Title

Addressing Outcome Measure Variability in Myasthenia Gravis Clinical Trials.

Journal

Neurology, 101(10)

ISSN

0028-3878

Authors

Guptill, Jeffrey T
Benatar, Michael
Granit, Volkan
et al.

Publication Date

2023-09-01

DOI

10.1212/wnl.0000000000207278

Addressing Outcome Measure Variability in Myasthenia Gravis Clinical Trials

Jeffrey T. Guptill, MD, Michael Benatar, MD, PhD, Volkan Granit, MD, Ali A. Habib, MD, James F. Howard, Jr., MD, Carolina Barnett-Tapia, MD, PhD, Richard J. Nowak, MD, Ikjae Lee, MD, Katherine Ruzhansky, MD, Mazen M. Dimachkie, MD, Gary R. Cutter, PhD, and Henry J. Kaminski, MD, for MGNet Clinical Trial Outcome Measure Working Group

Correspondence
Dr. Guptill
jeffrey.guptill@duke.edu
or Dr. Benatar
mbenatar@med.miami.edu

Neurology® 2023;101:442-451. doi:10.1212/WNL.0000000000207278

Abstract

An increasing number of clinical trials are enrolling patients with myasthenia gravis (MG). A lack of standardization in the performance of outcome measures leads to confusion among site research teams and is a source of variability in clinical trial data. MGNet, the NIH-supported Rare Disease Clinical Research Network for MG, views standardization of MG outcome measures as a critical need. To address this issue, a group of experts summarized key outcome measures used in MG clinical trials and a symposium was convened to address issues contributing to outcome measure variability. Consensus recommendations resulted in changes to outcome measure instructions and, in some cases, modifications to specific instruments. Recommended changes were posted for public commentary before finalization. Changes to the MG-Activities of Daily Living, MG–Quality of Life–15r, and MG–Impairment Index were limited to adding details to the administration instructions. Recommendations for proper positioning of participants and how to score items that could not be performed because of non-MG reasons were provided for the MG Composite. The Quantitative MG (QMG) score required the most attention, and changes were made both to the instructions and the performance of certain items resulting in the QMG-Revised. The Postintervention Status was believed to have a limited role in clinical trials, except for the concept of minimal manifestation status. As a next step, training materials and revised source documents, which will be freely available to study teams, will be created and posted on the MGNet website. Further studies are needed to validate changes made to the QMG-Revised.

From the Duke University School of Medicine (J.T.G.), Durham, NC; argenx US (J.T.G.), Boston, MA; University of Miami School of Medicine (M.B., V.G.), FL; Biohaven Pharmaceuticals (V.G.), New Haven, CT; University of California, Irvine (A.A.H.); The University of North Carolina School of Medicine (J.F.H.), Chapel Hill; Division of Neurology (C.B.-T.), Department of Medicine, University of Toronto, Ontario, Canada; Yale University School of Medicine (R.J.N.), New Haven, CT; Columbia University (I.L.), New York, NY; Medical University of South Carolina (K.R.), Charleston; Kansas University Medical Center (M.M.D.), Kansas City; School of Public Health (G.R.C.), University of Alabama at Birmingham; and George Washington University School of Medicine & Health Sciences (H.J.K.), DC.

Go to Neurology.org/N for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

The Article Processing Charge was funded by the authors.

Coinvestigators are listed at links.lww.com/WNL/C753.

Glossary

MG = myasthenia gravis; **MG-ADL** = MG-Activities of Daily Living Scale; **MGFA** = Myasthenia Gravis Foundation of America; **MGII** = MG Impairment Index; **MG-QOL-15** = MG–Quality of Life–15; **MG-QOL-15r** = MG-QOL-15 revised version; **MMS** = minimal manifestation status; **PIS** = Postintervention Status; **QMG** = quantitative MG; **RDCRN** = Rare Disease Clinical Research Network.

Several outcome measures are commonly used in myasthenia gravis (MG) clinical trials.[1] Prior task forces provided consensus recommendations on the use of outcome measures in MG clinical research.[2,3] For many years, the quantitative MG (QMG) score was the accepted key outcome measure for MG clinical trials. The phase 3 clinical trial of eculizumab in patients with refractory generalized MG with AChR antibodies marked a transition to a greater emphasis on patient-reported outcome measures in the field.[4] In this study, the MG–Activities of Daily Living Scale (MG-ADL) score was the primary outcome measure, and the MG-ADL has served as a primary or key secondary efficacy endpoint for several subsequent phase 2 and 3 studies.[5,6]

There has been a steadily increasing number of therapeutics under development for patients with MG. An observation by site investigators participating in clinical trials, and a source of frustration for sponsors and sites alike, is a lack of standardization in the training and performance of MG outcome measures. This lack of standardization is a source of variability in clinical trial data. This also leads to confusion among site research teams and complicates comparability of results across studies. Variability in outcome measure administration could also lead to trials measuring aspects of the disease differently and in a way that is not transparent in publications or trial reports.

The NIH-supported Rare Disease Clinical Research Network (RDCRN) for MG (MGNet) views standardization of MG outcome measures as a critical need for clinical trials and convened a group of experts to address the issue. After an outcome measure symposium, MG outcome measures were refined with the specific goal of improving the clarity of instructions and scoring, thereby improving the consistency of how outcome measures are performed and reducing the variability in outcome measure data.

Methods

We identified the most frequently used MG-specific outcome measures in clinical trials. These included the MG-ADL,[7] QMG Score,[8] Myasthenia Gravis Foundation of America (MGFA) Postintervention Status (PIS),[2] MG Composite,[9] MG–Quality of Life–15 (MG-QOL-15)[10] and MG-QOL-15 revised version (MG-QOL-15r),[11] and MG Impairment Index (MGII).[12] As an initial step, each outcome measure was summarized regarding administration, domains evaluated, psychometric properties, translations, and aspects contributing to a lack of standardization (Figure).

