Myasthenia Gravis Impairment Index

Responsiveness, meaningful change, and relative efficiency

OPEN 📖

Carolina Barnett, MD, PhD

Vera Bril, MD, FRCP

Moira Kapral, MD

Abhaya V. Kulkarni, MD, PhD

Aileen M. Davis, PhD

Correspondence to
Dr. Barnett:
c.barnetttapia@utoronto.ca

ABSTRACT

Objective: To study responsiveness and meaningful change of the Myasthenia Gravis Impairment Index (MGII) and its relative efficiency compared to other measures.

Methods: We enrolled 95 patients receiving prednisone, IV immunoglobulin (IVIg), or plasma exchange (PLEX) and 54 controls. Patients were assessed with the MGII and other measures—including the Quantitative Myasthenia Gravis Score, Myasthenia Gravis Composite, and Myasthenia Gravis Activities of Daily Living—at baseline and 3–4 weeks after treatment. Statistical markers of responsiveness included between-groups and within-group differences, and we estimated the relative efficiency of the MGII compared to other measures. Patient-meaningful change was assessed with an anchor-based method, using the patient's impression of change. We determined the minimal detectable change (MDC) and the minimal important difference (MID) at the group and individual level.

Results: Treated patients had a higher change in MGII scores than controls (analysis of covariance $p < 0.001$). The ocular domain changed more with prednisone than with IVIg/PLEX (effect size 0.67 and 0.13, analysis of covariance $p = 0.001$). The generalized domain changed more with IVIg/PLEX than with prednisone (effect size 0.50 and 0.22, analysis of covariance $p = 0.07$). For the total MGII score, the individual MDC95 was 9.1 and the MID was 5.5 for individuals and 8.1 for groups. Relative efficiency ratios were >1 favoring the MGII.

Conclusions: The MGII demonstrated responsiveness to prednisone, IVIg, and PLEX in patients with myasthenia. There is a differential response in ocular and generalized symptoms to type of therapy. The MGII has higher relative efficiency than comparison measures and is viable for use in clinical trials. *Neurology*® 2017;89:2357–2364

GLOSSARY

ANCOVA = analysis of covariance; **AUC** = area under the curve; **CI** = confidence interval; **ES** = effect size; **IVIg** = IV immunoglobulin; **MDC** = minimal detectable change; **MG** = myasthenia gravis; **MG-ADL** = Myasthenia Gravis Activities of Daily Living; **MGC** = Myasthenia Gravis Composite; **MGII** = Myasthenia Gravis Impairment Index; **MID** = minimal important difference; **PIC** = patient impression of change; **PLEX** = plasma exchange; **QMGS** = Quantitative Myasthenia Gravis Score; **RCT** = randomized controlled trial; **ROC** = receiver operator characteristic; **SEM** = standard error of measurement; **SRM** = standardized response mean.

The Myasthenia Gravis Impairment Index (MGII) is a novel measure of myasthenia gravis (MG) severity, with demonstrated feasibility, reliability, and construct validity.[1] Strengths of the MGII are that it was developed following current recommendations,[2] incorporating patient input at different stages, and has less floor effect than other commonly used measures.[1] A floor effect means that some symptomatic patients might have scores at the lower end of the scale, making it difficult to document change after an intervention. Therefore, the MGII might be more sensitive to detect change than other measures, but responsiveness has not been assessed.

Different methods to determine responsiveness reflect different views of what relevant means.[3,4] Statistical measures of responsiveness—such as *t* tests—may detect differences that are not meaningful for patients. The minimal detectable change (MDC) is the smallest change that is significantly beyond error of measurement.[3,5] Therefore, the MDC is useful to understand if a change in score is likely true change—more than error or not just error—but does not provide the patient's perspective. The minimal important difference (MID) is the smallest change that patients consider meaningful[5] and it should be larger than the MDC to be interpretable. Different methods to estimate the MID have been proposed. These include distribution-based methods—e.g., effect size (ES) and the standard error of measurement (SEM)—and anchor-based methods using an external criterion, such as patients' ratings of change. The former reflect statistical change and the latter—by incorporating the patients' views—imply meaningfulness from a patient's perspective.[3,5] Therefore, anchor-based methods are preferred for determining meaningful change.[2] The MDC and MID can be calculated at the group level—the smallest important mean change in a group such as in a clinical trial—or individual level: the smallest change to determine whether an individual has improved.[5] Table e-1 at Neurology.org shows different definitions of responsiveness and meaningful change.

