Neurol Ther (2023) 12:1573–1590
https://doi.org/10.1007/s40120-023-00464-x

ORIGINAL RESEARCH

# Measuring Overall Severity of Myasthenia Gravis (MG): Evidence for the Added Value of the MG Symptoms PRO

Antoine Regnault · Thomas Morel · Christine de la Loge · Flora Mazerolle · Henry J. Kaminski · Ali A. Habib

Received: August 1, 2022 / Accepted: March 14, 2023 / Published online: May 11, 2023


## ABSTRACT

***Introduction***: Accurate measurement of myasthenia gravis (MG) severity is required for appropriate clinical monitoring of patients with MG and assessment of the benefit of new treatments in clinical trials. Our objective was to explore how MG severity can be measured and to determine how the newly developed MG Symptoms Patient-Reported Outcome (PRO) instrument complements the available measures of MG severity.

***Methods***: The conceptual coverage of the Quantitative MG (QMG), MG Composite (MGC), MG-Activities of Daily Living (MG-ADL), and MG Symptoms PRO was scrutinized against core symptoms of MG: muscle weakness in three muscle groups (ocular, bulbar, and respiratory), muscle weakness fatigability, and physical fatigue. Post hoc analyses of the MG0002 study, a Phase 2a clinical trial of rozanolixizumab in adults with moderate to severe generalized MG, included correlation and Rasch model analyses.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s40120-023-00464-x.

A. Regnault · F. Mazerolle
Modus Outcomes, A Division of THREAD, Lyon, France

A. Regnault
e-mail: antoine.regnault@threadresearch.com

F. Mazerolle
e-mail: flora.mazerolle@threadresearch.com

T. Morel (✉)
UCB Pharma SRL, Allée de la Recherche, 60, 1070 Brussels, Belgium
e-mail: Thomas.Morel@ucb.com

C. de la Loge
PCOM Analytics, Avallon, France
e-mail: cdelaloge@pcom-analytics.com

H. J. Kaminski
Department of Neurology and Rehabilitation Medicine, George Washington University, Washington, DC, USA
e-mail: hkaminski@mfa.gwu.edu

A. A. Habib
Department of Neurology, University of California, Irvine, CA, USA
e-mail: aahabib@hs.uci.edu

***Results***: The qualitative appraisal highlighted that only the MG Symptoms PRO captured physical fatigue. Data from 541 assessments (43 unique patients) were used for the analyses. Correlations ranged between 0.56 and 0.74 for the MG-ADL, QMG, MGC, and MG Symptoms PRO Muscle Weakness Fatigability score, and between 0.20 and 0.71 for the MG Symptoms PRO scores focusing on independent muscle groups. Analyses with the Rasch model estimated a meaningful continuum of severity of MG, including all items, except ocular muscles, from the four instruments. The QMG and MG Symptoms PRO had the broadest coverage of the MG severity continuum. Muscle fatigability and physical fatigue were more characteristic of low severity while bulbar weakness indicated more severe MG.

***Conclusion***: The severity of MG can be reflected in a meaningful continuum underpinned by the MG-specific outcome measures. Only ocular muscle manifestations were shown to reflect a possibly different facet of MG severity. With its modular nature and comprehensive content, the MG Symptoms PRO provides complementary information to the outcome measures widely used in MG.

***Trial Registration***: ClinicalTrials.gov identifier: NCT03052751.

## PLAIN LANGUAGE SUMMARY

Myasthenia gravis (MG) is a chronic disease affecting the communication between nerves and muscles. People with MG experience muscle weakness that worsens after activity and improves after rest. MG can affect different groups of body muscles (e.g., around the eyes, in the limbs, and face or throat).

We show that the various symptoms of MG can be used to summarize the overall severity of the disease: people with mild and moderate MG often report only the fast onset of weakness in their limb muscles and mild physical fatigue, while those with more severe MG report more severe fatigue and also difficulties associated with weakness in facial and throat muscles (leading to difficulty with swallowing or speaking) and in respiratory muscles (making breathing difficult). This ordering of MG manifestations will help create more accurate methods to assess the severity of MG that can be used to evaluate new treatments or to monitor patients in the clinic.

We also suggest that weakness of muscles around the eyes (leading to eyelid drooping or double vision) may represent a unique aspect of MG, and may not provide as much information to summarize the severity of MG as other symptoms. However, this needs further investigation as our study did not include participants who had weakness in eye muscles as their only symptom.

We also document the ability of the MG Symptoms Patient-Reported Outcome questionnaire, a new questionnaire completed by patients, to provide useful information for measuring the severity of MG.

**Keywords:** Clinician-reported outcomes; Myasthenia gravis; Outcome measure; Patient-reported outcomes; Symptoms

**Key Summary Points**

***Why carry out this study?***

Myasthenia gravis (MG) is a multi-faceted disease characterized by various cardinal symptoms, including muscle weakness and fatigability in various muscle groups (ocular, bulbar, respiratory, axial, and limb muscles), as well as physical fatigue; therefore, the measurement of overall severity of MG should account for these multiple manifestations.

These post hoc analyses of data from a Phase 2 clinical trial explored the definition of severity of MG based on all its manifestations, as captured by the existing outcome measures in MG. They also aimed to document the ability of a newly developed patient-reported outcome (PRO) measure, the MG Symptoms PRO, to bridge the gaps of the more established outcome measures in MG in their coverage of the overall MG severity.

***What was learned from the study?***

A meaningful continuum of overall MG severity was revealed, simultaneously reflecting muscle weakness in limbs, physical fatigue, bulbar, and respiratory muscle weakness. Ocular signs and symptoms appeared to represent a unique aspect in the overall picture of MG severity as they did not match well with this continuum.

The analyses also demonstrated the added value of the MG Symptoms PRO, which provides greater granularity and flexibility in the coverage of the overall MG severity compared with the more established outcome measures in MG.

## INTRODUCTION

Myasthenia gravis (MG) is a rare autoimmune disorder of the neuromuscular junction that is characterized by fluctuating fatigable weakness of voluntary muscles, including common symptoms like weakness of the ocular muscles, bulbar muscles, and generalized muscle weakness (limb, neck, and respiratory muscles). MG-related muscle weakness tends to increase during periods of activity and to improve after rest.

Clinical monitoring of patients with MG and demonstration of the benefit of new treatments in clinical trials both require that MG severity is measured accurately. This can involve different measurement approaches, such as the use of biomarkers or clinical outcome assessments [1]. The latter uses physician, trained evaluator, or patient assessments of the severity of typical clinical manifestations of MG: muscle weakness and fatigability in various muscle groups (ocular, bulbar, respiratory, axial, and limb muscles). The most widely accepted outcome measures of MG severity [2, 3] are the Quantitative MG score (QMG) [4], MG Composite score (MGC) [5], and MG-Activities of Daily Living (MG-ADL) [6]. These three outcome measures provide a summary of the severity of MG, combining the clinical manifestations of MG in a single number ("multi-component indices"), with the underlying assumption that the clinical manifestations for each muscle group reflect a common concept of overall MG severity.

While these instruments were all developed to assess the same concept, the reported correlation of the widely used outcome measures in MG is highly variable [5, 7–10], suggesting that they capture different facets of MG severity. The lack of consistently strong associations between these measures suggests both the need for a better understanding of what constitutes MG severity and for further research to develop better measures for use in clinical trials and clinical practice.

