8 |
| SF-36 Energy and Fatigue | The same | 235 | 0.4 | 12.5 | −0.5 | 5.9 | 0.24 | 2.5 | 66.7 | 67.8 | 1.7 | 0.8 |
| SF-36 Mental Health | A little better | 154 | 2.4 | 12.5 | −0.5 | 5.9 | 0.24 | 2.5 | 66.7 | 67.8 | 1.7 | 0.8 |
| SF-36 Physical Functioning | A little worse | 128 | −3.1 | 11.5 | −1.6 | −7.8 | −0.25 | −3.5 | 67.8 | 69.9 | 1.7 | 0.8 |
| SF-36 Emotional Problems | The same | 235 | 0.5 | 23.6 | −5.5 | 6.5 | 0.03 | N/A | N/A | N/A | N/A | N/A |
| SF-36 Emotional Problems | A little worse | 128 | −3.1 | 11.5 | −1.6 | −7.8 | −0.25 | −3.5 | 67.8 | 69.9 | 1.7 | 0.8 |
| SF-36 Emotional Problems | The same | 235 | 0.5 | 23.6 | −5.5 | 6.5 | 0.03 | N/A | N/A | N/A | N/A | N/A |

The number of return visits did not appreciably alter these values.

*Relative changes in the BFMD-RS scores were calculated using the following formula: (score_{baseline} − score_{follow-up}) / score_{baseline}.

CI, confidence interval; LR+, positive likelihood ratio; LR−, negative likelihood ratio; N/A = not applicable.

All enrolled patients were asked to return for follow-up every 12 months. At follow-up, the BFMD-RS, BFMD-DS, and SF-36 were reassessed. The magnitude of the perceived change in motor symptoms and disease-related difficulties since the last visit was measured using the Patient-rated Global Impression of Disease.
Improvement (PGI-I) scale (1 = very much better; 2 = much better; 3 = a little better, 4 = no change; 5 = a little worse; 6 = much worse; and 7 = very much worse).²⁹

### Statistical Analysis

The detailed methods for estimating MCID were described previously.³⁰ Briefly, MCID values were determined following the recommended strategy,³¹ including the simultaneous use of both anchor- and distribution-based methods. Anchor-based methods estimate MCID by using an independent and clinically relevant tool that is simultaneously interpretable by itself and has a sufficiently strong correlation with the evaluated instrument.³¹,³² Spearman’s correlation method was used to test whether correlation coefficients between the PGI-I and changes in BFMD-RS, BFMD-DS, or SF-36 reached the required minimum of 0.3³¹ Correlation coefficients were 0.443, 0.357, and 0.374, respectively. Although all correlation coefficients exceeded the required minimum, ordinal regression modeling was also performed between the PGI-I (dependent value) and changes in scores of the assessed scales, effect sizes, MCID values, and results of ROC curve analysis for the whole study population are shown in Table 2. Controlling for changes in treatment for dystonia during follow-up was associated with a PGI-I score of 3, 4, or 5, was 517. The median number of follow-up visits was four, with the median intervisit interval of 12 months. Baseline characteristics of the study cohort are represented in Table 1.

Changes in treatment for dystonia during follow-up are shown in Supporting Information Table S1. A total of 136 patients (68.7%) were treated with DBS at the last follow-up.

Significant ordinal logistic regression models could be developed between the PGI-I and changes in BFMD-RS (Nagelkerke pseudo-R-square: 0.412; P < 0.01) and BFMD-DS scores (Nagelkerke pseudo-R-square: 0.389; P < 0.05). We could also build a significant ordinal logistic regression model between the PGI-I and changes in scores of the SF-36 (Nagelkerke pseudo-R-square: 0.461; P < 0.01).

Mean changes in BFMD-RS, BFMD-DS, and SF-36 scores, effect sizes, MCID values, and results of ROC curve analysis for the whole study population are shown in Table 2. Controlling for TOR1A gene testing (TOR1A positive [ n = 59] vs. negative cases) did not alter the calculated thresholds considerably (Table 3).

### Discussion

The concept of MCID is increasingly used in biomedical research for judging whether statistical significance implies clinical relevance. However, MCID scores for the BFMD-RS, BFMD-DS, and SF-36 had not yet been evaluated in the population with segmental and generalized isolated dystonia. Therefore, we aimed to calculate MCID thresholds for these instruments. Following the recommendations of Revicki and colleagues,³¹ the cut-off values on the BFMD-RS and BFMD-DS for minimal, yet clinically meaningful, improvement and worsening were 16.6% and 0.5 points and 21.5% and 0.5 points, respectively. Cut-off scores for the PCS, MCS,
and the Global Score of the SF-36 for observing clinically meaningful improvement and relevant deterioration regarding HRQoL in dystonia could be set at 5.5 and 5.5, 6.5 and 7.5, and 7.5 and 8.5 points, respectively.

As far as the authors are aware, present MCID estimations for the BFMD-RS and BFMD-DS cannot be compared to those of other studies because this is the first report on such threshold values for these instruments. Previous studies investigating DBS for isolated dystonia used a threshold value of >25% improvement in the BFMD-RS for considering clinical efficacy and >50% improvement for identifying “good” responders.\(^9\)\(^{13}\)\(^\text{-}\)\(^{17}\)\(^{35}\) Furthermore, patients having a 25% to 50% decrease in the BFMD-RS have been reported as partial responders.\(^3\)\(^6\)\(^\text{-}\)\(^\text{38}\) Using the recommended methods for MCID estimation, we found that the threshold for minimal, yet clinically relevant, improvement in BFMD-RS scores may lie at an even lower level (16.6%). Our MCID threshold may give some explanations for those patients who reported perceived improvement despite <25% improvement on the BFMD-RS after DBS.