We studied responsiveness and meaningful change of the MGII in patients with MG requiring treatment with prednisone, IV immunoglobulin (IVIg), or plasma exchange (PLEX), where clinical change is expected within short periods of time.[6,7] We followed published recommendations[8] and used different approaches to determine statistical and patient-meaningful change. Finally, we estimated the MID and the MDC to aid in the interpretation of change scores. As a secondary objective, we compared the relative efficiency of the MGII to other measures.

METHODS **Standard protocol approvals, registrations, and patient consents.** The University Health Network Research Ethics Board approved the study and all patients provided informed consent.

Patients. Patients with MG attending the Prosserman Family Neuromuscular Clinic, Toronto General Hospital, between June 2014 and June 2016, were eligible if their physician initiated prednisone, increased prednisone dose by ≥50%, or prescribed a course of IVIg or PLEX. We assessed patients at baseline and 3–4 weeks after IVIg/PLEX or after starting prednisone. Patients received IVIg or PLEX according to standard protocols.[6] The dose of prednisone was determined by the treating physician. We included as controls those patients who participated in the MGII reliability study[1] who did not receive an intervention and had a second visit within 3 weeks.

Measures. The MGII has 22 patient-reported and 6 examination items. Total score ranges from 0 to 84, higher scores indicating more severe impairments. The MGII can be divided into 2 subscores, ocular (8 items) and generalized (20 items), and missing responses are imputed using the mean score of the item domain (ocular/generalized). At follow-up, patients indicated if they felt better, worse, or unchanged, as patient impression of change (PIC). Patients who felt better or worse answered a 5-level Likert scale, ranging from "almost the same" to "much better/worse." Patients were also assessed with the Myasthenia Gravis Composite (MGC),[9] the Quantitative Myasthenia Gravis Score (QMGS),[10] and the Myasthenia Gravis Activities of Daily Living (MG-ADL)[11] for MG severity. Patients completed the NeuroQoL-fatigue short form[12] to quantify fatigue and disease-specific (MG-QOL15)[13] and generic (EQ-5D-5L)[14] quality of life questionnaires. Table e-2 provides details of these measures.

Analyses. Demographic data were evaluated with means ± SD, or counts and proportions. There are no formal methods to calculate sample size for responsiveness studies, but guidelines recommend a minimum of 50 patients.[8] We aimed to enroll 50 patients per treatment group (prednisone and IVIg/PLEX) to ensure a broad range of treatment response. We used R statistical software v.3.1.2 and considered *p* values <0.05 as significant.

Statistically significant change. *Between-groups responsiveness.* We compared the mean change in MGII scores between treated patients and controls, expecting higher change in treated patients (unpaired *t* test); because of differences in baseline scores among treatment groups, we used analysis of covariance (ANCOVA) to compare the 3 groups (prednisone, IVIg/PLEX, and controls). We combined IVIg and PLEX in one group, as they have similar efficacy.[6,15] We also calculated between-groups ES[8,16] for the total, ocular, and generalized scores.

We studied the subgroup of patients with pure ocular disease, comparing the mean change in scores in treated vs untreated patients (unpaired *t* test). We expected significant change in the total and ocular scores but not in the generalized score.