In addition to these widely used outcome measures, the MG Symptoms Patient-Reported Outcome (PRO) is a novel PRO measure specific to MG which was developed using the current best standards, which build on extensive direct

input from patients [11, 12], and which has demonstrated promising psychometric performances [13]. Integrating the patient voice in clinical decision-making, through PRO measures, has been identified as critical in many contexts, from clinical research [14–16] to routine clinical practice [17, 18], and even more specifically in rare diseases [19]. Thus, the MG Symptoms PRO has been developed to exclusively capture the symptoms of MG as perceived by the patients (and not as evaluated by clinicians). A patient-reported measure also allows the assessment of the severity of MG symptoms experienced by the patients over a longer period (and not only at the time of the examination). Contrary to the other outcome measures of MG severity, the MG Symptoms PRO uses independent measures of severity for each symptom group, with a separate scale for muscle weakness in the sentinel muscle groups (ocular, bulbar, respiratory), as well as a scale for muscle fatigability of all muscle groups [13]. It also includes a standalone scale for physical fatigue, including concepts not only of general energy and stamina but also physical manifestations of fatigue in the form of heaviness and weakness in the limbs and body in general, which has recently been flagged as an important symptom for patients with MG [2, 20]. This modular approach, focusing individually on each different symptom, enables the characterization of clinical manifestations of MG, which may provide a more versatile—and sensitive—measurement system for the severity of MG.

Our objective was to use MG-specific outcome measures data collected in a Phase 2 clinical trial to gain a better understanding of how the typical clinical manifestations of MG may inform the overall severity of the disease, which will eventually allow better measurement of MG severity. Moreover, in this research, we explored how the modular approach of the MG Symptoms PRO complements the most widely accepted outcome measures to assess MG severity.

# MATERIAL AND METHODS

## Study Sample

The analyses were conducted using data from the MG0002 study, a Phase 2a, multicenter, randomized, double-blind, placebo-controlled study investigating the clinical efficacy of rozanolixizumab in adults with moderate to severe generalized MG [21].

The full analysis set (FAS) of the MG0002 study included 43 randomized subjects treated with rozanolixizumab 7 mg/kg or placebo at 17 sites in the USA, Canada, and Europe. A total of 542 individual assessments of the MG outcome measures were collected in the study over 13 visits during a 50-day treatment period, plus a 49-day observation period. All analyses reported here were performed using the FAS, independent of the treatment group (i.e., data from the treatment arms were pooled together for these analyses).

## Standard Protocol Approvals, Registrations, and Patient Consents

The MG0002 trial (ClinicalTrials.gov Identifier: NCT03052751) was performed in accordance with the principles of the Declaration of Helsinki and the International Conference on Harmonisation Guidance for Good Clinical Practice. The study protocol, amendments, and patient-informed consent were reviewed by national, regional, or independent ethics committees or institutional review boards. Written informed consent was obtained from all patients.

Full details on the study design and results of the MG0002 study have been published previously [21].

## Outcome Measures of MG Symptom Severity

### *QMG Scale*

The QMG scale is a clinical outcome assessment of MG severity, which combines both clinician-reported outcome (ClinRO) items and performance outcome (PerfO) items that reflect the

patient's ability to perform specific tasks at the time of the assessment [4]. It includes 13 items covering ocular, bulbar, respiratory, and limb function, all assessed using 4-point response scales. These items combine timed tests of endurance (in seconds), grip strength (in kg), and forced vital capacity (in % predicted) all transformed into ordinal scores, with ordinal ratings of facial muscle strength and swallowing. A total score ranging from 0 to 39 can be calculated, with higher values indicating more severe disease. The QMG score showed correlations ranging from weak to moderate with markers of severity of MG, such as Myasthenia Gravis Foundation of America (MGFA) classification and electrophysiological parameters [22].

### *MGC Score*

The MGC score is a clinical assessment combining ClinRO, PerfO, and PRO items [7]. It includes 10 items assessing ocular function [2], muscle strength [3], patient-reported bulbar function [4], and breathing [1]. The items do not ask about a specific period of reference; rather, they reflect the targeted sign at the time of the examination (for ClinRO and PerfO items) and the targeted symptom in general, based on patient history (for PRO items). All items are assessed using 4-point response scales. A scoring algorithm assigning weights to each response level of the items leads to a total score that ranges between 0 and 50, with higher values indicating more severe disease [23]. The MGC scale showed adequate reliability in a cohort of MG patients [5].

### *MG-ADL*

The MG-ADL is a PRO instrument developed to assess the symptoms and activities in MG [6]. It includes 8 items covering ocular, bulbar, respiratory, and limb symptoms of MG. Similar to the QMG and MGC, no reference period is specified to rate the items, so the responses are assumed to reflect the targeted symptom at the time of the assessment. All items are graded on a 0–3 response scale, and a total score ranging from 0 to 24 can be calculated by summing them, with higher values indicating more severe MG. MG-ADL showed moderate to strong correlations with other MG-specific outcome measures, depending on the studies and good test–retest reliability [6, 10].

### *MG Symptoms PRO instrument*

The MG Symptoms PRO is a measure of the proximal symptoms of MG and was developed for use in clinical trials using a patient-centered mixed methods psychometric approach [13]. A draft version of the instrument was included in the MG0002 study for refinement of the item sets, development of a scoring system, and evaluation of measurement properties. Previous analyses of these data showed that this version of the MG Symptoms PRO had promising psychometric performance both in Rasch measurement theory and classical test theory frameworks [13].

Four additional items were developed after the MG0002 study to bridge conceptual gaps in terms of ocular muscle weakness (items related to eye movements: "difficulty moving eyes from side to side" and "difficulty moving eyes up and down") and respiratory muscle weakness, that were identified based on the qualitative and quantitative analyses of data collected within the study. The version resulting from the MG0002 study includes 42 items that cover five cardinal symptomatic concepts of MG, which are each independently assessed by a subscale: three related to muscle weakness—Ocular Muscle Weakness (3 items), Bulbar Muscle Weakness (10 items), Respiratory Muscle Weakness (3 items)—plus two additional scores related to Physical Fatigue (15 items) and Muscle Weakness Fatigability (9 items). All items ask about the experience of patients over the previous 7 days. A score is calculated for each subscale ranging from 0 to 100, with a higher score indicating higher severity of the symptom considered.

## Analyses

### *Qualitative Mapping of MG Outcome Measure Content*

The items of each MG symptom severity measure (MG-ADL, QMG, MGC, and MG Symptoms PRO) were reviewed and qualitatively mapped

to five core symptoms of MG identified in the conceptualization of the MG patient experience from the development of the MG Symptoms PRO [13]: muscle weakness in three specific muscle groups (ocular muscles, bulbar muscles, and respiratory muscles), physical fatigue (defined as general lack of energy and weakness in the limbs and body), and muscle weakness fatigability (defined as the patient perception of increasing weakness with enduring activity) in any muscle group. The mapping was performed by the authors with the contribution of the lead researcher for the development of the conceptual model of the experience of patients with MG (Dr. Sophie Cleanthous). The objective of this qualitative appraisal was to ascertain and compare the comprehensiveness of the conceptual coverage of each outcome measure and to identify any gaps in their conceptual coverage.

### *Statistical Analyses*

The correlation between each pair of MG outcome measure scores at baseline was estimated using the Spearman rank-order correlation coefficient. Correlations between 0.3 and 0.5 were considered weak (between 10% and 25% of shared variance), moderate between 0.5 and 0.7 (between 25% and 50% of shared variance), and strong above 0.7 (more than 50% of shared variance). The MG outcome measures are all intended to capture some form of MG severity, and should therefore have at least moderate correlations. However, a moderate correlation between two measures would indicate that a fair amount of information is not shared by the two measures, suggesting that they measure different characteristics of MG severity. Additionally, MG-ADL and MG Symptoms PRO are pure PRO measures, while the QMG and MGC are composite measures combining ClinRO, PerfO, and PRO (for the MGC) measures. Additionally, the MG Symptoms PRO captures the experience of patients in their daily life over a period of 7 days, whereas the MG-ADL, QMG, and MGC focus on the signs and symptoms observed for either an unspecified reference period or at the time of the examination. For these reasons, the correlations of the MG Symptoms PRO and, to a lesser extent, the MG-ADL with the QMG and MGC were expected to be lower than the correlation between the QMG and MGC.