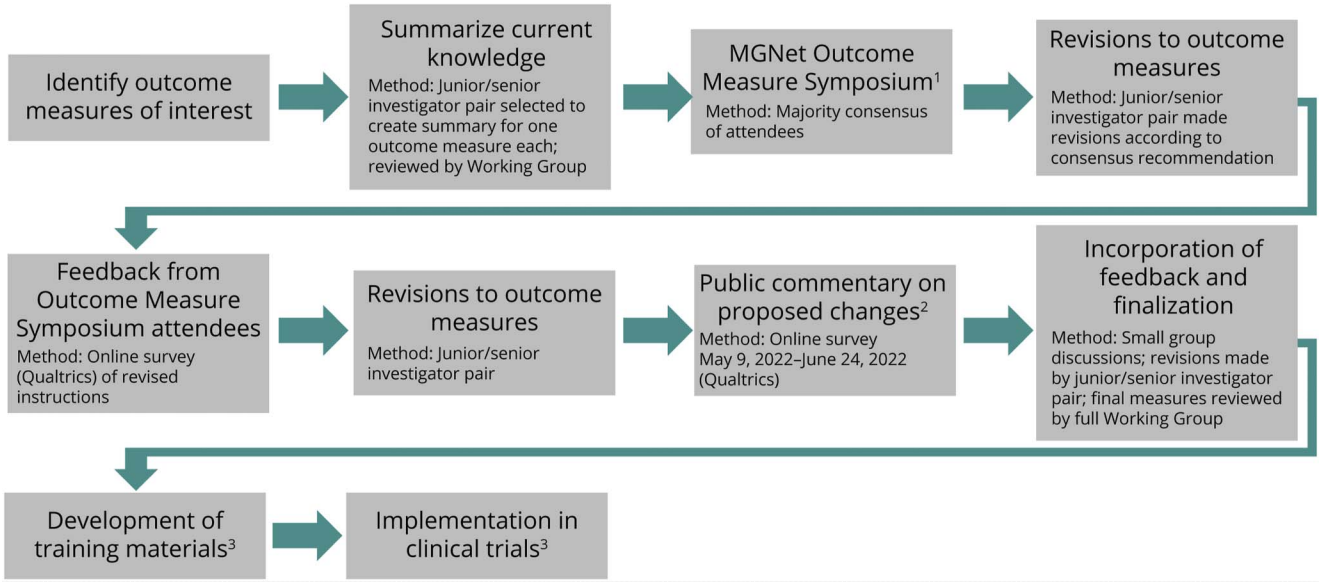
Outcome measure summary findings were presented at an MGNet Symposium (December 2020). Attendees included MGNet investigators, patient advocacy groups, patients, the NIH, and representatives from the RDCRN Data Management and Coordinating Center, and industry (eAppendix 1, links.lww.com/WNL/C754). These attendees were broadly inclusive of stakeholders involved in MG clinical trials, including thought leaders in the field with experience in designing clinical trials and those who have developed and extensively used MG outcome measures. Issues related to variability and standardization of each outcome measure in a clinical trial setting were discussed by attendees and informal consensus (majority agreement) was achieved for the best approach to standardize each outcome measure. Modifications were made to each outcome measure to align with the informal consensus achieved at the outcome measure symposium. The revised outcome measures and/or their instructions for administration were then reviewed by symposium attendees and approved. In cases where consensus was not reached at the symposium, options for how to revise a measure were presented to attendees and voted on separately. A simple majority determined how these areas were addressed.

The approved outcome measures were then posted on the MGNet website for a 4-week public comment period. Relevant stakeholders were informed of the opportunity for public commentary through the MGFA 14th International Conference on Myasthenia Gravis and Related Disorders,[13] posting on relevant professional communication platforms (e.g., American Association of Neuromuscular & Electrodiagnostic Medicine Connect and American Academy of Neurology Synapse) and direct email communication (Figure). Community neurologists with an interest in MG outcome measures would have had an opportunity to provide input at the MGFA Conference (if in attendance) and during the public commentary period. Comments from the public were reviewed and outcome measures were further modified, with additional input from MGNet Symposium attendees, experts in outcome measure development, and patients with MG. Patient-facing outcome measures were optimized for an eighth-grade reading level.[14]

Results

A summary of each outcome measure and available translations can be found in eAppendix 2 (links.lww.com/WNL/C755). There was consensus among symposium attendees that the general approach to modifying outcome measures was

Figure Overview of Methods for Standardizing MG Outcome Measures



[1]Refer to eAppendix 1 (links.lww.com/WNL/C754) for list of attendees. [2]Groups notified of public commentary: AANEM Connect, AAN Synapse, Alexion, Argenx, Biosensics, Cabaletta Bio, Clinical and Translational Science Award sites, Conquer MG, Horizon Therapeutics, Janssen, MGNet clinical sites, Muscular Dystrophy Association, Muscle Study Group, Myasthenia Gravis Foundation of America (posted on website), Rare Disease Clinical Research Networks (including NIH staff), Rick's Real Neuromuscular Friends, Signant, and UCB Pharma. [3]Next steps of process, currently pending. AAN = American Academy of Neurology; AANEM = American Association of Neuromuscular & Electrodiagnostic Medicine; MG = myasthenia gravis.

not to change the outcome measure itself, unless deemed absolutely necessary. No new items were to be added. The focus was on strengthening the outcome measure instructions/administration to enhance standardization and address situations that arise during clinical trials and are either not accounted for in current instructions or addressed in inconsistent ways by study sponsors.

One area of considerable discussion for patient-reported outcome measures (MG-ADL, MGII, and MG-QOL-15r) was whether they should capture only those symptoms/signs that are attributable to MG or to capture function/status "as is," recognizing the potential (indeed, likelihood) that comorbidities might introduce some confounding, given that patients may struggle to determine what is attributable to MG and what is not. Consensus was reached that patients should try to respond with symptoms related to MG, largely because prior validation studies had used this approach.[7,11] Future research, however, could explore whether patients should be instructed to answer questions "as they are" to minimize these potential confounding effects.