Within-group responsiveness. We performed paired *t* tests (baseline–visit 2) for the total, ocular, and generalized MGII scores, for all treated patients and by treatment group (prednisone and IVIg/PLEX). We calculated the relative efficiency of the MGII compared with the QMGS, the MGC, and the MG-ADL, using the paired *t* test statistic as follows: [*t* statistic MGII/*t* statistic comparison measure].[2,17,18] With the MGII in the numerator, a ratio >1 indicates that the MGII is more efficient—requires a smaller sample size to detect a specific effect size—than the comparator. We also estimated the standardized response mean (SRM) as shown in table e-1.[16]

We studied longitudinal construct validity through correlations between the change in MGII scores and other measures.[8] We expected moderate (*r* 0.4–0.7) correlations with change in the MGC, QMGS, MG-ADL, and NeuroQoL-fatigue and low to moderate correlations (*r* 0.2–0.5) with change in quality of life

(EQ-5D and MG-QOL15) scores. Confirming ≥75% of predefined hypotheses is a marker of construct validity.[19]

Patient meaningful change. We compared the change in MGII scores across PIC categories (better, unchanged, and worse) in treated patients, adjusting for baseline values (ANCOVA). We expected a significant difference in change scores with improved patients changing more than those unchanged or worse. We calculated the correlation between the PIC and the change scores, and correlations ≥0.3 provide evidence that the anchor (PIC) is appropriate.[18]

Interpretation of change scores: MDC and MID. We estimated the MDC with 90% and 95% confidence intervals (CIs), using the SEM, as seen in table e-1.[4,5,20] SEM values for the MGII are 3.3 for the total and 2.2 for the ocular and generalized scores.[1] The MDC at the group level is the individual MDC divided by the squared root of the sample size.[21] We planned different anchor-based methods to estimate the MID using the PIC option "a little better" as

marker of minimal improvement. We calculated the MID at a group level as the mean change in patients feeling "a little better," and also the 75% percentile for potential misclassification bias.[4,22] Additionally, we built receiver operator characteristic (ROC) curves, considering patients feeling "a little better" and higher as improved; the point of best sensitivity and specificity is the MID at the individual level.[5,23] For each estimate of the individual-level MDC and MID, we calculated sensitivity, specificity, positive predictive values, and negative predictive values to detect a responder. Because MID values are affected by baseline scores and can also differ between improvement and worsening, we planned to calculate the MID for worsening and by different baseline scores, given enough patients.[5,24]

RESULTS Of 111 eligible patients, 95 received the prescribed treatment and returned for the second assessment. Fifty-four patients from the reliability

Table 1	Demographic and clinical characteristics at baseline			
	Controls (n = 54)	Prednisone (n = 50)	IVIg/PLEX (n = 45)	p Value
Age, y	61 ± 13	57 ± 15	51 ± 18	0.05
Female sex	26 (48)	22 (44)	32 (71)	0.39
Disease duration, mo	141 ± 153	93 ± 114	65 ± 78	0.008
Generalized disease	47 (87)	34 (68)	45 (100)	0.69
Thymoma	10 (19)	12 (24)	12 (24)	0.22
Thymectomy	29 (54)	29 (58)	24 (53)	0.94
AChR Ab[a]	20 (59)	16 (67)	27 (73)	0.51
MuSK	0 (0)	1 (1)	0 (0)	0.90
MGII total score	17 ± 13	26 ± 16	37 ± 16	<0.001
MGII ocular score	6.4 ± 6	12.5 ± 6	10.2 ± 7	<0.001
MGII generalized score	10.3 ± 9	13.5 + 14	27 ± 11	<0.001
QMGS score	6.9 ± 4.5	9.1 ± 5.6	13.6 ± 7.7	<0.001
MGC score	4.8 ± 4.3	8.1 ± 6.2	13.1 ± 7.7	<0.001
MG-ADL score	3.6 ± 3.2	5.7 ± 3.7	7.7 ± 3.8	<0.001
MGFA class				<0.001
R	3 (5)	0 (0)	0 (0)	
MM	5 (9)	0 (0)	0 (0)	
I	10 (19)	20 (40)	0 (0)	
II	29 (54)	17 (34)	14 (31)	
III	7 (13)	11 (22)	25 (56)	
IV	0 (0)	2 (4)	6 (13)	
Treatments				
Pyridostigmine	35 (65)	33 (66)	36 (80)	0.20
Prednisone	31 (57)	17 (34)	15 (33)	0.02
Azathioprine	12 (22)	9 (18)	7 (16)	0.55
Mycophenolate	12 (22)	6 (12)	7 (16)	0.26
No medications	2 (4)	9 (18)	6 (13)	0.05