Analyses were conducted using the Rasch model to describe how the clinical manifestations of MG are distributed over the underlying continuum of overall MG severity and to evaluate each outcome measure's coverage of this continuum of overall MG severity created in the analysis, reflecting the information from all items from the four MG severity measures. The Rasch model is a mathematical model that puts subjects in the study sample and the items of the instrument under scrutiny on a common metric corresponding to the underlying continuum of the measured construct [24–26]. Our analyses included all items from the different outcome measures of MG symptom severity in a single model, hence creating a continuum of overall MG symptom severity. The fundamental initial hypothesis was therefore that the severity of each typical clinical manifestation was reflective to some extent of a higher-order concept of the overall severity of MG. The QMG, MGC, and MG-ADL are three composite scores that were designed to reflect the overall severity of MG across the various manifestations of the disease. It can therefore be reasonably assumed that their items can be mapped on a single metric. The MG Symptoms PRO was developed to capture the severity of the different symptomatic components of MG independently: ocular muscle weakness, bulbar muscle weakness, respiratory muscle weakness, muscle weakness fatigability, and physical fatigue. While these item sets were not originally planned to be combined into an overall measure of MG severity, they target concepts that are similar to those included in the QMG, MGC, and MG-ADL. Therefore, they should all be related to overall MG severity and theoretically warrant the mapping of these items to the common metric in the Rasch model.

We used the Rasch model as the structural probabilistic model for exploring how the typical manifestation of MG can define a meaningful continuum of overall severity, and how the items composing each outcome measure of MG cover this continuum. For this purpose, we observed how the estimates from the Rasch model for groups of items either referring to the

same core symptom of MG or coming from the same outcome measure were distributed over the continuum provided by the Rasch model.

In addition, the application of the Rasch model also provides several criteria for the appraisal of an item set intended to be used for measurement purposes. In our context, we focused our interpretation on the following properties, which are the most directly relevant to the characterization of the continuum by the items included in the model:

- Item fit to the Rasch model: observed responses are compared with expected responses by the model using statistical indices and a graphical examination of the item characteristic curve [27]. Statistical indices include standardized fit residuals, which are recommended to lie in the range $-2.5$ to $+2.5$ [26], and chi-square tests. Item misfit, especially underdiscrimination (i.e., standardized residuals greater than 2.5), would highlight items that may not reflect the same concepts as the full item set.
- Targeting of the items to the study participants: the distribution of person and item parameters over the common continuum are compared to flag any mismatch that would indicate that the items from the MG outcome measures were not fully appropriate to capture the severity of the patients in the sample.

We also considered the following other properties of the Rasch model which may be less critical for our objective, but are still informative on the measurement performance of the item set:

- Local dependence: standardized residual correlations between each pair of items should be low (typically below 0.3). Higher correlations could indicate pairs of items measuring a somewhat different construct than the overall item set.
- Suitability of the response options of each item: for each item, the ordering of successive response categories of the response scale will be considered: "disordered thresholds" would indicate a response scale that does not work as intended (i.e., increasing response not reflecting increasing levels of MG symptom severity) [28].

The Rasch model was applied to "stacked data" from the 13 visits of the study, with 43 unique patients, and approximately 559 measurement points were expected to be available, allowing for reliable estimations from the Rasch model. The model was initially applied to all unique items from the MG-ADL, QMG, MGC, and MG Symptoms PRO (items addressing talking, chewing, swallowing, and breathing that are virtually identical in the MG-ADL and MGC were included just once). Based on the results of the initial model, an alternative analysis was conducted excluding ocular muscle weakness items, which were not fitting the Rasch model well.

Data preparation, descriptive analyses, and correlation were performed using SAS v9.4 (SAS Institute, Cary, NC, USA), and the Rasch model was estimated using RUMM 2030 software (RUMM Laboratory, Perth, Australia).

## RESULTS

### Comparison of Conceptual Coverage of MG Symptom Severity Measures

The qualitative item-to-concept mapping of the MG symptom severity measures (Table 1) showed that ocular muscle weakness, bulbar muscle weakness, and respiratory muscle weakness are covered by all four measures of MG symptom severity. In contrast, muscle weakness fatigability is captured across all targeted muscle groups by the MG Symptoms PRO and QMG (except respiratory muscles), but only for neck by the MGC and for limbs by the MG-ADL. Finally, the MG Symptoms PRO covers multiple facets of physical fatigue (general fatigue, limb and axial weakness, limb heaviness), while the QMG and MGC only assess limb weakness and the MG-ADL does not assess this symptom at all. This qualitative mapping also emphasizes that the MG Symptoms PRO includes significantly more items than the other three outcome measures; therefore, it provides greater

**Table 1** Number of items covering each of the core symptoms of MG for each of the MG symptom severity measures

| Core MG symptoms | Subdomain | MG-ADL | MGC | QMG | MG Symptoms PRO |
|---|---|---|---|---|---|
| Ocular muscle weakness | Blurry vision | | | | 1 |
| | Double vision | 1 | 1 | 1 | 1 |
| | Eye movements | | 1 | | (2[a]) |
| | Eyelid drooping | 1 | 1 | 1 | 1 |
| Bulbar muscle weakness | Chewing | 1 | 1 | | 1 |
| | Facial muscle weakness | | | 1 | |
| | Facial/mouth drooping | | | | 1 |
| | Liquid control in mouth | | | | 1 |
| | Speech/voice problems | 1 | 1 | | 5 |
| | Swallowing | 1 | 1 | 1 | 2 |
| Respiratory muscle weakness | Breathing problems | 1 | 1 | 1 | 1(+ 2[a]) |
| Muscle weakness fatigability | Ocular muscles | | | | 2 |
| | Bulbar muscles—speech/voice | | | 1 | 2 |
| | Bulbar muscles—chewing/swallowing | | | | 2 |
| | Neck muscles | | 1 | 1 | |
| | Lower limb muscles | 1[b] | | 2 | 1 |
| | Upper limb muscles | 1[b] | | 2 | 1 |
| | Respiratory muscles | | | | 1 |
| Physical fatigue | General feeling of physical fatigue | | | | 6 |
| | Lower limb weakness/heaviness | | 1 | | 2 |
| | Upper limb weakness/heaviness | | 1 | 2 | 2 |
| | Whole body weakness/heaviness | | | | 2 |
| | Neck weakness | | | | 1 |
| | Movement limitations due to fatigue | | | | 2 |

[a]Additional items have been developed to bridge conceptual gaps in terms of ocular and respiratory muscle weakness during the development process of the MG Symptoms PRO but were not included in the MG0002 study
[b]MG-ADL items about "impairment of ability to brush teeth or comb hair" and "impairment of ability to arise from a chair" were categorized as indicators of muscle weakness fatigability for upper and lower limbs, respectively, while they can also be considered as assessment of impact of muscle weakness

granularity in terms of coverage of the different manifestations of MG, even within each symptom category. Additionally, the period of reference for the evaluation of the signs and symptoms is clearly specified in the instructions of the MG Symptoms PRO (over the 7 days preceding the assessment), whereas the timeframe for the evaluation of other measures is not as clearly defined. Therefore, the MG Symptoms PRO captures the severity of MG symptoms experienced by patients over a 7-day period, while the other instruments may not

capture the daily variations in the patient experience as fully.

## Sample Description

Forty-three patients were included in the full analysis set of the MG0002 study (Table 2). Almost half were in MGFA Class III (i.e., moderate weakness affecting muscles other than ocular muscles; may also have ocular muscle weakness of any severity; 49%), with 35% in Class IIIa (signs predominantly affecting limb, axial muscles) and 14% in Class IIIb (signs predominantly affecting oropharyngeal, respiratory muscles).