Other topics of emphasis with strong consensus included the following: (1) recording the time of day for assessments and maintaining consistent timing of assessments throughout a trial due to variability in the disease over the course of a day; (2) maintaining the same order of assessments and same raters throughout a study; (3) that MG trials should be as inclusive as possible and that patients with fixed deficits preventing completion of a specific item should be allowed to participate if an acceptable standardized method of handling these items can be determined; (4) the need to avoid missing data to the extent possible; (5) the need for trial statistical analysis plans to include instrument-specific approaches for handling missing data (e.g., how to handle a permanent injury that occurs during the course of a trial and prevents an assessment and for which recovery is not expected); (6) the importance of withholding pyridostigmine (or other cholinesterase inhibitor) for at least 12 hours for clinician-assessed outcome measures with the time, dose, and form (regular or long acting) of the last pyridostigmine dose clearly documented; and (7) the general principle that the primary outcome measure should be completed first at a study visit; it is recommended that instruments with muscle testing that can cause fatigue should also be performed early after arrival to study site.

Additional outcome measure–specific summaries and recommendations are discussed further below. The revised instructions for each outcome measure, which incorporate the changes recommended during the process described in the Methods, are found in eAppendix 3 (links.lww.com/WNL/C756). The revised outcome measures are also available on the MGNet website (mgnet.rarediseasesnetwork.org/resources/researchers-clinicians), which will remain the best source for the most up-to-date instructions.

MG-ADL

The MG-ADL is an 8-item patient-reported outcome measure that assesses MG-specific symptoms and their impact on

Table 1 Challenges and Key Recommendations for Patient-Centered MG Outcome Measures (MG-ADL and MG-QOL-15r)

Challenge	Recommendation
MG-ADL	
Uncertainty about acceptable input and help from the study team or family members	• There was consensus, including favorable feedback during public commentary, that MG-ADL should be self-reported. This will ensure that the responses reflect the patient's experience without interpretation from others. This is a departure from the original instrument, but this approach has been validated.[17] • If a patient asks for clarification about a question, the study team should simply remind them to answer questions based on the average severity of their MG-related symptoms. The study team should not help the patient to tease apart relative contributions from MG and non-MG causes. • If the patient makes a mark on the boundary line between 2 scores or marks 2 scores, the patient should be asked to clearly mark 1 score only. If the patient cannot decide between 2 scores, record the higher score.
Inconsistent duration of recall period	The symptom recall period should be standardized to the prior 7 d because this has worked well in prior trials; is a more reasonable time frame over which function can readily be recalled; and enables assessment of therapeutic response over a short period of time for rapid-acting treatments. Moreover, several other validated and commonly used patient-reported scales (e.g., PROMIS, ASCQ, and various Neuro-QoL domains) use a 7-d recall period.
Lack of consistent instructions on self-administration of MG-ADL	• Patients should be "trained" on self-administration during the first study visit according to instructions developed by MGNet. • Responses should reflect "average functioning" over the recall period. The recommendation for responding with "average function" was supported by input from patients who believed the instruction was clear, resonated best with what was important to them, and was less likely to be biased by their clinical status at a single moment. • If the patient is in doubt about how to answer an item, they should be instructed to choose the option that is most appropriate most of the time over the past 7 d • Patients should be instructed to not leave any items blank (if an item is left blank, the patient should be asked to complete the missing item)
MG-QOL-15r	
Uncertainty about acceptable input from the study team or family members	• There was consensus that the MG-QOL-15r should be self-reported without input from the study team or family members. • Instructions should be provided for self-completion of the MG-QOL-15r (eAppendix 3, links.lww.com/WNL/C756).
Unclear instructions regarding the recall period	• The symptom recall period should be standardized to "the past 4 wks." There was strong feedback during public commentary that the existing wording ("over the past few weeks")[26] was too vague and interpreted variably by patients. The recommended wording is supported by its use in the original validation of the instrument.[10] • There was also consensus that the MG-QOL-15r should not be administered too frequently (i.e., more frequently than every 4 wks).
Unclear instructions about scoring an item when the subject cannot decide between 2 grades	There was some discordance about whether to have an instruction for patients to select the higher (more severe) score if they were undecided about a particular item. In this situation, public commentary supported the recommendation that participants should be instructed to "choose the option that is most appropriate for you most of the time over the past 4 wks."

Abbreviations: ASCQ = Adult Sickle Cell Quality of Life; MG = myasthenia gravis; MG-ADL = MG-Activities of Daily Living Scale; MG-QOL-15 = MG–Quality of Life–15; MG-QOL-15r = MG-QOL-15 revised version; PROMIS = Patient-Reported Outcome Measurement Information System; QoL = quality of life.

daily activities.[7] The MG-ADL is a common primary endpoint in recent clinical trials[4,15] and may be more sensitive to clinical change than the QMG.[16] In clinical trials to date, the MG-ADL has been administered with varying degrees of instruction and guidance, or patients complete the instrument without any study team interaction. Other areas contributing to a lack of standardization include the following: a variable time frame for recall of symptoms (e.g., 7 vs 14 days), inconsistent instruction as to whether only MG symptoms should be considered in their responses, and a lack of clarity about how patients summarize their function over the specified time frame (e.g., do they consider "worst" or "average" function). While there is no evidence to support one time frame vs another, we recommend consistency across all studies. Key recommendations are summarized in Table 1.

MG-QOL-15r

The MG-QOL-15r is a 15-item patient-reported outcome measure assessing physical, psychological, and social domains commonly affected by MG.[10] A revised version, which reduces the number of responses for each item from 4 to 3, has been validated and is commonly used.[11] Areas of inconsistency in the administration of the MG-QOL-15r include differences in recall time (past week vs 2 weeks, etc), self-administration or administration by study team, and whether to include an instruction for individuals having difficulty deciding between 2 scores on an item to score higher. Key recommendations are summarized in Table 1.