Abbreviations: AChR Ab = acetylcholine receptor antibodies; I = pure ocular; II = mild generalized; III = moderate generalized; IV = severe generalized; IVIg = IV immunoglobulin; MG-ADL = Myasthenia Gravis Activities of Daily Living; MGC = Myasthenia Gravis Composite; MGFA = Myasthenia Gravis Foundation of America; MGII = Myasthenia Gravis Impairment Index; MM = minimal manifestations; MuSK = Muscle Specific Kinase antibodies; PLEX = plasma exchange; QMGS = Quantitative Myasthenia Gravis Score; R = remission.
[a] Antibody data available for 95 patients: 34, 24, and 37 patients in the control, prednisone, and IVIg/PLEX group, respectively. The percentage presented reflects positive antibodies on those tested.

study were included as controls. Demographic data are reported in table 1. Of the treated patients, 50 (53%) received prednisone—mean dose 23 ± 10 mg per day—and 45 (47%) IVIg/PLEX. All MGII items had <6% missing responses.

Statistically significant change. *Between-groups responsiveness.* The mean change in MGII score was −6.9 ± 11.8 for all treated patients and −0.6 ± 5.3 for controls (*t* test, *p* < 0.0001); the mean change for prednisone was −7.0 ± 10.3 and −6.8 ± 13.2 for IVIg/PLEX (3 groups [including controls] ANCOVA *p* < 0.0001). The ocular score changed more in the prednisone than the IVIg/PLEX group, even after correcting for baseline differences (−4.7 vs −1.5, ANCOVA *p* = 0.001). The change in generalized score was smaller in the prednisone compared to the IVIg/PLEX group, not reaching statistical significance (−2.3 vs −5.4, ANCOVA *p* = 0.07). Responsiveness statistics are in table 2.

In the pure ocular disease subgroup (n = 23), 16 patients received prednisone and 7 were controls. The mean change in total score was −3.6 ± 6.9 for prednisone and 0.3 ± 3.2 for controls (*t* test *p* = 0.07, ES = 1.3). The mean change in the ocular score was −4.1 ± 5.9 for prednisone vs 0.6 ± 1.9 for controls (*p* = 0.009, ES = 1.1). The generalized score did not change in treated or control patients (−0.3 ± 2.2 vs 0.5 ± 3.3, *p* = 0.8).

Within-group responsiveness. The paired *t* tests for the total, generalized, and ocular scores showed significant differences for all treatments (*p* < 0.05). The ocular score SRM was higher for prednisone (0.85) than immunomodulation (0.31). The generalized score SRM was higher for immunomodulation (0.51) than prednisone (0.31). All the relative efficiency ratios were >1 (table 3).

As expected, the MGII change scores had moderate correlations with change in other MG severity measures (*r* 0.48–0.71) and lower correlations with quality of life measures (*r* 0.29–0.39; table e-3).

Patient meaningful change. Of the 95 treated patients, 63 (66%) felt better, 18 (19%) unchanged, and 14 (15%) worse at follow-up. The correlations between PIC and change in total, ocular, and generalized MGII scores were 0.43, 0.33, and 0.32, respectively (*p* ≤ 0.001). The mean change in total, ocular, and generalized MGII scores was significantly different in patients who were better than those unchanged or worse (ANCOVA *p* < 0.0001, table 4).

The mean change in MGII scores in responders—patients who were a "little better" or higher (n = 58)—was −10.7 ± 11.9 for the total, −4.3 ± 4.8 for the ocular, and −6.4 ± 9.6 for the generalized scores.

MCD and MID. At the individual level, the MDC_{95} was 9.1 for the total score and 6.0 for the ocular and generalized scores; the MDC_{90} was 7.7 for the total and 5.1 for the ocular and generalized scores. At the group level, the MDC_{95} was 1.5 for the total and 0.9 for the generalized and ocular scores; the MDC_{90} was 1.1 for total and 0.8 for the ocular and generalized scores.