The distribution of responses to MG-ADL and MGC items suggested a possible floor effect, as most patients were rated in the two lowest categories, except for the ocular muscle items (Fig. 1). In contrast, the responses to the QMG items were less distributed towards the lower end of the response scale. In the MG Symptoms PRO, the patient responses to the ocular muscle weakness, bulbar muscle weakness, and respiratory muscle weakness items were mostly "none" or "mild" (Fig. 2, left panel). The responses to the items pertaining to physical fatigue were more frequently "a little of the time" and "some of the time", suggesting that physical fatigue was more frequently experienced in the study sample (Fig. 2, right panel). For the muscle weakness fatigability items, patients mostly endorsed "none of the time" or "a little of the time", except for the items related to limb muscle fatigability ("my arms felt weaker the longer I used them in my usual activities" and "my legs felt weaker the longer I used them in my usual activities"), which a substantial number of patients reported experiencing "some of the time", "most of the time", and even "all of the time".

## Correlations Between Measures of MG Symptom Severity

The composite measures of overall MG symptom severity (MG-ADL, QMG, and MGC) consistently had correlations between 0.57 and 0.71 (Fig. 3). They also had correlations between 0.56

**Table 2** Demographics and baseline disease characteristics of patients in MG0002 study [21]

| | **Total ($n$ = 43)** |
|---|---|
| Age (in years), mean (SD) | 51.9 (15.1) |
| Gender, female—$n$ (%) | 27 (62.8%) |
| Country—$n$ (%) | |
| Canada | 11 (25.6%) |
| Belgium | 10 (23.3%) |
| United States of America | 9 (20.9%) |
| Denmark | 7 (16.3%) |
| Germany | 3 (7.0%) |
| Czech Republic | 2 (4.7%) |
| Spain | 1 (2.3%) |
| Auto-antibody class—$n$ (%) | |
| AChR | 40 (93.0%) |
| MuSK | 1 (2.3%) |
| QMG score at baseline, mean (SD) | 15.6 (3.9) |
| MGC score at baseline, mean (SD) | 15.6 (6.2) |
| MG-ADL score at baseline, mean (SD) | 7.1 (3.1) |
| MG Symptoms PRO scores at baseline, mean (SD) | |
| Bulbar Muscle Weakness | 24.0 (17.7) |
| Muscle Weakness Fatigability | 43.8 (22.4) |
| Physical Fatigue | 49.2 (21.7) |
| Ocular Muscle Weakness | 31.5 (20.7) |
| Respiratory Muscle Weakness | 38.0 (25.8) |
| MGFA classification—$n$ (%) | |
| Class I | 0 (0.0%) |
| Class IIa | 7 (16.3%) |
| Class IIb | 12 (27.9%) |
| Class IIIa | 15 (34.9%) |
| Class IIIb | 6 (14.0%) |
| Class IVa | 2 (4.7%) |
| Class IVb | 1 (2.3%) |

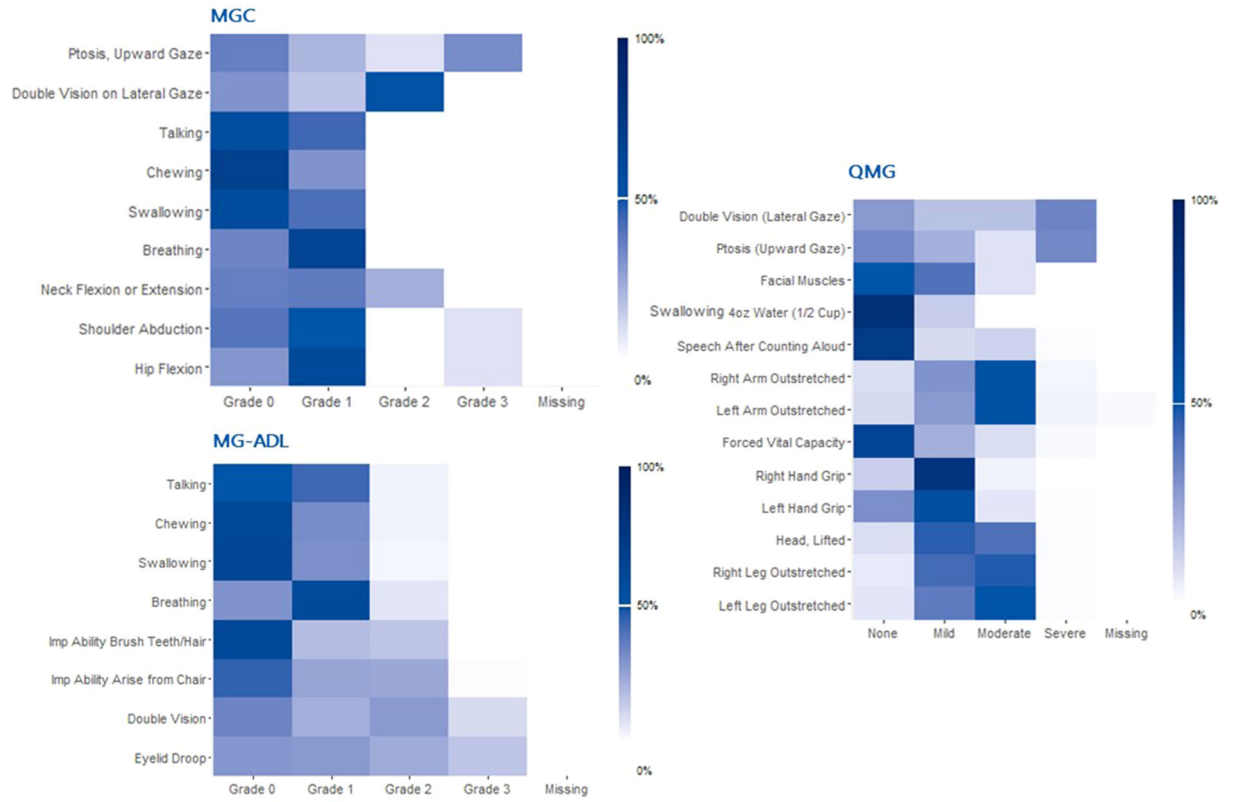


**Fig. 1** Distribution of the responses to the items of the Myasthenia Gravis Composite (MGC), Myasthenia Gravis-Activities of Daily Living (MG-ADL), and Quantitative Myasthenia Gravis (QMG) scores. ($n = 541$ assessments from 43 unique patients). *Darker colors* indicate higher percentages of patients who endorsed the response. Grades 0–3 refer to the 4-level response categories of MGC/MG-ADL items

and 0.74 with the MG Symptoms PRO Muscle Weakness Fatigability score, which includes items covering all muscle groups.

Overall, the correlations of MG Symptoms PRO scores with the composite measures of overall MG symptom severity were weak to moderate and systematically lower than the correlations observed between the composite measures. Only the MG Symptoms PRO Physical Fatigue score showed lower levels of correlations, which was expected as the composite measures do not include items assessing fatigue.

## Characterization of the Overall Severity of MG

### *Rasch Model of the Full Item Set from the Four MG Symptom Severity Measures*

The application of the Rasch model to all items from the four MG severity measures showed a good match between the items and the study sample, with a slightly broader coverage of the items toward the direction of higher severity (see full RMT outputs in supplementary material 1). This suggested that, overall, the MG severity measures cover the full breadth of MG severity and even include items that capture more severe levels of disease than those observed in the MG0002 study. Importantly, the major finding of this first application of the Rasch model was that the ocular muscle weakness items were not cohesive with the items reflecting the other symptoms, as they showed strong misfit to the Rasch model and strong correlations between their standardized residuals (see full RMT outputs in supplementary material 1).