Table 2 Challenges and Key Recommendations for the QMG

Challenge	Recommendation
Lack of clarity about pyridostigmine timing	Pyridostigmine and extended-release formulations should be held for at least 12 and 24 h before the QMG, respectively
Missing raw data	For all measurements, record the raw data (e.g., time) and the grade
Lack of standardized patient instructions	A script has been developed for all items to standardize instructions for the participant (eAppendix 3, links.lww.com/WNL/C756)
Lack of clarity about "coaching" the patient	For items that are particularly effort dependent (e.g., arm and leg outstretched items), participants should be given encouragement to elicit their best performance.
Insufficient instructions to assess ptosis and diplopia	• Flow charts for the assessment of diplopia and ptosis have been developed. Diplopia and ptosis should be assessed on the left and right at each visit, and the worst side should be scored at each visit. Most trials previously assessed diplopia on both sides only at the initial visit, and the side that was most effected at that visit was measured at every visit thereafter. This change acknowledges the fluctuating nature of the disease where diplopia may shift over time. • Diplopia should be scored only if the participant has binocular "double" vision. Blurry vision should not be scored. • A ptosis scoring approach based on a clock face has been developed to make the assessment more objective and provide source documentation.
No standardized speed for single breath count.	An audible metronome/app should be used to standardize a counting rate of 1/s during the speech assessment.
Arm, leg, and neck fatigability tests are insufficiently standardized	• A goniometer should be used to ascertain proper limb position before starting the arm and leg outstretched test. • For arm fatigability, an approach for measuring arm droop from 90° has been developed. If the arm drops 10° or more, the test should be stopped. • If 1 arm or leg cannot be assessed because of a non-MG related problem, the affected limb should not be scored; instead, the score from the unaffected limb should be used to impute the grade for the affected limb. This approach is supported by excellent observed correlations between right and left arm/leg in prior studies.[27] • The head lift assessment has been modified to improve standardization and account for limitations in neck mobility that are present in some participants. • Similar to the arm test, if the leg drops ≥10°, the test should be stopped. Ideally, the test is performed with the bed next to a wall with a mark at the point of a 10° excursion from the initial 45° position. Recognizing that all clinics may not be able to accommodate this setup, a 3-inch (approximately 8 cm) downward excursion is also acceptable and should be used throughout the trial.
Lack of clarity about the number of FVC attempts, normative data, and common confounding situations	• The same spirometer and mouthpiece should be used by all sites. In general, oval mouthpieces are easier for patients. Study planning should address how spirometry will be performed for patients with lower facial weakness that prevents a tight lip seal around the mouthpiece. Potential options include a face mask or a mouthpiece with a flange that minimizes air leak around the mouthpiece. However, neither option will solve the issue for all patients. • Three trials for FVC are recommended (up to 5 trials if quality issues), and the best value should be scored. Instructions for assessing the quality of FVC measurements were added. There may be value in updating to use SVC, as is being performed in the field of ALS,[28] and to using more current norms (e.g., NHANES III),[29] but this decision should be made at trial outset, and the same normative data should be maintained throughout the study at all sites. The best normative data may change over time (the same may be true for grip strength[27,30]), and to avoid the need to update the QMG instructions with specific norms, no specific normative data are recommended. If source data with raw values are maintained, it would be possible to compare trial data based on other normative values.
No standard rest and testing sequence for handgrip.	Three handgrip trials should be performed for each hand, alternating side to side. Rest 1 min between trials, and the best value should be used. If a patient cannot perform the grip assessment due to a reason other than weakness (e.g., injury), use the score from the unaffected arm to grade the affected arm.
The original instructions assumed the right hand is dominant in all patients	Grip strength scoring should incorporate dominant and nondominant hands rather than right/left hand.

Abbreviations: ALS = amyotrophic lateral sclerosis; FVC = forced vital capacity; MG = myasthenia gravis; NHANES III = Third National Health and Nutrition Examination Survey; QMG = quantitative MG; SVC = slow vital capacity.

QMG

The QMG is a clinician-administered assessment of strength and fatigable weakness in the domains of ocular, bulbar, respiratory, and limb/axial muscles. It requires special equipment and takes approximately 30 minutes to complete.[8,17] Of all the outcome measures addressed, the QMG had the most concerns related to variability in performance (Table 2). Challenging issues included positioning of

Table 3 Challenges and Key Recommendations for Hybrid MG Outcome Measures (MG Composite and MGII)

Challenge	Recommendation
MG composite	
Variable practices in handling overlapping items between MG-ADL and QMG	Overlapping items from the MG-ADL and QMG should be administered in a standardized way with the MG Composite (i.e., same instructions). If these outcome measures are performed on the same day as the MG Composite, the items do not need to be repeated; responses may be transferred from the MG-ADL and times from the QMG assessments transcribed to determine scoring for the MG Composite. Instructions for the MG Composite were harmonized with the MG-ADL and QMG.
Uncertainty about positioning during the assessment of certain muscles	Hip flexion should be tested in the supine position with the knee flexed to better isolate the hip flexors. Neck flexion and extension should be tested in the supine and prone positions, respectively. The position should be captured with an explanation if a patient cannot perform an assessment in the recommended position (e.g., patient unable to lay flat due to dyspnea).
Uncertainty as to whether raters need to be neurologists	The MG Composite may be performed by any trained and certified evaluator to allow the most flexibility for study teams.
The interpretation of weakness grading is highly variable	There was an extensive discussion about how to translate manual muscle testing to score normal/mild/moderate/severe weakness and the extent to which these grades should be tied to MRC scoring. Consensus was reached to use modified MRC scoring to grade weakness,[31] largely because of clinician familiarity with this system. Given concerns about the floor effect of MRC grading and that mild weakness is functionally limiting for patients, the recommended strength scoring is as follows: Mild weakness = MRC grade 4+ Moderate weakness = MRC grade 4 to 4− Severe weakness = MRC grade 3 or less
MGII	
Uncertainty about administering the MGII remotely	The MGII has since been validated for use in telephone visits.[32]
The modifications in the QMG instructions may also affect MGII	Changes to the QMG administration instructions for clinician-assessed items carry over to the MGII, and more detailed instructions for the MGII have been developed.
Uncertainty about acceptable input from the study team or family members	The self-reported items of the MGII are intended to be patient reported. The participant should be given the paper or electronic form with the standard instructions and then be left to complete using their best judgment. This is consistent with the consensus approach for the MG-ADL and MG-QOL-15r.
Questions about the recall period in light of the changes in MG-ADL and MG-QOL-15r	The time frame of the recall period for patient-reported items should be 2 wks; this differs from the 7-d period recommended for the MG-ADL because this is how the MGII was validated.