Eleven patients were "a little better" and the MID—group level—was 8.1 points (median 9.0; 75th percentile 2.5) for the total MGII score, 2.6 points (median 2; 75th percentile 1) for the ocular score, and 5.5 points (median 5; 75th percentile 1) for the generalized score. We could not estimate the MID by baseline values given the small group who were "a little better."

The ROC curve for the change in total MGII score had an area under the curve (AUC) of 0.76 (CI 0.66–0.86; figure e-1). The optimum cutpoint was 5.5 with 64% sensitivity and 73% specificity. The ROC for the generalized MGII score had AUC of 0.74 (CI 0.63–0.83) and a cutpoint of −2.5 had 59% sensitivity and 76% specificity. The ocular MGII score ROC had an AUC of 0.66 and was not reliable for cutpoints. Only 14 patients felt worse, so we could not estimate the MID for worsening. Table 5 summarizes individual MDC and MID cutpoints, with their sensitivity, specificity, and predictive values.

DISCUSSION This study provides evidence that the MGII is sensitive to detect change in patients with MG receiving prednisone, IVIg, or PLEX, as treated

Table 2	Between-groups responsiveness of the Myasthenia Gravis Impairment Index and subscores						
		Prednisone (n = 50)		IVIg/PLEX (n = 45)		All treated (n = 95)	
	Controls (n = 54), change	Change	ES	Change	ES	Change	ES
Total score	−0.6 ± 5.3	−6.9 ± 10.3	0.37	−6.8 ± 13.2	0.37	−6.9 ± 11.8	0.37
Ocular score	−0.7 ± 2.8	−4.7 ± 5.4	0.67	−1.5 ± 4.8	0.13	−3.1 ± 5.4	0.40
Generalized score	0.1 ± 4.2	−2.3 ± 7.4	0.22	−5.4 ± 10.6	0.50	−2.9 ± 6.7	0.28

Abbreviations: ES = between-groups effect size (change treatment − change control/SD baseline); IVIg = IV immunoglobulin; PLEX = plasma exchange. Analysis of covariance showed that the change in total, ocular, and generalized scores was significantly different among controls, prednisone, and IVIg/PLEX groups (*p* < 0.001), correcting for baseline values.

	Table 3		Within-groups responsiveness and relative efficiency of myasthenia gravis impairment measures									
	All treated (n = 95)				**Prednisone (n = 50)**				**IVIg/PLEX (n = 45)**			
Measure	**Change**	**SRM**	**t Test**	**Efficiency**	**Change**	**SRM**	**t Test**	**Efficiency**	**Change**	**SRM**	**t Test**	**Efficiency**
MGII	−6.9 ± 11.8	0.58	−5.7	1.0	−6.9 ± 10.3	0.69	−4.8	1.0	−6.8 ± 13.2	0.52	−3.5	1.0
QMGS	−1.3 ± 3.6	0.36	3.4	2.8	−1.4 ± 2.5	0.56	4.0	1.4	−1.1 ± 4.5	0.24	1.7	4.2
MGC	−2.3 ± 5.9	0.39	3.7	2.4	−1.5 ± 4.4	0.34	2.4	4.0	−3.1 ± 7.2	0.43	2.9	1.5
MG-ADL	−1.4 ± 2.7	0.52	4.7	1.5	−1.5 ± 2.5	0.60	4.1	1.4	−1.2 ± 2.9	0.41	2.7	1.7

Abbreviations: IVIg = IV immunoglobulin; MG-ADL = Myasthenia Gravis Activities of Daily Living; MGC = Myasthenia Gravis Composite Scale; MGII = Myasthenia Gravis impairment Index; PLEX = plasma exchange; QMGS = Quantitative Myasthenia Gravis Score; SRM = standardized response mean (change score [T2 − T1]/SD difference).
Efficiency = relative efficiency coefficient = (t statistic MGII/t statistic other measure)². All coefficients were >1, indicating higher efficiency of the MGII; t test = t statistic of the paired t test (baseline–visit 2).