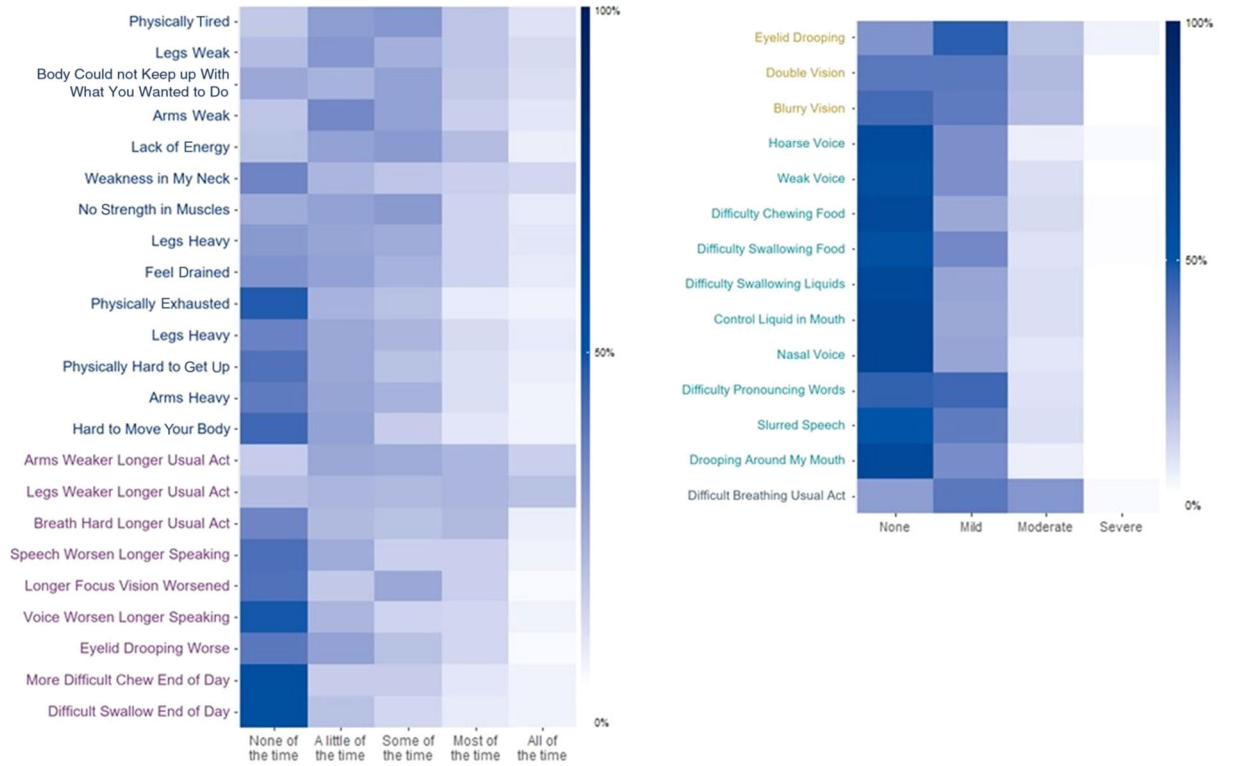


**Fig. 2** Distribution of the responses to the items of the Myasthenia Gravis Symptoms Patient-Reported Outcomes (MG Symptoms PRO) instrument ($n = 541$ assessments from 43 unique patients). *Darker colors* in the cell indicate higher percentages of patients who endorsed the response. *Item label colors* indicate the core MG symptom to which the item relates: *blue* physical fatigue; *purple* muscle weakness fatigability; *yellow* ocular muscle weakness; *turquoise* bulbar muscle weakness; *gray* respiratory muscle weakness

### *Rasch Model Excluding Ocular Muscle Weakness Items*

The Rasch model was then applied to all items from the four MG symptom severity measures, excluding items about ocular muscle weakness (see full RMT outputs in supplementary material 2). The good targeting of the item set to the study sample was maintained. Fewer items showed underdiscrimination, mostly from MGC: "hip flexion", "shoulder abduction", and "neck flexion or extension" did not discriminate the level of overall MG severity as much as expected. In other words, patients who presented lower severity of MG according to their profile of MG outcome ratings tended to have a higher rating than expected for these three items, and patients with higher levels of severity tended to have less severe ratings than expected. This may indicate that these items do not perfectly line up with the MG severity captured by the other items from the outcome measures. Other items had standardized residual statistics indicating misfit, but overdiscriminating, indicating that these items were very specific to a specific location of the continuum, which is not a major issue for the characterization of the continuum based on the item content. Six out of eight MG-ADL items showed "disordered thresholds", suggesting possible issues with the format of the response scale of the MG-ADL and several pairs of items showed some residual correlations greater than 0.3, almost systematically from a same instrument or a same symptom group. These minor deviations to the model did not critically preclude the interpretability of the continuum using the item parameter estimates from the model.

| Score | Ocular Muscle Weakness | Bulbar Muscle Weakness | Respiratory Muscle Weakness | Physical Fatigue | Muscle Weakness Fatigability | MG-ADL | MGC |
|---|---|---|---|---|---|---|---|
| MG Symptoms PRO Bulbar Muscle Weakness | 0.59 [0.35 ; 0.76] | | | | | | |
| MG Symptoms PRO Respiratory Muscle Weakness | 0.31 [0.01 ; 0.56] | 0.46 [0.18 ; 0.67] | | | | | |
| MG Symptoms PRO Physical Fatigue | 0.26 [-0.05 ; 0.52] | 0.26 [-0.04 ; 0.52] | 0.20 [-0.11 ; 0.47] | | | | |
| MG Symptoms PRO Muscle Weakness Fatigability | 0.48 [0.20 ; 0.68] | 0.71 [0.51 ; 0.83] | 0.32 [0.01 ; 0.56] | 0.46 [0.17 ; 0.66] | | | |
| MG-ADL | 0.63 [0.40 ; 0.78 | 0.62 [0.38 ; 0.77] | 0.35 [0.05 ; 0.59] | 0.40 [0.11 ; 0.62] | 0.74 [0.55 ; 0.85] | | |
| MGC | 0.39 [0.10 ; 0.62] | 0.50 [0.23 ; 0.70] | 0.37 [0.07 ; 0.60] | 0.20 [-0.11 ; 0.47] | 0.56 [0.30 ; 0.74] | 0.57 [0.32 ; 0.74] | |
| QMG | 0.51 [0.25 ; 0.70] | 0.46 [0.17 ; 0.66] | 0.31 [0.00 ; 0.55] | 0.34 [0.03 ; 0.57] | 0.58 [0.33 ; 0.75] | 0.64 [0.42 ; 0.79] | 0.71 [0.52 ; 0.83] |

**Fig. 3** Correlation between Myasthenia Gravis Symptoms Patient-Reported Outcome (MG Symptoms PRO) scores, Myasthenia Gravis-Activities of Daily Living (MG-ADL) score, Myasthenia Gravis Composite (MGC) score, and Quantitative Myasthenia Gravis (QMG) score, with 95% confidence interval, at baseline of study MG0002 ($n = 43$). Spearman correlation coefficient; in white, low correlations ($< 0.30$); in light blue, weak correlations (0.30–0.49); in blue, moderate correlations (0.50–0.69); in dark blue, high correlations ($\geq 0.70$)

This analysis with the Rasch model allowed the creation of a unique continuum of MG symptom severity, reflecting all the components (items) of the four outcome measures included in the model. Figure 4 illustrates this continuum going from the mildest severity (on the left) to the highest severity (on the right). In Fig. 4, the continuum is expressed in the internal metric of the Rasch model (in "logits"), which does not have a direct interpretation (e.g., the '0' comes from an estimation constraint and does not carry any meaning). The distribution of all observations from the full study sample over the overall MG symptom severity continuum is represented in the upper panel of the figure, while the lower panel shows how the items of each outcome measure are reflective of various levels of overall severity of MG. Indeed, each box in the lower panel along the *x*-axis represents an item parameter estimate from the Rasch model.