Abbreviations: MG = myasthenia gravis; MG-ADL = MG-Activities of Daily Living Scale; MGII = MG Impairment Index; MG-QOL15 = MG–Quality of Life–15; MG-QOL-15r = MG-QOL-15 revised version; MRC = Medical Research Council; QMG = quantitative MG.

individuals and standardizing instructions (particularly ocular items), how to score MG-related weakness for several items in specific situations, concerns about the number of trials to perform and the normative data used to assess forced vital capacity, and how to score items that cannot be completed because of non-MG factors. It was recognized that patient factors can sometimes limit performance of certain items, including factors that limit performance of items tested bilaterally (e.g., shoulder and leg items limited by cervical/lumbar spine pain). Because there is no simple solution for this, it is recommended that the patient perform to the best of their ability on that day. Recommended changes to the evaluation of ptosis were extensive. The changes to the QMG were sufficiently extensive to justify indicating "revised" in the name of the instrument (QMG-Revised) to avoid confusion. The case report form was updated to include the timing of last cholinesterase inhibitor, handedness of the participant, signature of the evaluator, and the specific cause if an arm or a leg was not tested because of a non-MG–related cause. Key recommendations are summarized in Table 2.

MG Composite

The MG Composite assesses disease-specific symptoms and examination findings derived from the MG-ADL (patient reported), QMG, and MG-Manual Muscle Test (clinician determined). It consists of 10 weighted items and takes approximately 5 minutes to complete.[9] Areas of uncertainty for the MG Composite included proper positioning of individuals for clinician-assessed items and how to grade weakness. Additional areas of variability include the following: whether overlapping items that are assessed in other outcome measures and measured on the same day need to be repeated for the MG Composite (e.g., diplopia in the QMG); whether raters need to be neurologists; and how to score items that cannot be performed for MG or non-MG reasons. Key recommendations are summarized in Table 3.

MG Impairment Index

The MGII assesses MG-specific impairments through patient self-report (22 items) and clinical examination (6 items).[12] It assesses ocular, bulbar, respiratory, and limb domains and takes

approximately 10 minutes to complete. Patients are instructed to consider only symptoms related to MG, and certain clinician-evaluated items (arm endurance, leg endurance, and neck endurance) follow the same instructions as the QMG, including patient positioning. However, scoring of items performed in the QMG and MGII differ. The MGII has been administered in the clinic, and during the symposium, it was uncertain whether it would be suitable to telemedicine assessments. In many cases, the QMG is administered at the same visit as the MGII; in light of the changes to the QMG, the instructions for carrying over QMG scores to the MGII needed to be addressed. Key recommendations are summarized in Table 3.

MGFA PIS

The MGFA PIS is a clinician-assessed instrument developed to measure the effects of a therapeutic intervention on disease status.[2] It has 8 major categories and can be used both in clinical trials and in the clinic. Major considerations regarding the PIS were lack of definition for criteria when defining improvement or worsening (i.e., change in QMG or MG Composite score vs overall impression), its relevance for clinical trials, and a lack of category standardization across use in trials. The anchor time point for assessing categories (e.g., change from initial visit vs last visit vs worst ever) has also been variably assessed.

Key Recommendations for MGFA PIS

1. In the setting of interventional clinical trials, the full MGFA PIS is not recommended. Several categories defined in the PIS, such as Pharmacologic Remission that requires that a "patient has had no symptoms or signs of MG for at least 1 year,"[2] are not relevant to most interventional studies that evaluate treatment effect over a shorter time frame. In addition, improved/worse status are usually redundant with other analyses performed on quantitative measures (e.g., change in QMG score). Of importance, there was clear consensus on the value of the PIS in other clinical research settings that were beyond the scope of these recommendations focused on outcome measures for clinical trials.
2. An alternative approach to the PIS categories of improved/worse is a clinician and patient global impression of change score. A 7-point Likert scale for clinician-reported global assessment of disease severity or change has been included in several trials to date and achieved consensus.[18] Validated patient-reported global assessments, such as the Single Simple Question (which has been studied in MG[19]), might also be used to supplement physician-reported global assessments, but further research would be needed before doing so.
3. Minimal manifestation status (MMS) is widely accepted as a critical treatment goal for patients with MG[20] and remains an important concept that is worth retaining for clinical trials. In addition to measuring MMS as an outcome, MMS can also serve as a guide for steroid tapering in clinical trials.[15,21] It was noted during public commentary that the definition of MMS is a source of confusion. The original definition of MMS is "the patient has no symptoms or functional limitations from MG but has some weakness on examination…" for a duration of at least 1 year.[2] In clinical trials using MMS, the time requirement has been dropped because trial durations are often less than 1 year. There was also significant disagreement about whether to preserve, discard, or modify the existing subgrades of MMS. It was noted that, outside of open-label extension studies, baseline treatments are usually not altered. Thus, the important concept is the clinical status, and the subcategories are less important and not recommended for use in clinical trials. Of note, a separate task force is currently addressing this issue for the clinic, and further recommendations for clinical use may be forthcoming.[22]
4. The anchor time point for assessing any category of the PIS in a clinical trial is the last assessment before initiation of the experimental therapeutic (typically the baseline or randomization visit).