Some previous studies have already evaluated MCID thresholds for the SF-36 in patients with asthma,\(^3\)\(^9\) chronic obstructive pulmonary disease,\(^3\)\(^9\) rheumatoid arthritis,\(^4\)\(^0\) and chronic fatigue syndrome,\(^4\)\(^1\) and in patients undergoing total hip\(^4\)\(^2\) or knee replacement surgeries.\(^4\)\(^3\) where MCID values for SF-36 subscales varied between 0.4 and 25 points and between 8.3 and 20 points for improvement and decline, respectively. Although the SF-36 is a general health status measure, application of the aforementioned MCID estimations to dystonia patients may be misleading and inappropriate because MCID is highly dependent on characteristics of the study population.\(^3\)\(^9\)\(^\text{-}\)\(^\text{44}\)

The concept of MCID was established to overcome the issue of statistical significance not necessarily implying clinical importance. Therefore, our estimations may be useful in judging the clinical relevance of results from previous and future studies by using the BFMD-RS, BFMD-DS, and SF-36 to measure the effectiveness of treatments for isolated dystonia, especially DBS. According to our MCID calculations, the beneficial effects of neurosurgical procedures for isolated dystonia on the severity of dystonia and disability related to dystonia intensify with time, and, after reaching a plateau, they remain clinically relevant during long-term follow-up. In addition, improvements developing a long time after neurosurgical interventions exceed the MCID thresholds established in the present study in a more pronounced manner compared to those measured soon after the surgery (Supporting Information Tables S2 and S3).

The strength of the present approach lies partly in the simultaneous use of anchor- and distribution-based methods, resulting in similar MCID estimations. In addition, we provide MCID scores for scales evaluating changes in severity of dystonia both objectively and from patients’ perspectives. The high number of included patients may also ensure reliability and enhance the wider applicability of the calculated thresholds. However, our estimations should be used with caution because differences in characteristics of study populations may exist. Based on disease severity, our patient population is largely characteristic for patients suitable for DBS therapy, but less likely proper for a focal dystonia cohort referred for botulinum neurotoxin treatment. Given that the present study cohort did not include patients with dystonia of genetic origin other than TOR1A gene, acquired dystonia, and pediatric patients, our MCID thresholds are not necessarily feasible in such populations. In addition, although we tried to investigate a relatively homogeneous population (idiopathic and inherited isolated dystonia syndromes with segmental and generalized distribution), a considerable heterogeneity may exist in both motor and nonmotor symptoms, which could not be specifically addressed during the study.

References

1. Girach A, Vinagre Aragon A, Ziu P. Quality of life in idiopathic dystonia: a systematic review. J Neurol 2019;266:2897–2906.
2. van den Dool J, Tijssen MA, Koelman JH, Engelbert RH, Visser B. Determinants of disability in cervical dystonia. Parkinsonism Relat Disord 2016;32:48–53.
3. Burke RE, Fahn S, Marsden CD, Bressman SB, Moskowitz C, Friedman J. Validity and reliability of a rating scale for the primary torsion dystonias. Neurology 1985;35:73–77.
4. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473–483.
5. Albanese A, Sorbo FD, Comella C, et al. Dystonia rating scales: critique and recommendations. Mov Disord 2013;28:874–883.
6. Eltahawy HA, Saint-Cyr J, Giladi N, Lang AE, Lozano AM. Primary dystonia is more responsive than secondary dystonia to pallidal interventions: outcome after pallidotomy or pallidal deep brain stimulation. Neurosurgery 2004;54:613–619; discussion, 619–621.
7. Gruber D, Kuhn AA, Schoenecker T, et al. Pallidal and thalamic deep brain stimulation in myoclonus-dystonia. Mov Disord 2010;25:1733–1743.
8. Sako W, Morijaki R, Mizobuchi Y, et al. Bilateral pallidal deep brain stimulation in primary Meige syndrome. Parkinsonism Relat Disord 2011;17:123–125.
9. Cif L, Gonzalez-Martinez V, Vasques X, et al. Staged implantation of multiple electrodes in the internal globus pallidus in the treatment of primary generalized dystonia. J Neurosurg 2012;116:1144–1152.
10. Bruggemann N, Kuhn A, Schneider SA, et al. Short- and long-term outcome of chronic pallidal neurostimulation in monogenic isolated dystonia. Neurology 2015;84:895–903.
11. Lettieri C, Rinaldo S, Devigili G, et al. Clinical outcome of deep brain stimulation for dystonia: constant-current or constant-voltage stimulation? A non-randomized study. Eur J Neurol 2015;22:919–926.
12. Zhan S, Sun F, Pan Y, et al. Bilateral deep brain stimulation of the subthalamic nucleus in primary Meige syndrome. J Neurosurg 2018;128:897–902.
13. Vidalith M, Vercueil L, Houeto JL, et al. Bilateral deep-brain stimulation of the globus pallidus in primary generalized dystonia. N Engl J Med 2005;352:459–467.
14. Kupsch A, Benecke R, Muller J, et al. Pallidal deep-brain stimulation in primary generalized or segmental dystonia. N Engl J Med 2006;355:1978–1990.