patients had higher change in MGII scores than controls. The ocular score was more responsive to prednisone than to IVIg/PLEX, which could be in part explained by baseline differences since the prednisone group had slightly higher baseline ocular score (12.5 vs 10.2); however, the difference in change scores remained significant after correcting for baseline values. Therefore, the different responsiveness of subscores possibly reflects a specific treatment effect. In a randomized controlled trial (RCT) of IVIg vs placebo, patients with pure ocular disease responded less than those with generalized disease,[25,26] supporting our findings. Using subscores might help to compare the efficacy of different treatments on different body regions affected by MG.

The MGII had higher relative efficiency than the other measures tested, meaning that it will require a smaller sample size to detect the same effect magnitude. The change in QMGS scores was smaller in this study than in previous RCTs of IVIg/PLEX.[6,25] In the present study, the inclusion criterion was only that the treating physician prescribed IVIg/PLEX, while the RCTs had more stringent criteria, which might result in more responsive patients. In addition, patients in RCTs often do better than patients in real-world settings. In fact, in the IVIg vs placebo study, no treated patients were worse at follow-up,[25] while in this study 9 (20%) of the IVIg/PLEX patients reported worsening. Despite this limitation, the MGII

was more sensitive to detect change than the other measures used.

We found that the total and ocular scores are responsive in patients with pure ocular disease—without significant change in the generalized score—confirming minimal input of the generalized items in pure ocular patients' outcome.[1] This supports the use of the ocular score as an outcome measure for pure ocular patients.

The MGII can also detect change that is meaningful for patients, and the change scores were significantly different across the PIC categories of better, worse, and unchanged. The correlation between the PIC with the MGII change scores supports its use as anchor. Previous studies on MG measures[27,28] have used the change in the MG-QOL15 and physician impression of change as criteria for improvement. However, the MID for the MG-QOL15 is unknown and given the short-term interventions, we did not expect major changes in quality of life.

To aid in the interpretation of change scores, we estimated the MDC and the MID at the group and individual level. The estimates at the group level are useful for clinical trials, to interpret the mean change in scores in a treatment group and to calculate sample size. The MID for groups is 8.1 points—above the MDC_{95} at the group level (1.5)—reflecting change above error of measurement. Therefore, we suggest that a difference of 8.1 points can be used to estimate

	Table 4		Mean change scores in treated patients, according to patient impression of change			
MGII component	**Better (n = 63)**	**Unchanged (n = 18)**	**Worse (n = 14)**	**ANCOVA p value**	**Little better (n = 11)**
Total score	−10.0 ± 11.7	−3.2 ± 8.8	2.4 ± 8.8	<0.001	−8.1 ± 8.9
Ocular score	−4.2 ± 4.7	−2.8 ± 5.2	1.1 ± 6.6	<0.001	−2.6 ± 2.8
Generalized score	−5.8 ± 9.5	−0.4 ± 8.0	1.3 ± 5.4	<0.001	−5.5 ± 8.6

Abbreviations: ANCOVA = analysis of covariance; MGII = Myasthenia Gravis Impairment Index.
ANCOVA demonstrated significantly different change in total, ocular, and generalized. MGII scores in patients reporting to be better, unchanged, and worse. Little better is a subgroup of those who were overall better.

Table 5 Performance of different minimal detectable change (MDC) and minimal important difference (MID) cutpoints for the Myasthenia Gravis Impairment Index at the individual level

Measure	Cutpoint	Sensitivity, %	Specificity, %	PPV, %	NPV, %
Total score					
MDC 95	9.1	50	84	83	52
MDC 90	7.7	57	81	83	55
MID individual	5.5	64	73	79	56
Ocular score					
MDC 95	6.0	26	84	69	42
MDC 90	5.1	31	78	69	42
MID individual	NA[a]				
Generalized score					
MDC 95	6.0	43	92	89	51
MDC 90	5.1	43	92	89	51
MID individual	2.5	59	76	79	54