Discussion

An increased number of therapeutics under development for MG in recent years has led to important observations about outcome measure training and performance in the clinical trial setting. The MGNet Clinical Trial Outcome Measure Working Group was convened to synthesize our collective experience and to use the cumulative expertise of the group to make recommendations that apply "lessons learned." Of importance, the Working Group focused on issues related to standardization of MG-specific outcome measures specifically for clinical trials. The Working Group did not specify which outcome measures should be used or is "best"; this decision needs to be made independently for each trial depending on the goals and often with input from regulatory authorities. The scope did not include the use of these outcome measures in the clinic or other clinical research settings. The MGFA Clinical Classification was also considered outside the scope because it is not commonly used as a clinical trial outcome measure. In addition, we have not addressed the potential need for new outcome measures.

Improving standardization of MG outcome measures will have many benefits for future clinical trials. First, training will be consistent across trials and MGNet will develop a standardized set of training tools that will be freely available for future use. This should reduce costs previously borne by individual sponsors to develop their own training materials and training plans and potentially yield faster startup times. For example, MG outcome measure training completed for 1 trial should be valid for another trial within a specified time frame to reduce redundancy. To date, detailed training materials for MG outcome measures have been maintained by study sponsors and are not widely available. Because MGNet outcome measure instructions, case report forms, and training materials will be freely available, the enhanced accessibility should reduce this barrier to clinical trial implementation, potentially increasing the number of trained sites and facilitating participation among sites that have not traditionally been involved in MG clinical trials. This is an important need, given the competition for patients among a growing number of

clinical trials in this rare disease and the potential for inexperienced sites to have more variability in their data.

Ultimately, implementing these recommendations is expected to lead to less noisy data and fewer errors at sites. This standardization should have benefits for study design, such as the potential for sample size reductions (e.g., smaller standard deviations), although this would need to be proven in future studies. Finally, standardization could also permit greater comparison across studies and, hopefully, eventually pooling of data in a repository akin to Pooled Resource Open-Access ALS Clinical Trials for amyotrophic lateral sclerosis.[23] Implementation of these standardization recommendations may make comparisons with trials completed under historical outcome measure approaches more challenging.

Increasing outcome measure standardization necessitated adding clarifying language to the instructions. The Working Group was challenged with finding the right balance between providing appropriate guidance without being excessive. This was a particular concern for patient-reported outcomes where there is a risk that patients may not read the instructions if they are too long or they could become excessively burdensome. Specifically for the MG-ADL, it was decided to separate each item into its own block that contains the item followed by any instructions.

The public commentary period revealed several important themes. Several comments suggested adding or extensively modifying existing questions. Examples included the following: (1) adding items to assess vision in primary or downgaze; (2) developing a more patient-centric presentation of questions and responses for the MG-ADL; (3) shifting the focus of the MG-ADL from choking specifically to a more general assessment of swallowing; (4) increasing the focus of outcome measures on ocular symptom impact; and (5) assessing anxiety and mood specifically related to living with MG and/or an MG exacerbation. These comments suggest that there are residual issues with the current outcome measures used in clinical trials that cannot be easily resolved. Prior studies have highlighted other limitations with outcome measures currently in use.[1,6,24] More holistic assessments may be needed that measure aspects of the disease not currently captured and that are developed with extensive patient input.[25]

A potential limitation of our approach was the lack of a formal consensus process. We did, however, take several steps to ensure the rigor of our process, which included multiple and overlapping opportunities for broad stakeholder input, both in the presence of peers (e.g., during symposium) and anonymously (e.g., online surveys). The incorporation of a public commentary provided an additional opportunity for broad feedback from a diverse set of stakeholders.

The next steps for the MGNet Clinical Trial Outcome Measure Working Group include the development of a comprehensive set of training materials and case report forms for each outcome measure that reflects the recommended changes. MGNet anticipates holding training sessions for study teams at future meetings and offering initial and renewal certifications for clinical trial raters. Wide dissemination of this information to study teams is expected to improve efficiency in MG clinical trial outcome measure training. Unlike most of the recommendations that were limited to additional instructions, rather than changes to the outcome measure itself, there were extensive changes to the QMG instrument. Thus, a study is needed to validate the impact of these changes on the performance of the QMG-Revised. In addition, there will be a need for official translations (including assessment of local dialects and cultural adaptations) and, in some cases, validation of these instructions for use in other countries. MGNet is willing and able to centrally host official translations that can then be made available to study teams.

Acknowledgment

The Working Group thanks the peer reviewers, whose thorough and thoughtful review increased the quality of this manuscript. The Working Group also thanks Helen Girma for administrative support and the Myasthenia Gravis Foundation of America for supporting this program.

Study Funding

The work was supported by MGNet (NIH U54 NS115054), a member of the Rare Disease Clinical Research Network Consortia (RDCRN), and the Data Management and Coordinating Center (NIH U2CTR002818). Funding support for the RDCRN is provided by the National Center for Advancing Translational Sciences (NCATS) and the National Institute of Neurological Disorders and Stroke.