By construction, an item parameter in this model is the location on the continuum where two consecutive response categories of a given item are equally probable ("item thresholds"), which is indicative of a level of severity that can be discriminated by the response to this item. The boxes, therefore, represent the coverage of the continuum by the items composing the scales and allow the characterization of the continuum. Items from the QMG and MG Symptoms PRO cover the broader range of the full MG symptom severity continuum. When considering all items from its five scales together, the MG Symptoms PRO did not have any gaps in the coverage of the continuum, while the QMG had a few gaps in the middle of the continuum. The MG-ADL and, to a lesser extent the MGC, did not include items covering the mildest severity of MG.

Additionally, a color code was used for the items (squares) in the figure to represent each

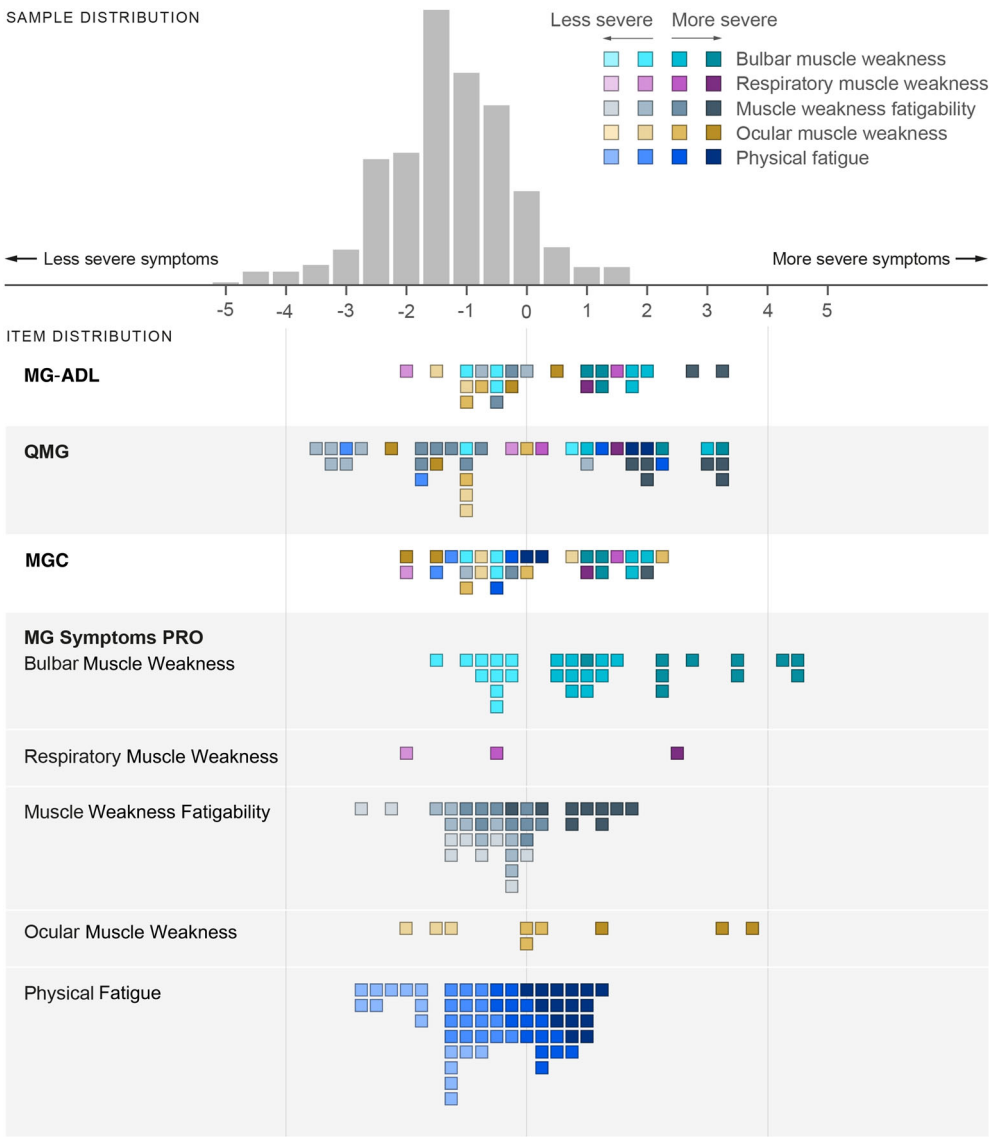


**Fig. 4** Distribution of MG0002 individual assessments (*upper panel in gray*) and MG symptom severity outcome measure items on MG symptom severity measurement continuum ($n = 541$ assessments from 43 unique patients). MG symptom severity continuum was estimated from the Rasch model analysis of the full item set, excluding ocular muscle weakness. Ocular muscle weakness items were mapped to the continuum for full disclosure; their parameters were estimated in the metric of the model by applying a post hoc "anchored" Rasch model. The *upper panel* shows the distribution of the 541 assessments over the MG severity continuum; in the *lower panel*, each *box* represents an "item threshold" (i.e., the point of the continuum where the most probable response between two adjacent response categories for an item changes)

cardinal symptom of MG, and therefore characterize the typical symptoms for various levels of severity of MG. Muscle weakness fatigability items appeared to be the ones able to characterize milder severity of MG, while bulbar muscle weakness items were more characteristic of more severe MG. Respiratory muscle weakness and ocular muscle weakness items were spread over a wide range of severity but did not cover the mildest levels of severity.

Finally, physical fatigue items ranged from the milder levels of MG to severe ones (but did

not reach the most severe extremity of the continuum), providing good coverage of the levels of severity observed in the MG0002 study (the figure included in supplementary material 3 provides an alternative visualization of the same data that emphasizes the coverage by each symptom). Of note, the ocular muscle weakness items were mapped on the metric of this second model in a post hoc "anchored" Rasch model (i.e., a model in which all items were included, but only the parameters for the ocular muscle weakness items were estimated, while all others were set to the value estimated in the second model). By doing so, the ocular muscle items do not contribute to the creation of the overall MG severity continuum, but we can still have a sense of how they relate to the severity reflected by the other key symptoms of MG, hence providing a more comprehensive picture of the coverage of the MG symptom severity by all four measures.

## DISCUSSION

The newly developed MG Symptoms PRO complements the set of currently widely used measures of disease activity and severity of MG. It provides greater granularity for the assessment of the cardinal symptoms of MG (weakness of the ocular muscles, bulbar muscles, and generalized muscle weakness in limb, neck, and respiratory muscles), providing more extensive coverage of the severity of MG, capturing even for mild symptoms that may be imperfectly captured by other measures. It also fills an important gap in the measurement of fatigue. Finally, its modular nature, in which each symptom is assessed independently, allows the possibility of a more meaningful interpretation by focusing on the symptoms that are particularly relevant for a patient at a specific level of MG severity.