Abbreviations: NPV = negative predictive value; PPV = positive predictive value.
MID individual: Based on the receiver operator characteristic curves, where responders were those "a little better" and higher (n = 58).
[a] The receiver operator characteristic curve for ocular score had area under the curve <0.7 and thus was not reliable.

sample size for trials. For a 2-arm parallel study, with α = 0.05 and 80% power, 86 patients—43 per group—would be required, which is feasible. An alternative would be using the mean change in those patients who were considered responders (10.7 points). The estimates at the individual level help determine whether a patient has had meaningful change, to classify patients who are responders to an intervention. The MID for individuals is 5.5 points, and this is smaller than the MDC_{95} and MDC_{90} (9.1 and 7.0, respectively). Therefore, while a cutpoint of 5.5 has the best sensitivity and specificity to define a responder, there is some risk of misclassification due to error of measurement. Using the MDC to define responders will reduce sensitivity but increase specificity, so choosing a cutpoint depends on the clinical scenario and whether one wants to maximize sensitivity or specificity (table 5).

A strength of this study is that we used a framework of responsiveness and important change to clarify what the different definitions of relevant change mean. We also used an anchor-based approach using patient self-report for the MID estimation avoiding the use of distribution-based methods, which are no longer recommended.[5] To make the results meaningful for clinicians and researchers working with patients with MG, we considered multiple approaches to understand what magnitude of change is meaningful at a group and individual patient level. In addition, we studied responsiveness to specific interventions and had a control group that was stable between assessments that occurred within the same timeframe as the treated group.

A limitation is that although most treated patients felt better, few were just "a little better," most reporting larger improvement. Therefore, the MID calculations at the group level—patients reporting being a little better—should be interpreted with caution. This small group also prevented us from estimating MID values according to baseline scores and because few patients felt worse, we could not estimate the MID for worsening.[4,5] In addition, we only had one MuSK-positive patient, so these findings might not be generalizable to that specific population.

We assessed patients 3 weeks after treatment to ensure homogeneous follow-up time, but we recognize that in the prednisone group the generalized score could have improved more with longer treatment. In addition, this limits the generalizability to chronic treatments (e.g., azathioprine) that have a much slower effect and where the MID values might be different. Since patients might experience adaptation to their symptoms with time, it is possible that they have different views of minimal improvement for interventions that require a long lime to act. Therefore, future studies are needed to understand the MID for slow-acting treatments. For long-term outcomes, studying patient-acceptable symptom states—score thresholds where patients feel good enough rather than better—might be more meaningful.[29]

The MGII is sensitive to detect statistically and patient meaningful change in patients with MG receiving prednisone, IVIg, or PLEX. The MGII is more efficient to detect statistically significant change than the other measures studied and sample

size estimations are feasible for future intervention studies. In addition, the MGII subscores are responsive and the ocular component can be used for clinical trials in pure ocular disease. The ocular and generalized scores may have differential responsiveness according to the intervention and this should be studied with other treatments. The MDC and MID differ for groups and individuals and cutpoints for improvement should be chosen based on the clinical scenario. These findings support the use of the MGII to detect change in patients receiving interventions for MG.