Disclosure

J.T. Guptill has consulted for Immunovant, Alexion, Apellis, Momenta, Ra Pharma, Becton Dickinson, Cabaletta Bio, Regeneron, argenx, Sanofi, Janssen, UCB, Toleranzia, and Piedmont Pharmaceuticals. He received industry grant support from UCB pharma for a fellowship training grant. He has served as a site investigator for Alexion, Janssen, UCB Pharma, Argenx, and Takeda. He received grant/research support from NIH (NIAID, National Institute of Neurological Disorders and Stroke, NIMH), Myasthenia Gravis Foundation of America, and Centers for Disease Control and Prevention. He is currently an employee of argenx. M. Benatar has consulted for Alexion, Immunovant, Takeda, UCB, Ad Scientam, and Sanofi. He receives research funding from Alexion and Immunovant. He has served as the site principal investigator for MG trials sponsored by Alexion, UCB, and the NIH. V. Granit received honoraria as a consultant or advisory board member from Alexion Pharmaceuticals, Argenx, Immunovant Inc., and Amylyx Pharmaceuticals Inc. He is employed by Biohaven Pharmaceuticals. A.A. Habib reports no relevant disclosures. J.F. Howard reports research support (paid to his institution) from Alexion Pharmaceuticals, Argenx, Cartesian Therapeutics, the Centers for Disease Control and Prevention (Atlanta, GA), the Myasthenia Gravis Foundation of America, the Muscular Dystrophy Association, the NIH (including the National Institute of Neurological Disorders and Stroke and the National Institute of

Arthritis and Musculoskeletal and Skin Diseases), PCORI, Ra Pharmaceuticals (now UCB Biosciences), and Millennium Pharmaceuticals/Takeda Pharmaceuticals; honoraria from Alexion Pharmaceuticals, Argenx, F. Hoffman-LaRoche Ltd., Immunovant Inc., Ra Pharmaceuticals (now UCB Biosciences), Regeneron Pharmaceuticals, and Sanofi US and nonfinancial support from Alexion Pharmaceuticals, Argenx, Ra Pharmaceuticals (now UCB Biosciences), and Toleranzia AB. C. Barnett-Tapia has received honoraria as consultant or member of advisory board from Alexion, Sanofi, and Argenx. She is the primary developer of the MGII and may receive royalties for its use. R.J. Nowak reports no relevant disclosures. I. Lee reports no relevant disclosures. K. Ruzhansky has served on advisory boards for Alexion, Argenx, Immunovant, and UCB/Ra, has served as a site PI for Alexion, Argenx, UCB, and Janssen, and has grant funding from MGFA. M.M. Dimachkie serves or recently served as a consultant for Abcuro, Amazentis, ArgenX, Astellas, Catalyst, Cello, Covance/Labcorp, CSL-Behring, EcoR1, Janssen, Kezar, MDA, Medlink, Momenta, NuFactor, Octapharma, Priovant, Ra Pharma/UCB, Roivant Sciences Inc, Sanofi Genzyme, Shire Takeda, Scholar Rock, Spark Therapeutics, Abata/Third Rock, UCB Biopharma, and UpToDate. He received research grants or contracts or educational grants from Alexion, Alnylam Pharmaceuticals, Amicus, Biomarin, Bristol-Myers Squibb, Catalyst, Corbus, CSL-Behring, FDA/OOPD, GlaxoSmithKline, Genentech, Grifols, Kezar, Mitsubishi Tanabe Pharma, MDA, NIH, Novartis, Octapharma, Orphazyme, Ra Pharma/UCB, Sanofi Genzyme, Sarepta Therapeutics, Shire Takeda, Spark Therapeutics, The Myositis Association, UCB Biopharma/RaPharma, Viromed/Healixmith, & TMA. G.R. Cutter reports no relevant disclosures. H.J. Kaminski is a consultant for Roche, Cabaletta Bio, and UCB Pharmaceuticals and is a CEO and CMO of ARC Biotechnology, LLC, based on US Patent 8,961,98. He is the principal investigator of the Rare Disease Network for Myasthenia Gravis (MGNet) National Institute of Neurological Disorders and Stroke, U54 NS115054, Targeted Therapy for Myasthenia Gravis. R41 NS110331 to ARC Biotechnology, and coinvestigator for R43NS124329 MV2C2 antibody as a new therapeutic for myasthenia gravis to Mimivax, LLC. Go to Neurology.org/N for full disclosures.

Publication History

Received by *Neurology* September 7, 2022. Accepted in final form February 23, 2023. Submitted and externally peer reviewed. The handling editor was Associate Editor Anthony Amato, MD, FAAN.

Appendix 1 Authors

Name	Location	Contribution
Jeffrey T. Guptill, MD	Duke University School of Medicine, Durham, NC; argenx US, Boston, MA	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data
Michael Benatar, MD, PhD	University of Miami School of Medicine, FL	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data
Volkan Granit, MD	University of Miami School of Medicine, FL; Biohaven Pharmaceuticals, New Haven, CT	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data
Ali A. Habib, MD	University of California, Irvine	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; analysis or interpretation of data
James F. Howard, Jr., MD	The University of North Carolina School of Medicine, Chapel Hill	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data
Carolina Barnett-Tapia, MD, PhD	Division of Neurology, Department of Medicine, University of Toronto, Ontario, Canada	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; and analysis or interpretation of data
Richard J. Nowak, MD	Yale University School of Medicine, New Haven, CT	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; and analysis or interpretation of data
Ikjae Lee, MD	Columbia University, New York, NY	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; and analysis or interpretation of data
Katherine Ruzhansky, MD	Medical University of South Carolina, Charleston	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; and analysis or interpretation of data
Mazen M. Dimachkie, MD	Kansas University Medical Center, Kansas City	Drafting/revision of the article for content, including medical writing for content; analysis or interpretation of data
Gary R. Cutter, PhD	University of Alabama at Birmingham School of Public Health	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; and analysis or interpretation of data

Appendix 1 *(continued)*

Name	Location	Contribution
Henry J. Kaminski, MD	George Washington University School of Medicine & Health Sciences, DC	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data

Appendix 2 Coinvestigators

Coinvestigators are listed at links.lww.com/WNL/C753.