Our analyses identified a continuum underpinned by all non-ocular items of the outcome measure data collected in the MG0002 study, which we propose constitute a reasonable representation of the overall severity of MG. The application of the Rasch model assumed a hypothesized continuum of overall severity of MG that is reflected by each of the manifestations of MG captured by the outcome measures included in the analysis. Overall, the clinical and conceptual underpinnings that prompted this hypothesis were substantiated by our quantitative findings. The continuum exposed by our analysis reflects the signs and symptoms captured by all the commonly used outcome measures in MG (MG-ADL, QMG, and MGC), as well as the newly developed MG Symptoms PRO in a clinically meaningful way. In this continuum of MG severity, the lowest levels of severity were characterized by muscle weakness in limbs and physical fatigue, while bulbar muscle weakness manifestations were estimated to be typical of most severe MG. Previous research applying the Rasch model independently to the MG-ADL, QMG, and MGC items also indicated that they could be mapped on a continuum of MG severity [6, 29, 30]. Our analysis with the Rasch model concluded that the severity of ocular muscle weakness signs and symptoms (ptosis, diplopia, blurry vision) was not accurately reflecting the severity of MG characterized according to the other muscle groups. Previous applications of the Rasch model to individual MG-specific outcome measures had already identified that the ocular items did not fit well with the others [29], and that their contribution to the characterization of severity was unstable across the different analyses, as they were specific to either very mild [6], very mild and very severe MG [30], or very mild and moderate MG [29]. This singularity of ocular muscle weakness in the assessment of overall MG severity can probably be related to the differential susceptibility of ocular muscles to the autoimmune process of MG, and to neuromuscular transmission disorders generally [31, 32].

Our analyses also provided useful insight into both the commonly used measures in MG and the newly developed MG Symptoms PRO. The QMG and MG Symptoms PRO had the widest coverage of the overall MG symptom severity, and were able to accurately capture mild, moderate, and severe MG. In contrast, the MG-ADL and MGC did not effectively capture mild MG, which confirms previous findings documenting the floor effect seen with MG-ADL [3, 9].

The correlations observed between the MG-ADL, QMG, and MGC scores were moderate (ranging from 0.5 to 0.8), which show that, while there is common variability (i.e., common information) captured by these scores, they also have substantial differences in the concept they measure. The levels of correlation estimated here are in line with the wide range of correlations previously observed for these instruments, ranging from as low as 0.33 (between QMG and MG-ADL) to 0.88 (between QMG and MGC, which share some very similar items) [5, 7–10]. In parallel, the MG Symptoms PRO scores assessing weakness in the different muscle groups (Ocular, Bulbar, and Respiratory Muscle Weakness) and Physical Fatigue had lower correlations with the commonly used measures of overall MG severity, which was expected as these scores only capture a specific facet of MG. This finding consolidates the modular approach of MG Symptoms PRO as a complementary way to measure MG symptom severity. Our analysis showed that the MG Symptoms PRO scales cover a wide range of overall MG severity (with scales including all core symptoms of MG, especially a comprehensive assessment of physical fatigue) and, at the same time, provide independent assessments of the severity of each symptom relevant to patients with MG.

The downside of this finer conceptual granularity is the length of the instrument. A shorter version could be developed by deleting some redundant items (as flagged by the conceptual mapping), but the modular nature of the instrument also enables another and perhaps more promising solution: it is possible to use only the items relevant to a specific targeted population, or to the individual patient. For example, patients with severe MG will probably find the Bulbar Muscle Weakness domain the most relevant to consider, while for patients with a milder presentation of MG, the Physical Fatigue domain is most likely better suited.

Another noteworthy finding of our analyses was that physical fatigue, which has been identified as an important symptom in MG [13, 20], was instrumental in the characterization of a wide range of severity from milder to moderate forms of MG. As the MG Symptoms PRO is the instrument measuring physical fatigue the most comprehensively, its coverage of the milder end of the continuum was improved compared to all other instruments, including the QMG. This feature of the MG Symptoms PRO can be especially beneficial in the context of clinical trials, where the objective is to detect the improvement of MG symptoms. Fatigue in MG includes more than physical fatigue: for example, general fatigue or mental fatigue have been reported [33, 34], including in the development of the MG Symptoms PRO [13]. Further research would be needed to explore how these other facets of fatigue fit into the overall severity of MG.

While our findings are an important addition to the understanding of MG severity measurements, there are some limitations to be considered. The first limitations of our analyses are related to sampling. Our analyses were conducted using data from a single study that included only 43 patients; however, multiple data collected for each patient were included in the analysis, leading to more than 500 unique assessments of each instrument. Additionally, the sample was from a clinical trial, so it was a selected sample of patients and may not be representative of a general MG population (e.g., participants in the clinical trial may have more severe MG). In particular, the sample excluded patients with ocular symptoms only; consequently, the findings may not apply to patients with ocular MG. Further research with larger samples, maybe with milder MG and with ocular symptoms only, would be warranted to confirm our results. Another caveat is that some deviations from the Rasch model were observed. While these deviations may be important in most applications of the Rasch model (e.g., for creating a new measure), they are less important in our context: our interpretation was on trends for groups of items (not individual items) to characterize the expression of the underlying continuum. Nonetheless, the estimates from the Rasch model may be marginally impacted by these deviations, so the invariance of the item parameter estimates in other samples might be further explored. Another limitation of our analysis is that we did not conduct any longitudinal analysis to explore how the items

from the different outcome measures change over time relative to one another. Hence, we did not explore the invariance of MG severity measurement over time when using MG-specific outcome measures. Finally, our analyses only included the MG-specific outcome measures available in the MG0002 study. Further research should investigate how other existing instruments fit in terms of coverage of the MG symptom severity continuum that we uncovered in our analysis. Specifically, the recently developed MG Impairment Index [35] was not included in our analyses. It was collected in a subgroup of patients in the MG0002 study, but this sample was so small that it could not reasonably be added to our analyses.

## CONCLUSION

In conclusion, our analyses of the MG-ADL, QMG, MGC, and MG Symptoms PRO data with the Rasch model [1] formed a meaningful continuum of overall MG severity reflecting all manifestations of the disease except ocular muscle manifestations [2], highlighted the singularity of ocular signs and symptoms in the overall picture of MG severity [3], confirmed some limitations of the widely used measures of MG severity (physical fatigue not being fully captured, moderate correlations between the various measures, floor effect of MG-ADL), and [4] demonstrated the added value of the MG Symptoms PRO to the available arsenal of measures of MG severity.

## ACKNOWLEDGEMENTS

***Funding.*** This research and the MG0002 clinical study mentioned herein, as well as the journal Rapid Service Fee, were funded by UCB Pharma. Ali A. Habib and Henry J. Kaminski's contributions were supported by the MGNet, a member of the Rare Disease Clinical Research Network Consortium (RDCRN), NIH U54 NS115054.

***Medical Writing, Editorial, and Other Assistance.*** Sophie Cleanthous (Modus Outcomes) provided support with the conceptual appraisal of outcome measures and interpretation of results. Lori Bacarella (Modus Outcomes) provided editorial support funded by UCB Pharma. Veronica Porkess, of UCB Pharma, provided editorial review and assistance. Irene de la Torre Arenas, of UCB Pharma, provided data visualization support.

***Author Contributions.*** Antoine Regnault designed the analysis and contributed to the conduct and interpretation of the analyses. Thomas Morel contributed to the design and interpretation of the analyses. Christine de la Loge contributed to the design and interpretation of the analyses. Flora Mazerolle contributed to the conduct and interpretation of the analyses. Henry J. Kaminski contributed to the interpretation of the analyses. Ali A. Habib contributed to the interpretation of the analyses. All others reviewed and had substantial contribution to the manuscript content.

***Disclosures.*** Antoine Regnault; Flora Mazerolle: Employees of Modus Outcomes, a patient-centered outcome research consultancy that received payment from UCB to conduct this research. Henry J. Kaminski: Principal investigator of the Rare Disease Clinical Research Network (U54 NS115054); a consultant for Alnylam Pharmaceuticals, Caballata Bio, Roche, and UCB Pharma; and is CEO and CMO of ARC Biotechnology, LLC, based on US Patent 8961981. Thomas Morel; Employee and shareholder of UCB Pharma. Christine de la Loge: Contractor to UCB Pharma. Ali A. Habib: Has received research support from argenx, Alexion, VielaBio, UCB Pharma, Genentech, Regeneron, and Sanofi. He has received consulting fees/honoraria from argenx, Alexion, and UCB Pharma.