AUTHOR CONTRIBUTIONS

STUDY FUNDING

DISCLOSURE

REFERENCES

1. Barnett C, Bril V, Kapral M, Kulkarni A, Davis AM. Development and validation of the Myasthenia Gravis Impairment Index. Neurology 2016;87:879–886.
2. US Department of Health and Human Services FDA Center for Drug Evaluation and Research; US Department of Health and Human Services FDA Center for Biologics Evaluation and Research; US Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health Qual Life Outcomes 2006;4:79.
3. Beaton DE, Bombardier C, Katz JN, Wright JG. A taxonomy for responsiveness. J Clin Epidemiol 2001;54:1204–1217.
4. Beaton DE. Understanding the relevance of measured change through studies of responsiveness. Spine 2000;25:3192–3199.
5. King MT. A point of minimal important difference (MID): a critique of terminology and methods. Expert Rev Pharmacoecon Outcomes Res 2011;11:171–184.
6. Barth D, Nabavi Nouri M, Ng E, Nwe P, Bril V. Comparison of IVIg and PLEX in patients with myasthenia gravis. Neurology 2011;76:2017–2023.
7. Muscle Study Group. A trial of mycophenolate mofetil with prednisone as initial immunotherapy in myasthenia gravis. Neurology 2008;71:394–399.
8. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PMM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. Qual Life Res 2003;12:349–362.
9. Burns TM, Conaway MR, Cutter GR, Sanders DB; Muscle Study Group. Construction of an efficient evaluative instrument for myasthenia gravis: the MG composite. Muscle Nerve 2008;38:1553–1562.
10. Barohn RJ, McIntire D, Herbelin L, et al. Reliability testing of the quantitative myasthenia gravis score. Ann NY Acad Sci 1998;841:769–772.
11. Wolfe GI, Herbelin L, Nations S, Foster B, Bryan WW, Barohn RJ. Myasthenia gravis activities of daily living profile. Neurology 1999;52:1487–1489.
12. Cook KF, Victorson DE, Cella D, Schalet BD, Miller D. Creating meaningful cut-scores for Neuro-QOL measures of fatigue, physical functioning, and sleep disturbance using standard setting with patients and providers. Qual Life Res 2014;24:575–589.
13. Burns TB, Conaway MR, Cutter GR, Sanders DB; The Muscle Study Group. Less is more, or almost as much: a 15-item quality-of-life instrument for myasthenia gravis. Muscle Nerve 2008;38:957–963.
14. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001;33:337–343.
15. Barnett C, Wilson G, Barth D, Katzberg HD, Bril V. Changes in quality of life scores with intravenous immunoglobulin or plasmapheresis in patients with myasthenia gravis. J Neurol Neurosurg Psychiatry 2012;84:94–97.
16. Norman GR, Wyrwich KW, Patrick DL. The mathematical relationship among different forms of responsiveness coefficients. Qual Life Res 2007;16:815–822.
17. Liang MH, Larson MG, Cullen KE, Schwartz JA. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. Arthritis Rheum 1985;28:542–547.
18. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61:102–109.
19. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34–42.
20. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. Health Qual Life Outcomes 2006;4:54.
21. Busija L, Osborne RH, Nilsdotter A, Buchbinder R, Roos EM. Magnitude and meaningfulness of change in SF-36 scores in four types of orthopedic surgery. Health Qual Life Outcomes 2008;6:55.
22. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific quality of life questionnaire. J Clin Epidemiol 1994;47:81–87.
23. Turner D, Schünemann HJ, Griffith LE, et al. Using the entire cohort in the receiver operating characteristic analysis maximizes precision of the minimal important difference. J Clin Epidemiol 2009;62:374–379.
24. Crosby RD, Kolotkin RL, Williams GR. An integrated method to determine meaningful changes in health-related quality of life. J Clin Epidemiol 2004;57:1153–1160.

25. Zinman L, Ng E, Bril V. IV immunoglobulin in patients with myasthenia gravis: a randomized controlled trial. Neurology 2007;68:837–841.

26. Barnett TC, Bril V, Davis AM. Performance of individual items of the quantitative myasthenia gravis score. Neuromuscul Disord 2013;23:413–417.

27. Muppidi S, Wolfe GI, Conaway M, Burns TM; MG Composite and MG-QOL15 Study Group. MG-ADL: still a relevant outcome measure. Muscle Nerve 2011;44:727–731.

28. Burns TM, Conaway M, Sanders DB; MG Composite and MG-QOL15 Study Group. The MG Composite: a valid and reliable outcome measure for myasthenia gravis. Neurology 2010;74:1434–1440.

29. Wijeysundera DN, Johnson SR. How much better is good enough? Patient-reported outcomes, minimal clinically important differences, and patient acceptable symptom states in perioperative research. Anesthesiology 2016;125: 7–10.