References

1. Thomsen JLS, Andersen H. Outcome measures in clinical trials of patients with myasthenia gravis. *Front Neurol*. 2020;11:596382.
2. Jaretzki A III, Barohn RJ, Ernstoff RM, et al. Myasthenia gravis: recommendations for clinical research standards. Task Force of the Medical Scientific Advisory Board of the Myasthenia Gravis Foundation of America. *Neurology*. 2000;55(1):16-23.
3. Benatar M, Sanders DB, Burns TM, et al. Recommendations for myasthenia gravis clinical trials. *Muscle Nerve*. 2012;45(6):909-917.
4. Howard JF Jr, Utsugisawa K, Benatar M, et al. Safety and efficacy of eculizumab in anti-acetylcholine receptor antibody-positive refractory generalised myasthenia gravis (REGAIN): a phase 3, randomised, double-blind, placebo-controlled, multicentre study. *Lancet Neurol*. 2017;16(12):976-986.
5. Howard JF Jr, Bril V, Vu T, et al. Safety, efficacy, and tolerability of efgartigimod in patients with generalised myasthenia gravis (ADAPT): a multicentre, randomised, placebo-controlled, phase 3 trial. *Lancet Neurol*. 2021;20(7):526-536.
6. Muppidi S, Silvestri NJ, Tan R, Riggs K, Leighton T, Phillips GA. Utilization of MG-ADL in myasthenia gravis clinical research and care. *Muscle Nerve*. 2022;65(6):630-639.
7. Wolfe GI, Herbelin L, Nations SP, Foster B, Bryan WW, Barohn RJ. Myasthenia gravis activities of daily living profile. *Neurology*. 1999;52(7):1487-1489.
8. Barohn RJ, McIntire D, Herbelin L, Wolfe GI, Nations S, Bryan WW. Reliability testing of the quantitative myasthenia gravis scorea. *Ann NY Acad Sci*. 1998;841:769-772.
9. 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):1553-1562.
10. Burns TM, Conaway MR, Cutter GR, Sanders DB, Muscle Study Group. Less is more, or almost as much: a 15-item quality-of-life instrument for myasthenia gravis. *Muscle Nerve*. 2008;38(2):957-963.
11. Burns TM, Sadjadi R, Utsugisawa K, et al. International clinimetric evaluation of the MG-QOL15, resulting in slight revision and subsequent validation of the MG-QOL15r. *Muscle Nerve*. 2016;54(6):1015-1022.
12. Barnett C, Bril V, Kapral M, Kulkarni A, Davis AM. Development and validation of the Myasthenia Gravis Impairment Index. *Neurology*. 2016;87(9):879-886.
13. Guptill JT, Granit V, Habib A, et al. Addressing outcome measure variability in myasthenia gravis clinical trials. *Muscle Nerve*. 2022;65:S11-S12.
14. Good Calculators. *Flesch Kincaid Calculator [online]*. Accessed August 28, 2022. goodcalculators.com/flesch-kincaid-calculator/.
15. Wolfe GI, Kaminski HJ, Aban IB, et al. Randomized trial of thymectomy in myasthenia gravis. *N Engl J Med*. 2016;375(6):511-522.
16. Howard JF Jr, Freimer M, O'Brien F, Wang JJ, Collins SR, Kissel JT. QMG and MG-ADL correlations: study of eculizumab treatment of myasthenia gravis. *Muscle Nerve*. 2017;56(2):328-330.
17. Bedlack RS, Simel DL, Bosworth H, Samsa G, Tucker-Lipscomb B, Sanders DB. Quantitative myasthenia gravis score: assessment of responsiveness and longitudinal validity. *Neurology*. 2005;64(11):1968-1970.
18. Busner J, Targum SD. The clinical global impressions scale: applying a research tool in clinical practice. *Psychiatry (Edgmont)*. 2007;4(7):28-37.
19. Abraham A, Breiner A, Barnett C, Katzberg HD, Bril V. The utility of a single simple question in the evaluation of patients with myasthenia gravis. *Muscle Nerve*. 2018;57(2):240-244.
20. Sanders DB, Wolfe GI, Benatar M, et al. International consensus guidance for management of myasthenia gravis: executive summary. *Neurology*. 2016;87(4):419-425.
21. Lee I, Kuo HC, Aban IB, et al. Minimal manifestation status and prednisone withdrawal in the MGTX trial. *Neurology*. 2020;95(6):e755-e766.
22. Ruzhansky K, Li Y, Wolfe G, et al. MGFA task force for standardization of myasthenia gravis outcome measures in clinical practice. *Muscle Nerve*. 2022;65:S15.
23. Atassi N, Berry J, Shui A, et al. The PRO-ACT database: design, initial analyses, and predictive features. *Neurology*. 2014;83(19):1719-1725.
24. McPherson T, Aban I, Duda PW, et al. Correlation of quantitative myasthenia gravis and myasthenia gravis activities of daily living scales in the MGTX study. *Muscle Nerve*. 2020;62(2):261-266.
25. Cleanthous S, Mork AC, 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):457.
26. Burns TM, Grouse CK, Conaway MR, Sanders DB; Mg composite and Mg-qol15 study group. Construct and concurrent validation of the MG-QOL15 in the practice setting. *Muscle Nerve*. 2010;41(2):219-226.
27. Barnett TC, Bril V, Davis AM. Performance of individual items of the quantitative myasthenia gravis score. *Neuromuscul Disord*. 2013;23(5):413-417.
28. Lechtzin N, Cudkowicz ME, de Carvalho M, et al. Respiratory measures in amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Frontotemporal Degener*. 2018;19(5-6):321-330.
29. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med*. 1999;159(1):179-187.
30. Peters MJ, van Nes SI, Vanhoutte EK, et al. Revised normative values for grip strength with the Jamar dynamometer. *J Peripher Nerv Syst*. 2011;16(1):47-50.
31. Medical Research Council. *Aids to the Examination of the Peripheral Nervous System*. The White Rose Press; 1976. Memordanum No. 45.
32. Menon D, Alnajjar S, Barnett C, et al. Telephone consultation for myasthenia gravis care during the COVID-19 pandemic: assessment of a novel virtual myasthenia gravis index. *Muscle Nerve*. 2021;63(6):831-836.