***Compliance with Ethics Guidelines.*** The MG0002 trial (ClinicalTrials.gov Identifier: NCT03052751) was performed in accordance with the principles of the Declaration of Helsinki and the International Conference on Harmonisation Guidance for Good Clinical

Practice. The study protocol, amendments, and patient-informed consent were reviewed by national, regional, or independent ethics committees or institutional review boards. Written informed consent was obtained from all patients.

***MG Symptoms PRO Copyright and Use.*** The copyright to the MG Symptoms PRO is owned by UCB Biopharma SRL, Belgium.

***Data Availability.*** Underlying data from this manuscript may be requested by qualified researchers six months after rozanolixizumab approval in the US and/or Europe, or global development is discontinued, and 18 months after trial completion. Investigators may request access to anonymized individual patient-level data and redacted trial documents which may include: Analysis-ready datasets, study protocol, annotated case report form, statistical analysis plan, dataset specifications, and clinical study report. Prior to use of the data, proposals need to be approved by an independent review panel at www.Vivli.org and a signed data sharing agreement will need to be executed. All documents are available in English only, for a prespecified time, typically 12 months, on a password protected portal.



## REFERENCES

1. Benatar M, Sanders DB, Burns TM, Cutter GR, Guptill JT, Baggi F, et al. Recommendations for myasthenia gravis clinical trials. Muscle Nerve. 2012;45(6):909–17.

2. Thomsen JLS, Andersen H. Outcome measures in clinical trials of patients with myasthenia gravis. Front Neurol. 2020;11.

3. Barnett C, Herbelin L, Dimachkie MM, Barohn RJ. Measuring clinical treatment response in myasthenia gravis. Neurol Clin. 2018;36(2):339–53.

4. Barohn RJ, McIntire D, Herbelin L, Wolfe GI, Nations S, Bryan WW. Reliability testing of the quantitative myasthenia gravis score a. Ann N Y Acad Sci. 1998;841(1):769–72.

5. Burns TM, Conaway M, Sanders DB, Composite MG, Group M-QS. The MG composite: a valid and reliable outcome measure for myasthenia gravis. Neurology. 2010;74(18):1434–40.

6. Muppidi S. The myasthenia gravis–specific activities of daily living profile. Ann N Y Acad Sci. 2012;1274:114–9.

7. Burns TM. The MG composite: an outcome measure for myasthenia gravis for use in clinical trials and everyday practice. Ann N Y Acad Sci. 2012;1274(1):99–106.

8. Howard JF Jr, Freimer M, O'Brien F, Wang JJ, Collins SR, Kissel JT, et al. QMG and MG-ADL correlations: study of eculizumab treatment of myasthenia gravis. Muscle Nerve. 2017;56(2):328–30.

9. McPherson T, Aban I, Duda PW, Farzaneh-Far R, Wolfe GI, Kaminski HJ, et al. Correlation of quantitative myasthenia gravis and myasthenia gravis activities of daily living scales in the MGTX study. Muscle Nerve. 2020;62(2):261–6.

10. Muppidi S, Wolfe GI, Conaway M, Burns TM, Mg C, Mg-Qol15 Study G. MG-ADL: still a relevant outcome measure. Muscle Nerve. 2011;44(5):727–31.

11. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1—eliciting concepts for a new PRO instrument. Value Health. 2011;14(8):967–77.

12. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—establishing and reporting the evidence in newly

developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2—assessing respondent understanding. Value Health. 2011;14(8):978–88.

13. Cleanthous S, Mork A-C, Regnault A, Cano S, Kaminski HJ, Morel T. Development of the Myasthenia Gravis (MG) Symptoms PRO: a case study of a patient-centred outcome measure in rare disease. Orphanet J Rare Dis. 2021;16(1):1–14.

14. Snyder C, Crossnohere N, King M, Reeve BB, Bottomley A, Calvert M, et al. The PROTEUS-trials consortium: optimizing the use of patient-reported outcomes in clinical trials. Clin Trials. 2022: 17407745221077691.

15. Weinfurt KP, Reeve BB. Patient-reported outcome measures in clinical research. JAMA. 2022;328(5): 472–3.

16. Brundage MD, Crossnohere NL, O'Donnell J, Cruz Rivera S, Wilson R, Wu AW, et al. Listening to the patient voice adds value to cancer clinical trials. JNCI: Journal of the National Cancer Institute. 2022.

17. Litchfield I, Greenfield S, Turner GM, Finnikin S, Calvert MJ. Implementing PROMs in routine clinical care: a qualitative exploration of GP perspectives. BJGP Open. 2021;5(1).

18. Campbell R, Ju A, King MT, Rutherford C. Perceived benefits and limitations of using patient-reported outcome measures in clinical practice with individual patients: a systematic review of qualitative studies. Qual Life Res. 2021:1–24.

19. Morel T, Cano SJ. Measuring what matters to rare disease patients–reflections on the work by the IRDiRC taskforce on patient-centered outcome measures. Orphanet J Rare Dis. 2017;12(1):1–13.

20. Alekseeva T, Gavrilov Y, Kreis O, Valko P, Weber K, Valko Y. Fatigue in patients with myasthenia gravis. J Neurol. 2018;265(10):2312–21.

21. Bril V, Benatar M, Andersen H, Vissing J, Brock M, Greve B, et al. Efficacy and safety of rozanolixizumab in moderate to severe generalized myasthenia gravis: a phase 2 randomized control trial. Neurology. 2021;96(6):e853–65.

22. Barnett C, Katzberg H, Nabavi M, Bril V. The quantitative myasthenia gravis score: comparison with clinical, electrophysiological, and laboratory markers. J Clin Neuromuscul Dis. 2012;13(4):201–5.

23. Burns TM, Conaway MR, Cutter GR, Sanders DB. Construction of an efficient evaluative instrument for myasthenia gravis: the MG composite. Muscle Nerve. 2008;38(6):1553–62.

24. Rasch G. Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests. 1960.

25. Andrich D. Rasch models for measurement. Newbury Park: Sage; 1988.

26. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. Health Technol Assess. 2009;13(12):iii, ix-x, 1–177.

27. Wright BD, Masters GN. Rating scale analysis. Mesa Press; 1982.

28. Andrich D. A rating formulation for ordered response categories. Psychometrika. 1978;43: 561–73.

29. Barnett C, Merkies IS, Katzberg H, Bril V. Psychometric properties of the quantitative myasthenia gravis score and the myasthenia gravis composite scale. J Neuromusc Dis. 2015;2(3):301–11.

30. Sadjadi R, Conaway M, Cutter G, Sanders DB, Burns TM, Group MCMQS. Psychometric evaluation of the myasthenia gravis composite using Rasch analysis. Muscle Nerve. 2012;45(6):820–5.

31. Serra A, Ruff R, Kaminski H, Leigh RJ. Factors contributing to failure of neuromuscular transmission in myasthenia gravis and the special case of the extraocular muscles. Ann N Y Acad Sci. 2011;1233(1):26–33.

32. Soltys J, Gong B, Kaminski HJ, Zhou Y, Kusner LL. Extraocular muscle susceptibility to myasthenia gravis unique immunological environment? Ann N Y Acad Sci. 2008;1132:220.

33. Andersen LK, Aadahl M, Vissing J. Fatigue, physical activity and associated factors in 779 patients with myasthenia gravis. Neuromuscul Disord. 2021;31(8):716–25.

34. Ruiter AM, Verschuuren JJ, Tannemaat MR. Fatigue in patients with myasthenia gravis. A systematic review of the literature. Neuromusc Disord. 2020;30(8):631–9.

35. Barnett C, Bril V, Kapral M, Kulkarni A, Davis AM. Development and validation of the myasthenia gravis impairment index. Neurology. 2016;87(9): 879–86.